



**FACULTAT DE FARMÀCIA
UNIVERSITAT DE BARCELONA**

**CARACTERITZACIÓ GENÈTICA I FUNCIONAL DE L'OPERÓ
glc, IMPLICAT EN EL METABOLISME DE GLICOLAT I
GLIOXILAT A *Escherichia coli***

MARIA TERESA PELLICER MOYA

Recollir els pellets, rentar-los amb cura amb etanol 70%, assecar-los i ressuspendre'ls amb 8 μ l de tampó de càrrega. Assegurar-se que el pellet està ben ressupès. Guardar les mostres a -20°C fins la seva electroforesi.

DNA binding buffer

Tris-HCl, pH 8.0	50 mM
KCl	100 mM
MgCl ₂	12.5 mM
EDTA	1 mM
Glicerol	20% (v/v)
DTT	1 mM

Solució de Ca²⁺ i Mg²⁺

CaCl ₂	5 mM
MgCl ₂	10 mM

Solució de parada

NaCl	200 mM
EDTA	30 mM
SDS	1%
Yeast tRNA	100 μ g/ml

Tampó de mostra

Blau de bromofenol	0.1%
Xylen cianol	0.1%
NaOH 0.1 M:formamida	1:2 v/v

4.3.3. Electroforesi de les mescles de digestió

Les mostres s'escalfen a 95°C abans de ser sembrades en un gel d'urea-poliacrilamida. Les mostres es sembren junt amb una reacció de seqüència del fragment que s'està estudiant. Aquests gels tenen la mateixa composició que els gels de seqüència (Mètodes 3.5.), però tenen la peculiaritat de tenir una forma de falca. El gruix del gel a la part superior és de 0.4 mm i a la part inferior és d'1.2 mm. La raó per la qual s'usen aquest tipus de gel ("wedge-shaped gel") és la d'intentar que les bandes de digestió més petites (situades a la part més gruixuda del gel) no abandonin el gel mentre es van separant le bandes de tamany més gran (que es troben a la part més prima del gel).

Els gels es corren fins que el blau de bromofenol arribi al final del gel (càrrega curta) o bé es fa una càrrega més llarga si es vol veure alguna zona de protecció que la càrrega curta no permet observar. Se sembren 4 µl dels 8 µl amb els que s'havia ressupès la mostra.

El gel s'asseca a 80°C durant 2 hores i s'exposa a una pel·lícula autorradiogràfica.

5. PREPARACIÓ I ANÀLISI D'RNA

En tots els experiments en els quals es manipulava RNA, el material i les solucions eren doblement autoclavats. Les solucions es feien a partir d'H₂O-DEPC autoclavada dos cops. Les substàncies de partida per realitzar les solucions no eren mai posades en contacte en cap recipient que no hagués estat autoclavat o esterilitzat per escalfament o/n a 200°C.

H₂O-DEPC

DEPC (Dietilpirocarbonat) 1 ml

H₂OmQ 1 l

En agitació o/n i autoclavar 2 cops.

5.1. Preparació d'RNA total

La preparació d'RNA total d'*Escherichia coli* es va dur a terme tal i com havia estat descrit per Belasco *et al* (1985). Els cultius de partida eren de 25 ml, crescuts en diferents condicions en funció de l'experiment que amb l'RNA obtingut es vulgués dur a terme. La quantificació i la determinació de la puresa de l'RNA es duia a terme determinant l'absorbància de la mostra a 280 i a 260 nm. La integritat de l'RNA després de la seva obtenció era comprovat sembrant 5 µg del mateix en un gel d'agarosa amb tampó TAE 1X (Mètodes 3.2.1.).

5.2. Anàlisi d'RNA per "Northern blot"

El "Northern blot" és una tècnica que permet determinar la presència d'un mRNA en un cultiu bacterià en unes condicions determinades, i per tant és una manera d'investigar si un gen s'expressa o no en aquestes circumstàncies. Mitjançant una anàlisi per "Northern blot" és difícil saber les unitats transcripcionals que componen el sistema que s'està estudiant, sobretot quan es duu a terme amb RNA total de procariotes, ja que aquest és molt inestable i es trenca fàcilment. Per tant, el tamany del transcrit observat no sempre és

indicatiu del tamany real del mateix. Per aquest motiu, en el cas dels procariotes, es pot considerar una tècnica més qualitativa que no pas quantitativa.

Per dur a terme els experiments de "Northern blot", 9 µg d'una solució d'RNA total eren sembrats en un gel desnaturalitzant d'agarosa/formamida en tampó MOPS 1X i eren sotmesos a una electroforesi durant 5 hores a 35 V constants, juntament amb marcadors de pes molecular per RNA. Les mostres d'RNA eren sembrades juntament amb el tampó de càrrega denaturant. Posteriorment l'RNA era transferit a una membrana de nylon (Schleicher und Schuell) per capilaritat, i fixat a la mateixa per escalfament a 80°C durant 1 hora.

Aquesta membrana de nylon que contenia l'RNA era posada en contacte amb una solució d'hibridació i amb una sonda marcada radiactivament pel mètode del "random primed" (Mètodes 3.3.2.1.). La sonda en qüestió era un fragment de DNA pertanyent a la zona interna codificant del gen l'expressió del qual es volia analitzar. La temperatura d'hibridació era de 42°C.

Després de diversos rentats amb solucions de diferents graus d'astringència i a diferents temperatures es va posar a contactar la membrana amb una pel·lícula autorradiogràfica.

Tampó MOPS 5X

MOPS	0.2 M
Acetat sòdic	50 mM
EDTA	5 mM
H ₂ O DEPC	qs

Ajustar a pH 7.0 amb NaOH. Abans de la seva utilització s'afegeix al tampó MOPS 1X un 6 % de formaldehid del 37 %.

Tampó de càrrega denaturant

Formamida desionitzada	50 %
Tampó MOPS 5X	qsp 1X
Formaldehid	2 %
Blau de Bromofenol saturat	5 %
Glicerol	12 %
Bromur d'etidi	0.01 µg/ml

Solució de transferència (SSPE 10X)

Fosfat sòdic, pH 7.0	100 mM
NaCl	3.6 M
EDTA	200 mM

Solució d'hibridació

Formamida desionitzada	50 %
Tampó fosfat pH 7.0	50 mM

SDS	1 %
Solució de Denhardt	2.5X
Sulfat de dextrà	10 %
DNA de salmó desnaturalitzat	0.5 mg/ml

Solució Denhardt (50X)

Ficoll 400	1 %
Polivinilpirrolidona 40000	1 %
BSA	1 %

Solució de rentat (SSC 20X)

Citrat sòdic, pH 7.0	300 mM
NaCl	3 M

Els rentats es duïen a terme amb SSC 2X ó SSC 1X als quals s'afegia SDS 0.5 %.

5.3. Determinació de l'inici de transcripció per "Primer Extension"

Per tal de determinar (els) l'inici(s) de transcripció dels gens que componen el sistema *glc* es va dur a terme la tècnica de "primer extension". La tècnica consisteix en la hibridació d'RNA total extret d'un cultiu, crescut en condicions inductores, amb un oligonucleòtid complementari a la regió interna del gen en qüestió. Posteriorment es procedeix a la còpia de la cadena complementària a l'RNA en forma de DNA mitjançant l'acció de la transcriptasa reversa. Aquesta elongació es duu a terme en presència de $\alpha^{35}\text{S}$ -dATP, de manera que la cadena formada queda marcada radioactivament. Un cop acabada la reacció, es procedeix a la seva electroforesi juntament amb una reacció de seqüència del gen obtinguda usant com a cebadors els mateixos primers.

El protocol seguit va ser el descrit per Lange *et al* (1995) i els tampons usats van ser els proporcionats per la casa comercial de la transcriptasa reversa (M-MLVTM Reverse Transcriptase, Gibco, BRL).

Procediment

1. Posar en contacte 50 μg d'RNA total amb 5 pmol de primer en presència de tampó de transcriptasa reversa durant 5 min a 70°C (volum final de la mescla 20 μl).
2. Congelar la mescla en nitrogen líquid i deixar-la descongelar lentament en gel.
3. Addicionar a la mescla descongelada 2.5 μl d'una barreja de dNTPs 5 mM (1.25 mM cadascun), 2 μl d' $\alpha^{35}\text{S}$ -dATP (10 $\mu\text{Ci}/\mu\text{l}$), 200 unitats de transcriptasa reversa i 2 μl de DTT (100 mM). Incubar durant 30 min a 37°C.

4. Addicionar a la mescla 1 μ l d'una solució de dNTPs 100 mM (25 mM cadascun). Incubar durant 30 min més a 37°C.

5. Afegir 175 μ l de TES. Fer una extracció fenol-cloroform i precipitar el DNA (cadena radioactiva elongada) amb etanol al 100 %.

6. Ressupendre el precipitat amb tampó de càrrega i guardar la mostra a -20°C fins al moment de la seva electroforesi.

7. Analitzar la cadena elongada en un gel d'urea-poliacrilamida juntament amb una reacció de seqüència del gen del qual es vol saber l'inici de transcripció.

Tampó de càrrega

Blau de bromofenol	0.1 %
Xylen cianol	0.1 %
NaOH 0.1 M:formamida	1:2 v/v

RESULTATS

IV. RESULTATS

1. SEQÜENCIACIÓ DE LA REGIÓ CROMOSÒMICA QUE CONTÉ EL SISTEMA GÈNIC *glc*

Per tal de poder conèixer quin és el gen que codifica la glicolat oxidasa i com estan organitzats els gens que formen part del sistema *glc*, es va procedir en primer terme a la seqüenciació del locus *glc*.

1.1. Localització del(s) gen(s) que codifica(quen) la glicolat oxidasa dins el locus *glc*

Estudis realitzats prèviament ja havien demostrat que el locus *glc* contenia el(s) gen(s) que codifica(quen) la glicolat oxidasa. Molina *et al* (1993) havien demostrat que el clon pLB10, que té un insert *Hind*III d'11.5 kb, contenia la regió cromosòmica del locus *glc* suficient per codificar l'enzim GOX i l'MSG. També havien demostrat que el clon pIM4 només expressava l'MSG. Per tal de localitzar el(s) gen(s) de la glicolat oxidasa dins del fragment genòmic clonat en el clon pLB10, es va construir una col·lecció de subclons per digestió amb diferents enzims de restricció i posterior clonatge dels fragments resultants en el vector pBluescript. Els plasmidis recombinants resultants (de pTP21 a pTP30 en la Figura 1) van ser

usats en estudis de complementació de les activitats enzimàtiques deficientes en la soca JA151 (Molina *et al*, 1994), un derivat *recA* de la soca DV21A05, delecionada en el locus *glc*.

Partint del clon pLB10 es va obtenir el fragment *Hind*III-*Eco*RI de 8.8 kb, que va ser clonat en pBluescript per donar el clon pTP21. El clon pTP22 va construir-se a partir del clon pLB10, digerit amb els enzims de restricció *Hind*III i *Sal*I, i posterior clonatge del fragment de 6.3 kb en pBluescript. El clon pT23 va ser construït per clonatge del fragment *Hind*III-*Bam*HI de 5.3 kb, obtingut a partir del clon pLB10, en pBluescript. Per construir el clon pTP24 es va partir del clon pTP21. Aquest va digerir-se amb *Cla*I i *Eco*RI. El fragment de 6.8 kb alliberat va ser clonat en pBluescript. El clon pTP25, que conté un insert de 3.8 kb comprès entre les dianes *Eco*RV va obtenir-se per digestió de pTP24 amb aquestes dianes i clonació en pBluecript. El clon pTP26 conté l'insert *Bam*HI de 3.1 kb obtingut a partir pTP25. El fragment *Kpn*I-*Sal*I d'1.3 kb va ser obtingut a partir del clon pTP24 i després de ser clonat en pBluescript va resultar en el clon pTP27. El clon pTP28, que va obtenir-se a partir del clon pTP23, conté un insert de 3.1 kb comprès entre les dianes *Cla*I i *Kpn*I. El clon pTP29, que conté un fragment de 0.9 kb limitat per les dianes *Eco*RV i *Bam*HI, va obtenir-se a partir del clon pTP23. Per últim, el clon pTP30 va obtenir-se per digestió del clon pTP26 amb *Mlu*I i *Bam*HI i subclonatge del fragment d'1.6 kb alliberat en pBluescript.

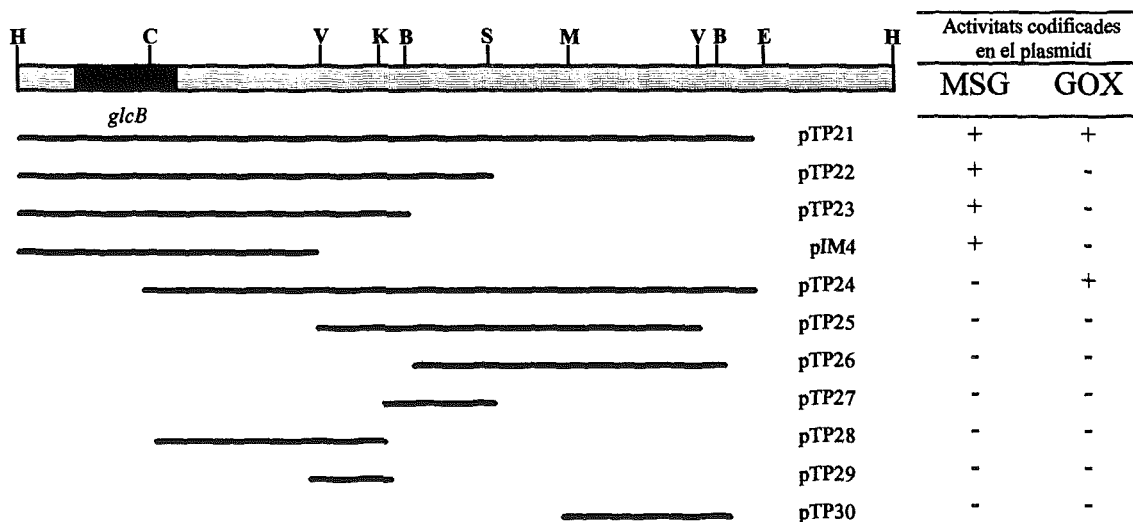


Figura 1. Construcció de subclons per la localització dins el locus *glc* del gen que codifica glicolat oxidasa. A la part superior es mostra el mapa físic parcial del fragment *Hind*III d'11.5 kb present en el clon pLB10 on s'indica la posició del gen *glcB* (Molina *et al*, 1994). Els inserts dels plasmidis usats per la localització del gen que codifica la GOX s'indiquen mitjançant línies. Tots els inserts dels plasmidis representats en aquesta figura han estat clonats en el vector pBluescript. Per cada plasmidi es mostra la presència (+) o absència (-) de les activitats malat sintasa G (MSG) i glicolat oxidasa (GOX). B: *Bam*HI; C: *Cla*I; H: *Hind*III; K: *Kpn*I; M: *Mlu*I; S: *Sal*I; V: *Eco*RV.

En la Figura 1 es mostra el patró d'expressió de les activitats MSG i GOX en extractes cel·lulars obtinguts a partir de cultius en hidrolitzat de caseïna de la soca JA151 transformada amb els diferents plasmidis recombinants.

Només els plasmidis que contenen la regió de 6 kb compresa entre el gen *glcB* i la diana *EcoRI* eren capaços de restaurar l'activitat GOX deficient en la soca JA151. Cal remarcar que els nivells d'ambdues activitats codificades en aquest locus, MSG i GOX, eren molt elevats (de l'ordre de 100 vegades respecte l'expressió de la soca tipus salvatge) i no incrementaven per la presència de glicolat en el cultiu, indicant que l'expressió dels gens era constitutiva des del promotor del vector.

1.2. Estratègia de seqüenciació

Per dur a terme la seqüenciació de les 6 kb que expressaven l'activitat GOX en un mutant Δglc i poder així identificar els gens localitzats en aquest fragment genòmic, es van escollir quatre sucblons (pTP26, pTP27, pTP28 i pTP29) que cubrien la regió entre la diana *Clal* interna a *glcB* i la diana *BamHI* més propera a la diana *EcoRI* (Figura 1). Aquests clons tenien una zona de solapament entre ells, per evitar així els possibles errors provocats per la proximitat de seqüències diana per un mateix enzim de restricció no detectables per digestió. De cadascun dels clons es van obtenir varies deleccions, i cadascuna d'elles va ser seqüenciada com a mínim 2 cops per cada cadena de DNA. En la Figura 2 es mostren els 5335 pb de DNA seqüenciats entre la ja anomenada diana de restricció *BamHI* i el codó d'inici de la traducció ATG de *glcB*, prèviament descrit per Molina *et al* (1994). A aquesta seqüència se li ha assignat el número d'accés L43490 en el banc de dades GenBank.

```

1 ggatcctcgaaaccaatacagggcatgaataaacatgctggaacgcattggggcggagccgtacaagcgttaccgattgtatgaaaagcaga 90
91 ttaataaccagctaactcagggttcotctccagcggcaccgcggagcgaatcaaatggtggccttcagctcgatatacgtggagattcttt 180
   * S L N M E L P V A S R I L H H G E L E I D H L N K
181 ttcaaggctccgcacatgatcgctgctgcgcgctggcgacgtgcggcagccgctgcaacaccgcggtttagatccgcgcatgctggcga 270
   K V T R V H D R A A R Q A V H P L R Q L V A N Y I R A H Q R
271 tcgatctgcttttttggctgtggtcgatggttagagattatttacgctggcacaactgaattaacatcagatcggatcaatgattgcagc 360
   D I Q K K Q Q P R H Y L N N V S A F V S N F M L D T L S Q L
361 gtaaacaccagcacctgatgtgagaagcctgacaaatggcgagatggaagcgtgatccaactgcgcatgttcgatcagcgaatatctct 450
   T F V L V Q N H S A Q C I A L H F A H D L Q A H E I L S I E
451 ttgttgttctcactggcggcgagcattttttcataacagcgggttatcacaacaaaatcagcctgcgttcccagcgttgccgcagcctt 540
   K N N E S A A L M K E Y C R T I V V F D A Q T G L T A A L R

```

541 gccgattgcctccagtaatgcgcaacgtcgagcagatcgtagcggcttcgcggtcggtactgaacagatggatcagcgggctggg 630
A S E G E L L A R V D L L D Y L T R P Q T S F L H I L P S T

631 tctgcaaccgattaagccgtgacacagagaatcgacccctgcgcggttcaataatcccgcccgccgagcagcggtcagccctca 720
D Q V R N L R A V R S D R G Q A T E I I G R G R L V T L G E

721 cgcagtcgggagcgtgagaagccgagcttttcacacatgcagcttcgaggggaagcggctgaccgacctcagtagcgcctcgataatt 810
R L A S R S F G L K E C L R R E S P L P Q G V K L V G D I I

811 aaccgttgatctctctgcaaccacttcgcaaataggcgcagcttcattcttcccgacccctcgtgacacaggtaggaccaattttt 900
L R E Y S E A V V E C I P R R E D K M
← *glcC*

901 ctgttcttaactcgcaaacacgcacatcacgtaagtgattatgttacatcaatttaacattgagtaaccaagacaaggtcacagagct 990

991 ggaaaaaaaaatggtctgaccggtagctaaagagatagacgasaacgaaaagcccgcttaataactgttccacagaagcagcgcgcaaaaat 1080

1081 cagctgccacacaacacaacaaagcgaagcctactcatgagcatcttgtagcaagagcgtcttgtagggcgtttaccgatgtagcaccgc 1170
M S I L Y E E R L D G A L P D V D R
glcD →

1171 acatcggtagctgatggcaactgcgtgagcatgtccctggaacttgagatcctgcataaccgatgaggagatcattccttacagagtgtgacggg 1260
T S V L M A L R E H V P G L E I L H T D E E I I P Y E C D G

1261 ttgagcgcgtatcgcaaccgctcattactggttctctgcctaagcaaatggaacaggtgacagcagcttctggctgtctgccatcgccgtg 1350
L S A Y R T R P L L V V L P K Q M E Q V T A I L A V C H R L

1351 cgtgtaccggtggtgaccocgtggtgcaggcacccggctttctggtggcgcgctgcccgtggaaaaaggtgtgttgttggtagggcgcgc 1440
R V P V V T R G A G T G L S G G A L P L E K G V L L V M A R

1441 tttaaagagatcctgacattaaaccogttggtcgcgcgcgcgctgacagccagcgtgagtaacctggcgatctccagggccgttgca 1530
F K E I L D I N P V G R R A R V Q P G V R N L A I S Q A V A

1531 ccgcataatctctactcagcaccggaacctctcacaatcgccgttccattggcggaatgtggctgaaaatcccgccggcgtccac 1620
P H N L Y Y A P D P S S Q I A C S I G G N V A E N A G G V H

1621 tgcctgaaatattggtctgaccgtacataacctgctgaaaattgaagtcaaacctggagcggcggagcactgacactggatcggacgcg 1710
C L K Y G L T V H N L L K I E V Q T L D G E A L T L G S D A

1711 ctggattoacctggttttgacctgctggcgtgttccacggatcggaaggtatgctcggcgtgaccaccgaagtgacggtaaaactgctg 1800
L D S P G F D L L A L F T G S E G M L G V T T E V T V K L L

1801 ccgaagccgcccggtggcggggttctgttagccagcttgactcgttagaanaagccgacttgcggttggtagacatcagcccaatggc 1890
P K P P V A R V L L A S F D S V E K A G L A V G D I I A N G

1891 attatcccgccggggtgagatgatggataacctgtcgatccgcgcggcgaagattttatcattgcccgttatcccgtagacgcgaa 1980
I I P G G L E M M D N L S I R A A E D F I H A G Y P V D A E

1981 gcgattttgttatgagcgtggacggcgtggagtctgacgtacaggaagactgagcggggttaacgacatcttgttgaagcgggcgcg 2070
A I L L C E L D G V E S D V Q E D C E R V N D I L L K A G A

2071 actgacgtccgtctggcaacaggaagcagagcgcgtacgtttctggccggctcgcaaaaatgcgttcccgccggtaggacgtatctcc 2160
T D V R L A Q D E A E R V R F W A G R K N A F P A V G R I S

2161 ccggattactactgcatggatggaccacatcccgctgcgcgcctcctggcgtactggaaggcattgcccgttatcgcagcaatatgat 2250
P D Y Y C M D G T I P R R A L P G V L E G I A R L S Q Q Y D

2251 ttacgtgttgcaacgtctttcatgcggagatggcaaacatgcaccogttaaactcttttgatgccaacgaaccgggtgaatttgcccgc 2340
L R V A N V F R A G D G N M H P L I L F D A N E P G E F A R

2341 gcggaagagctggcgggaagatcctcgaactctgcgttgaagttggcggcagcatcagtgggcaacatggcatcggcgagaaaaaatc 2430
A E E L G G K I L E L C V E V G G S I S G E H G I G R E K I

2431 aatcaaatgtagccagcttcaacagcagatgaaatcacgaccttccatgcggcgaaggcggcgtttgaccccgatggtttgctgaaccct 2520
N Q M C A Q F N S D E I T T F H A V K A A F D P D G L L N P

2521 gggaaaaacattcccagctacacogctgtgtgaatttgggtgccatgcatgtgcatcacggctatttacotttccctgaactggagcgt 2610
 G K N I P T L H R C A E F G A M H V H H G H L P F P E L E R

2611 ttctgatgctacogcagtgattacagccagggcgtgctggagcaggatgaatcaggcgattagcgataaaaacggcgtggtgattcagg 2700
 F * M L R E C D Y S Q A L L E Q V N Q A I S D K T P L V I Q G
glcE →

2701 gcagcaatagcaaagccttttaggtcgccctgtcacogggcaaacgctggatgttctgttcatcggcgattgtaattacgacccga 2790
 S N S K A F L G R P V T G Q T L D V R C H R G I V N Y D P T

2791 ccgagctggtgataaccgcgctgctggaacgcgctggtgacaattgaagcggcgtggaagcggggcaaatgctcccctgtgagc 2880
 E L V I T A R V G T P L V T I E A A L E S A G Q M L P C E P

2881 cggcgcattatggtgaagaagccacctggggcggtggtcgcctgcggtggtggggcgcgctgcccgtgagcggttcggtcccg 2970
 P R Y G E E A T W G G M V A C G L A G P R R P W S G S V R D

2971 atttgtcctcggcacgcgcattaccogggcgtgaaaacatctgcgttttgggtggcgaagtgatgaaaaacgttgcgggatacagtc 3060
 F V L G T R I I T G A G K H L R F G G E V M K N V A G Y D L

3061 tctcaccggttaatggtcggaagctacggttctctggcgtgctcaactgaaatctcaatgaaagtgttacgcgacccgcgctcctcctga 3150
 S R L M V G S Y G C L G V L T E I S M K V L P R P R A S L S

3151 gcctcgcctgggaaatcagcctgcaagaagccatgagtgaaatcgccgagtggaactccagccattaccattagtgottatgttaact 3240
 L R R E I S L Q E A M S E I A E W Q L Q P L P I S G L C Y F

3241 tcgacaatgctgtgtggtaccgcttgagggcggaaggtcggttaaagcagcgctgaaactgctgggtggcgaagaggttgcgggtc 3330
 D N A L W I R L E G G E G S V K A A R E L L G G E E V A G Q

3331 agttctggcagcaattgctgtaacaacaactgcccgttctctcgttacaggtaacctatggcgatttcattaccagtgatgogccga 3420
 F W Q Q L R E Q Q L P F F S L P G T L W R I S L P S D A P M

3421 tgatggatttaccogcgagcaactgatcgactggggcggttacgctggctgaaatcgacagccgaggacaatcaaatccatcgca 3510
 M D L P G E Q L I D W G G A L R W L K S T A E D N Q I H R I

3511 tcgcccgaacgctggcgtatcgaccccgtttagtcggagatgggtgcttggcccctatcgctcctttattccgctatcacc 3600
 A R N A G G H A T R F S A G D G G F A P L S A P L F R Y H Q

3601 agcagcttaaacagcagctcgacccttgcggcgtgttaaccocggtcgcatgtacgoggaactttgaggagcaggctatgcaaacccaa 3690
 Q L K Q Q L D P C G V F N P G R M Y A E L * M Q T Q
glcF →

3691 ttaactgaagagatgcccgaacgcgcgcgctggaagccagcagcctcctgcgcgctgttctactgcggatgttgatccgcaacc 3780
 L T E E M R Q N A R A L E A D S I L R A C V H C G F C T A T

3781 tgcccaacctatcagcttctggcgtgaaactggacggcgcgcggtcctctatctgattaaacaggtgctggaagcaacgaagtc 3870
 C P T Y Q L L G D E L D G P R G R I Y L I K Q V L E G N E V

3871 acgcttaaacacagcagcctcgcctcgcctcacttgcgtaattgtgaaaccacctgcctctggtgtgcgctatcacaattg 3960
 T L K T Q E H L D R C L T C R N C E T T C P S G V R Y H N L

3961 ctggatatggcgtgatattgtcgagcagaagtgaacgcccaactgocggagcgaatactgcggaaggattgcgccaggtagtgccg 4050
 L D I G R D I V E Q K V K R P L P E R I L R E G L R Q V V P

4051 cgctccggcgtcttccgctgagcaggtagggtggtgctgacagctttttaccggaaacaggtcagagcaaaaatgctgctgaa 4140
 R P A V F R A L T Q V G L V L R P F L P E Q V R A K L P A E

4141 acggtgaaagctaaacgcgctcgcgcgctgcccataagcgtcgggttttaattgtggaaggtgcgccagcctacgcttttcgcccac 4230
 T V K A K P R P P L R H K R R V L M L E G C A Q P T L S P N

4231 accaacggcgaactgcgcgagtgctggtcgtctggggtacagcgtatgocagtaaacgaagcagcgttctgtggcgggtgactat 4320
 T N A A T A R V L D R L G I S V M P A N E A G C C G A V D Y

4321 catcttaatgcagcaggaagggctggcacggcgcgcaataatattgatgctggtggcccgatggaagcaggtgocgaggaatt 4410
 H L N A Q E K G L A R A R N N I D A W W P A I E A G A E A I

```

4411 ttgcaaacgcgcoagcggctgcccggcgtttgtcaaaagagtatgggcagatgctgaaaaacgatgcgttatatgccgataaaagcacgctcag 4500
    L Q T A S G C G A F V K E Y G Q M L X N D A L Y A D K A R Q

4501 gtcagtgaaactggcggctgatttagtogaacttotgcgcgaggaacccgctggaaaaactggcaattcgccggcgataaaaagctggccttc 4590
    V S E L A V D L V E L L R E E P L E K L A I R G D K K L A F

4591 cactgtccgtgtaccctacaacatgcgcgcaaaagctgaacggcgaagtggaaaaagtgttgcctcgtcttgatttaccttaacggacggt 4680
    H C P C T L Q H A Q K L N G E V E K V L L R L G F T L T D V

4681 cccgacagccatctgtgctgcccgggtccagcgggaacatatacggttaacgcaccccgatctggcagccagctgcccgggatacaaaaatgaat 4770
    P D S H L C C G S A G T Y A L T H P D L A R Q L R D N K M N

4771 gcgctggaagcggcaaacccgaaatgatogtccaccgcaacattggttccagacgcacatctggcagcgcggctcgtacctctgtgcgt 4860
    A L E S G K P E M I V T A N I G C Q T H L A S A G R T S V R

4861 cactggattgaaattgtagaacaagcccttgaagggaataacaaaatgaaaactaaagtcattcttagccagcaaatggcgagtgcaat 4950
    H W I E I V E Q A L E K E * M K T K V I L S Q Q M A S A I
                                glcG →

4951 tattgccgcaggtcaggaagaggcgcagaaaaataaactggtctgtttccattgctgttgccgatgacggcggctcatctgctggcgttaag 5040
    I A A G Q E E A Q K N N W S V S I A V A D D G G H L L A L S

5041 tcgcatggacgattgcccgcgattgcccgttatatctcccaggagaaagcgcgtaccgcccgctggggcgtcgtgaaactaaggccta 5130
    R M D D C A P I A A Y I S Q E K A R T A A L G R R E T K G Y

5131 tgaagagatggtgaacaacggacgtaccgcgtctgactgcccgttattaacgtcgtggaagggcggctaccggttgttgggatgg 5220
    E E M V N N G R T A F V T A P L L T S L E G G V P V V V D G

5221 gcaaattattggtgcccgtggcgtttctggtttaaccggagcacagatgcccaggtcgcgaaagcggcagcagcgggttggcgaata 5310
    Q I I G A V G V S G L T G A Q D A Q V A K A A A A V L A K *

5311 agcgaaaacgaggagataaacaatg 5335
                                M
                                glcB →

```

Figura 2. Seqüència nucleotídica del fragment genòmic de 5.3 kb situat a 5' de *glcB* en el locus *glc*. En la figura només es mostra una cadena de DNA. Els gens *glcC* (transcrit en el sentit contrari a les agulles del rellotge) i *glcD*, *glcE*, *glcF* i *glcG* (transcrits en el sentit de les agulles del rellotge) han estat traduïts a aminoàcids en el codi d'una lletra; els símbols dels aminoàcids estan escrits sota el primer nucleòtid del corresponent codó. Els codons de parada estan indicats amb un asterisc. La direcció de transcripció està indicada amb una fletxa. També s'indica el codó d'inici de *glcB* (número d'accés X74547). Estan subratllats els possibles llocs d'unió a ribosoma. A aquesta seqüència li ha esta assignat el número d'accés L43490 del banc de dades GenBank.

En la seqüència es poden identificar cinc nous marcs oberts de lectura, que han estat anomenats com *glcD*, *glcE*, *glcF*, *glcG* en la mateixa cadena i immediatament seguits per *glcB*. A la cadena complementària a aquests 5 gens apareix un marc obert de lectura, que es transcriu divergentment a la resta, i que ha estat anomenat *glcC*.

Per cada marc obert de lectura s'ha detectat un triplet d'inici de la traducció i també una seqüència d'unió a ribosoma, rbs (ribosome binding site) o seqüència Shine-Dalgarno. En tots els gens aquesta seqüència és rica en bases púriques, complementàries a l'rRNA 16S de la subunitat 30 S dels ribosomes procariotes.

1.3. Anàlisi computacional de la seqüència aminoacídica deduïda dels diferents marcs oberts de lectura identificats

Les característiques dels marcs oberts de lectura que es descriuen en els apartats següents són derivades de l'anàlisi de la seqüència presentada en la Figura 2. L'anàlisi s'ha dut a terme usant el PC/GENE software package (Intelligenetics Inc.) i el University of Wisconsin Genetics Computer Group package program. L'estudi de *glcT* s'ha fet basant-se en la seqüència U28377 del banc de dades de seqüències GenBank (Blattner *et al*, 1997), corresponent a la regió cromosòmica d'*Escherichia coli* compresa entre els minuts 64.5 i 67.

Cal esmentar que en comparar la seqüència de 5335 bp de la Figura 2 obtinguda en aquest treball amb la publicada posteriorment per Blattner *et al* (1997) dins els projecte de seqüenciació d'*Escherichia coli*, existeix un únic nucleòtid de diferència. En la posició 3572 de la seqüència presentada en aquest treball (L43490) existeix un nucleòtid G que no apareix en la seqüència de Blattner. Aquest canvi provoca una fusió dels marcs oberts de lectura *glcE* i *glcF* que donaria lloc a l'expressió d'una proteïna de 84 kDa. Com es mostrarà més endavant per l'expressió de gens en maxicèl·lules (Resultats 2.1.), la seqüència que es mostra en aquest treball és la correcta.

1.3.1. Característiques de *glcC*

El gen *glcC* comença a la posició 866 i acaba a la posició 102 (Figura 2). Es pot observar la seqüència Shine-Dalgarno (GAGG) a 11 pb en direcció 5' del codó d'inici de la traducció. Aquest marc obert de lectura codifica una proteïna de 254 aminoàcids amb una massa molecular calculada de 28901 Da. La freqüència òptima d'utilització de codons ("codon usage") segons va ser descrita per Ikemura (1981) és de 0.64. La seqüència aminoacídica deduïda va ser comparada amb seqüències del banc de dades GenBank, i va mostrar ser similar a varies proteïnes reguladores procariotes. Es va trobar un motiu "hèlix-volta-hèlix" a l'extrem N-terminal de la seqüència aminoacídica derivada. Aquest motiu és típic de dominis d'unió al DNA de factors de transcripció procariotes, i va resultar ser similar a les següents proteïnes d'*Escherichia coli*: al de la proteïna reguladora de l'operó *lct*, implicat en la utilització d'L-lactat (Dong *et al*, 1993), al del repressor del complexe de la piruvat deshidrogenasa (Haydon *et al*, 1993) i al de la proteïna reguladora de reguló *exu* (seqüència no publicada amb número d'accés U18997). En la Figura 3 es mostra l'aliniament de seqüències de les proteïnes abans

esmentades, on es pot observar que és a l'extrem N-terminal on totes elles tenen aquest motiu d'unió al DNA.

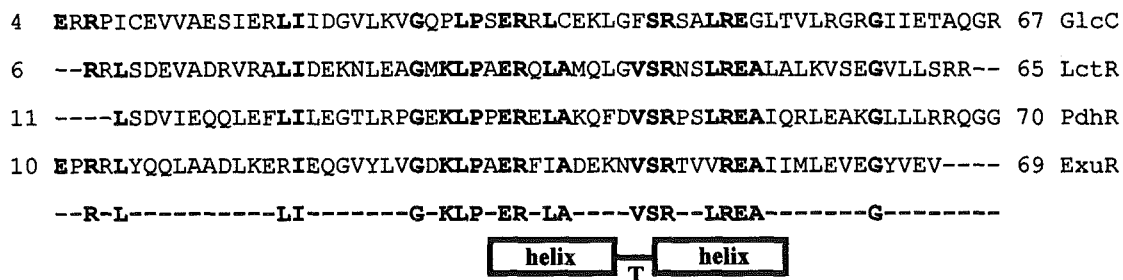


Figura 3. Anàlisi comparativa de la seqüència aminoacídica de diferents proteïnes reguladores que mostren alta similitut amb GlcC. En aquesta figura es mostren les regions corresponents al domini d'unió al DNA de les diferents proteïnes reguladores comparades. La seqüència inferior mostra la seqüència més conservada entre totes les proteïnes analitzades, i a sota es mostra la localització del motiu d'unió a DNA comú en tots els reguladors. LctR: proteïna reguladora de l'operó *lct*; PdhR: repressor del complexe de la piruvat deshidrogenasa; ExuR: proteïna reguladora del reguló *exu*; helix: hèlix; T(turn): volta. Els números indiquen les posicions dels aminoàcids en cada proteïna.

GlcC forma part de la família GntR de reguladors transcripcionals (Reizer *et al*, 1991). Aquesta família de proteïnes es caracteritza per tenir el motiu d'unió al DNA a l'extrem N-terminal. L'alta similitut de GlcC amb proteïnes reguladores procariotes sembla indicar que GlcC podria ser la proteïna reguladora del sistema *glc*.

1.3.2. Característiques de *glcD* i *glcE*

El gen *glcD*, comprès entre les posicions 1117 i 2616 (Figura 2), està separat de *glcC* per una regió de 251 pb. A 11 pb del codó ATG es pot observar una bona seqüència Shine-Dalgarno (GAAG). El gen codifica un polipèptid de 499 aminoàcids amb una massa molecular calculada de 53749 Da. El "codon usage" és de 0.49. La comparació amb seqüències proteïques de la base de dades GenBank va mostrar una significativa similitut amb les seqüències de D-lactat deshidrogenases de *Kluyveromyces lactis* (Lodi *et al*, 1994) i de *Saccharomyces cerevisiae* (Lodi i Ferrero, 1993), i amb menor grau amb el mateix enzim d'*Escherichia coli* (Rule *et al*, 1985). També van mostrar similitut amb els enzims d'*Escherichia coli* glicerol-3-fosfat deshidrogenasa anaeròbica (Cole *et al*, 1988) i fumarat reductasa (Cole, 1982)

El codó d'inici ATG de *glcE* està sobreposat al codó de terminació de *glcD*. La seqüència Shine-Dalgarno es troba a 12 pb del seu ATG. El gen codifica una proteïna de 350 aminoàcids, que té una massa molecular deduïda de 38352 pb. El seu "codon usage" és de 0.60. Aquest polipèptid presenta també la màxima similitut amb la seqüència de la D-lactat deshidrogenasa de *Saccharomyces cerevisiae*.

En l'anàlisi de l'estructura primària dels productes gènics de *glcD* i *glcE* es van identificar possibles llocs d'unió a nucleòtids de flavina. En el cas de GlcD es va trobar un consens parcialment conservat de la seqüència d'unió a flavines descrita per Branden i Tooze (1991). Aquesta seqüència consens té el motiu G-X-G-X-X-G separat d'un glutamat conservat a 17 residus; entre el consens i el glutamat conservat existeixen diverses positions invariables de residus hidrofòbics. En la Figura 4 es pot veure l'aliniament de seqüències i la conservació del motiu.

84	TRGAGTGLSGGALPLEKGVLLVMARFKE	111	GlcD
5	TVIMGGGLAGLLCGLQLQKHGLRCAIVT	32	GlpB
11	VIIIGGGATGAGIARDCALRGLRVILVE	38	GlpA
9	AVVIGAGGAGIARLAQISQSGQTCALLS	36	SdhA
8	VVVLGAGPAGYSAAFRCADLGLLETVIVE	35	Lpd

Figura 4. Seqüència consens del lloc d'unió al cofactor flavínic de diferents enzims d'*Escherichia coli*. GlcD: glicolat oxidasa; GlpA i GlpB: glicerol-3-fosfat deshidrogenasa anaeròbica; SdhA: succinat deshidrogenasa i Lpd: lipoamida deshidrogenasa. Les positions invariables per les glicines i el glutamat conservat estan marcats en negreta. Els símbols subratllats marquen les positions d'aminoàcids hidrofòbics semiconservades que completen el consens per la unió a flavines segons ha estat descrit per Branden i Tooze (1991). Els números indiquen les positions dels aminoàcids en cada proteïna.

L'anàlisi de la seqüència aminoacídica de GlcD i GlcE va permetre també identificar una seqüència amb una similitut destacable a la seqüència dels centres actius de varies oxidoreductases. Aquesta seqüència conté un residu d'histidina (Hys-140 per GlcD i Hys-91 per GlcE) separada d'un residu de cisteïna per 13 i 12 aminoàcids respectivament. El mateix patró es pot observar en els enzims d'*Escherichia coli* glicerol-3-fosfat deshidrogenasa anaeròbica (Cole *et al*, 1988), fumarat reductasa (Cole, 1982) o succinat deshidrogenasa (Wood *et al*, 1984).

En la literatura s'ha descrit que enzims que tenen com a cofactors FAD o FMN veuen augmentada la seva activitat enzimàtica si aquests s'afegeixen en el medi de reacció. Per assajar la possible activació de l'activitat GOX per l'addició en el medi de reacció de FAD i

FMN, es van usar concentracions d'ambdós en un rang de 2-50 µM i 0.2-5 mM respectivament. No es va poder observar cap augment en l'activitat glicolat oxidasa per addició externa d'aquests cofactors.

1.3.3. Característiques de *glcF*

S'han identificat 3 possibles codons d'inici de la traducció de *glcF* en les posicions 3532, 3556 i 3679 de la Figura 2. Els dos primers codons es solapen amb la zona codificant de *glcE* i no tenen cap seqüència Shine-Dalgarno ben conservada en posició 5'. El marc obert de lectura que comença a la posició 3679 i acaba a la posició 4902 té una possible seqüència d'unió a ribosoma a 12 pb de l'inici de traducció. L'ús d'aquest triplet iniciador ATG dona un producte gènic de 407 residus aminoacídics amb una massa molecular de 45083 Da. El seu "codon usage" és de 0.62.

22	LRACVHCGFCTATCPTYQLLG	42	GlcF
72	LDRCLTCRNCETTCP SGVRYH	92	GlcF
6	FENCIKCTVCTTACPV SRVNP	26	GlpC
53	LKYCINCKRCEVACPS SDVKIG	73	GlpC
145	LYECILCACCS TSCPSFWWNP	165	SdhB
145	FSGCINCGLCYAACP QFGLNP	165	FrdB
5	T DECISCGACAAECP VEAIHE	25	Fdx (CT)
34	ADTCIDCGACEAVCP TGAVKA	54	Fdx (CT)
5	E ETCISCGACAAECP VNAIEQ	25	Fdx (RR)
34	ADTCIDCGNCANVCP VGAPVA	54	Fdx (RR)
	+--C+-C--C---CP-----*		

Figura 5. Aliniament de diverses seqüències aminoacídiques de proteïnes que contenen centres ferrosulfurats. En aquesta figura es mostren els aliniaments de les següents proteïnes: GlcF: glicolat oxidasa d'*E.coli*; GlpC: glicerol-3-fosfat deshidrogenasa anaeròbica d'*E.coli* (Cole *et al*, 1988); SdhB: succinat deshidrogenasa d'*E.coli* (Darlison i Guest, 1984); FrdB: fumarat reductasa d'*E.coli* (Cole *et al*, 1982); Fdx (CT): ferredoxines de *Clostridium termosaccharolyticum* (Tanaka *et al*, 1973) i Fdx (RR): ferredoxines de *Rhodospirillum rubrum* (Matsubara *et al*, 1983). Els diferents clusters de cisteïnes estan aliniats i els residus altament conservats estan ressaltats en negreta. Els números indiquen les posicions dels aminoàcids en cada proteïna. La seqüència consens està representada a l'última línia del dibuix, on les cisteïnes estan indicades com a C, les prolines com a P, els residus hidrofòbics conservats com a + i els residus alifàtics com a *.

GlcF és una proteïna amb un alt contingut en cisteïnes. Existeixen 17 residus cisteïnics, la majoria dels quals estan organitzats en 2 centres típics de proteïnes ferrosulfurades implicades en processos respiratoris. Els centres tenen una estructura C-X-X-C-X-X-C-X-X-X-C. L'aliniament de la seqüència d'aminoàcids d'aquests centres ferrosulfurats amb la dels centres de diferents enzims respiratoris, va permetre identificar altres posicions conservades a part de les cisteïnes. Per exemple, trobem un residu hidrofòbic situat a dues posicions per davant de la primera cisteïna, que és seguida per un altre residu hidrofòbic. L'última cisteïna és seguida per una prolina i a sis posicions hi trobem un residu alifàtic (Figura 5).

Entre les proteïnes que en un aliniament de seqüència donen alta similitut amb GlcF trobem la subunitat GlpC de la glicerol-3-fosfat deshidrogenasa anaeròbica d'*Escherichia coli* (Cole *et al*, 1988) i la subunitat ferrosulfurada de la succinat deshidrogenasa de diferents espècies. Tots aquests enzims són proteïnes multimèriques amb mecanismes catalítics similars: oxiden diferents molècules i transfereixen els electrons via les subunitats ferrosulfurades a algun component de la cadena de transport electrònic.

1.3.4. Característiques de *glcG*

El codó d'inici de *glcG* està localitzat a 5 pb en direcció 3' respecte el codó de parada de *glcF*. La seva seqüència Shine-Dalgarno es troba situada a 11 pb del codó ATG d'inici de la traducció. El gen codifica un polipèptid de 134 aminoàcids amb una massa molecular deduïda de 13737 Da. El seu "codon usage" és de 0.71. GlcG no presenta alta similitut amb cap proteïna o grup de proteïnes, excepte amb una proteïna de 142 aminoàcids de funció desconeguda de *Citrobacter freundii* (ORFY situat al costat del reguló *dha*, número d'accés CFU09771 al GenBank) (Daniel *et al*, 1995).

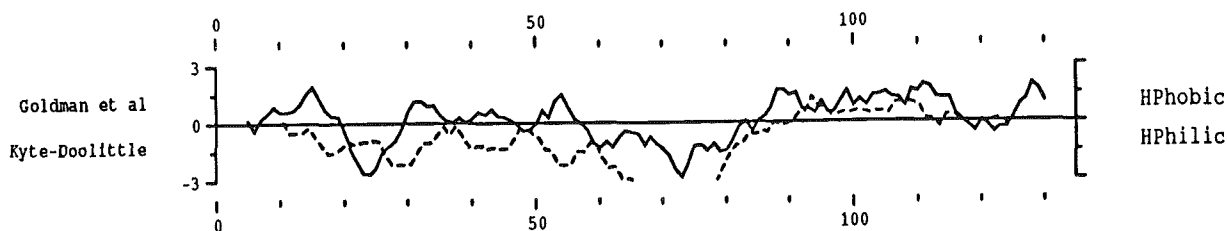


Figura 6. Perfil hidropàtic de GlcG. Aquest perfil ha estat calculat amb el programa PEPLOT (Genetics Computer Group de la Universitat de Wisconsin). Es mostra el perfil segons l'algoritme de Goldman *et al* usant una finestra de 20 aminoàcids (línia contínua, que identifica les possibles hèlix transmembrana no polars), i també segons Kyte-Doolittle (línia discontinua, que indica l'índex d'hidrofobicitat específica dels residus aminoacídics en una finestra de 9 aminoàcids).

Aquesta proteïna no presenta un perfil d'hidrofobicitat similar a proteïnes de membrana, però sí que és remarcable el fet que en el seu extrem C-terminal té una regió altament hidrofòbica amb una invaginació més hidrofílica (Figura 6).

1.3.5. Característiques de *glcT*

Gairebé simultàniament a la submissió al GenBank de la seqüència L43490 (Figura 2), Blattner *et al* (1997) van seqüenciar la mateixa zona del genoma d'*Escherichia coli* dins el projecte global de seqüenciació del genoma d'*Escherichia coli*. La zona per ells determinada era més àmplia que la que nosaltres havíem seqüenciat, i allà vam poder veure que en direcció 3' a *glcB* existeix un marc obert de lectura amb una alta similitud a proteïnes hidrofòbiques de membrana. Aquest ORF, anomenat ORF-f560 per Blattner *et al* (1997) té un codó inicial ATG a 355 pb del codó de parada de *glcB*, en la mateixa direcció de transcripció. La seva seqüència nucleotídica abarca des de la posició 75671 a la 73991 de la seqüència numerada U28377 i d'ella se'n deriva un polipèptid de 560 aminoàcids amb una massa molecular estimada de 61600 Da. A 11 pb en direcció 5' del triplet d'inici de la traducció existeix una possible seqüència d'unió a ribosoma GAGAA que és conservada respecte al consens. El seu "codon usage" és de 0.64.

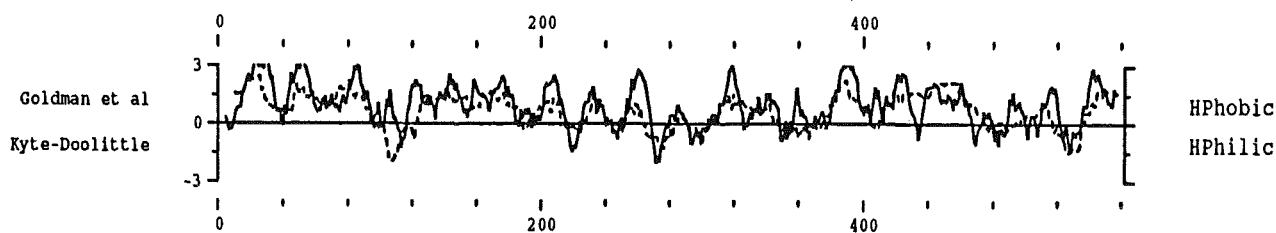


Figura 7. Perfil d'hidropaticitat de GlcT. Aquest perfil ha estat calculat amb el programa PEPLOT (Genetics Computer Group de la Universitat de Wisconsin). Es mostra el perfil segons l'algoritme de Goldman *et al* usant una finestra de 20 aminoàcids (línia contínua, que identifica les possibles hèlix transmembrana no polars), i també segons Kyte-Doolittle (línia discontinua, que indica índex d'hidrofobicitat específica dels residus aminoacídics en una finestra de 9 aminoàcids).

Els perfil d'hidropaticitat de la proteïna és inequívoc de proteïnes de membrana (Figura 7). Mostra que GlcT és una proteïna de membrana i es poden identificar un total de 13 regions

transmembrana potencials. Aquesta proteïna ha estat classificada com a membre de la família LldP de transportadors.

Identities = 182/271 (67%), Positives = 229/271 (84%)

```

1 MVTWTQMYMPMGGGLGLSALVALIPIIFFFFVALAVLRLKGHVAGAITLILSILIAIFAFKM 60
  M W Q Y P G + LS+L+A +PI+FFF AL L+LKG+VA + T+ +++ +A+ +KM
1 MNLWQQNYDPAGNIWLSSLIASLPILFFFFALIKLKLKGYVAASWTVAIALAVALLFYKM 60

61 PIDMAFAAAGYGFIYGLWPIAWIIIVAAVFLYKLTVASGQFDIIRSSVISITDDQRLQVLL 120
  P+ A A+ YGF YGLWPIAWII+AAVF+YK++V +QQFDIIRSS++SIT DQRLQ+L+
61 PVANALASVVYGYFFYGLWPIAWIIIAAVFVYKISVKTGQFDIIRSSILSITPDQRLQMLI 120

121 IGFSFGALLEGAAGFGAPVAITGALLVGLGFKPLYAAGLCLIANAPVAFGALGVPILVA 180
  +GF FGA LEGAAGFGAPVAIT ALLVGLGFKPLYAAGLCLI NTAPVAFGA+G+PILVA
121 VGFCFGAFLEGAAGFGAPVAITAALLVGLGFKPLYAAGLCLIVNTAPVAFGAMGIPILVA 180

181 GQVTGIDPFHIGAMAGRQLPFLSVLVPFVWLVAMMDGWKGVKETWPAALVAGGSFAVTQFF 240
  GQVTGID F IG M GRQLPF++++V FW++A+MDGW+G+KETWPA +VAGGSFA+ Q+
181 GQVTGIDSFEIGQMVGRQLPFMTIIVLFWIMAIMDGWGRGIKETWPAVVVAGGSFAIAQYL 240

241 TSNYIGPELDPDITSALVSIVSLALFLKVVWRP 271
  +SN+IGPELPDI S+LVS++ L LFLK W+P
241 SSNFIGPELPDIISLVLCLLTLFLKRWQP 271

```

Identities = 167/257 (64%), Positives = 200/257 (77%)

```

302 YSLGQIIRAWSPFLILTTLVVTIWTMCKPFKALFAPGGAFYSLVINFPQIPHLHQQVLKAAPI 361
  Y+ GQ++RAW+PFL LT VT+W++ PFKALFA GGA Y VIN +P+L + V + P+
294 YTAGQVLRWATPFLFLTATVTLWSIPPFKALFASGGALYEWVINIPVPYLDKLVARMPPV 353

362 VAQPTPMDAVFKFDPLSAGGTAFIAAIIISIFILGVGIKKGIGVFAETLISLKWPILSIG 421
  V++ T AVFKFD SA GTAI AA++SI L + I F TL L PI SIG
354 VSEATAYA AVFKFDWFSATGTAILFAALLSIVWLKMKPSDAISTFGSTLKEALPIYSIG 413

422 MVLAFAFVTNYSGMSTTLALVLAGTGVMFPFFSPFLGWLGWVFLTGSSTSSNALFGSLQST 481
  MVLAFAF++NYSG+S+TLAL LA TG F FFSFPLGWLGWVFLTGSSTSSNALF +LQ+T
414 MVLAFAFISNYSGLSSTLALALAHTGHAFTFFSPFLGWLGWVFLTGSSTSSNALFAALQAT 473

482 TAQQINVSDDLVAANTSGGVTGKMISQSI AVACAATGMVGRESELFRTYTKHSLIFAS 541
  AQQI. VSD LLVAANT+GGVTGKMISQSI A+ACAA G+VG+ES+LFR+TVKHSLIF
474 AAQQIGVSDLLLVAANTTGGVTGKMISQSI AIAACAAVGLVKGESDLFRFTYTKHSLIFTC 533

542 VIGIITLLQAYVFTGML 558
  ++G+IT LQAYV T M+
534 IVGVITLLQAYVLTWMI 550

```

Figura 8. Apilament de la seqüència d'aminoàcids de GlcT i LctP. En la fila superior es mostra la seqüència de GlcT i en la inferior la de LctP. La seqüència intermitja no numerada mostra les posicions comunes a ambdues proteïnes i les posicions conservades (+). Els números indiquen les posicions dels aminoàcids en la proteïna.

La comparació de la seqüència aminoacídica de GlcT amb el banc de dades GenBank, dóna com a resultat una alta similitud amb proteïnes de membrana. És destacable l'alt grau de

similitud que existeix entre GlcT i la permeasa *lctP* d'*Escherichia coli* (Dong *et al.*, 1993) (Figura 8). Aquestes dues proteïnes tenen un 65.5 % d'identitat de promig i un 80% dels residus aminoacídics conservats. La similitud es conserva al llarg de tota la seqüència, excepte en una regió central (entre les posicions 271 i 302 de GlcT i 271 i 294 d'LctP).

2. ANÀLISI ESTRUCTURAL I FUNCIONAL DELS GENS DEL SISTEMA *glc*

2.1. Expressió dels gens en maxicèl.lules

La identificació dels productes dels gens que integren el locus *glc* es va dur a terme mitjançant l'expressió d'aquests a partir de diferents clons que els contenen en un sistema de maxicèl.lules. La soca JA120 va ser transformada amb diferents clons que contenen o bé la totalitat o bé diferents parts del sistema *glc*, i també com a control amb el vector de clonatge pBluescript. Els plasmidis recombinants usats van ser pTP21, pTP22, pTP23, pIM4, pTP25 i pTP26. A la Figura 9 es mostra l'autoradiografia dels productes d'expressió (en l'època en que es van realitzar aquests experiments no s'havia clonat el gen *glcT*, per la qual cosa aquest no està inclòs en els estudis d'expressió).

En el panell A de la Figura 9 es mostra l'autoradiografia de l'electroforesi dels productes d'expressió en un gel d'SDS-poliacrilamida al 10 %. En el carril 1 s'observen les proteïnes expressades a partir del plasmidi pTP21. Aquest plasmidi conté els gens *glcC*, *glcD*, *glcE*, *glcF* i *glcB* i expressa bandes proteiques de 82, 56 i 40 kDa de massa molecular. En el carril 2 es mostra l'expressió del plasmidi pTP22 (que conté *glcE*, *glcF* i *glcB*), que respecte l'anterior no expressa la banda proteica de 56 kDa, mentre que el plasmidi pTP23 (que conté *glcF* i *glcB*) únicament expressa la proteïna de 82 kDa (carril 3). El plasmidi pIM4 (conté *glcB*) mostra el mateix patró que el plasmidi pTP23 (carril 4); és a dir, pIM4 només expressa una proteïna de 82 kDa. El carril 6 mostra com el plasmidi pTP25 (conté *glcD* i *glcE*) només expressa dues proteïnes de 56 i 40 kDa però no la de 82 kDa, mentre que el plasmidi pTP26 (conté *glcD*) només expressa la proteïna de 56 kDa (carril 7).

En l'electroforesi en un gel de poliactilamida al 10 % no era possible detectar el producte de l'expressió del gen *glcG*, ja que aquest donaria una proteïna aproximadament de 13 kDa, que queda emmascarada amb el front de l'electroforesi. Per aquest motiu es va expressar el plasmidi pTP28 i els seu producte es va córrer en una gel de SDS-poliactilamida del 13.5 %. L'autoradiografia resultant es mostra en el panell B de la Figura 9. Es pot

observar com apareix una banda de 13 kDa de massa molecular aparent. Per a la identificació d'aquesta banda es va optar per usar el plasmidi pTP28 i no el pIM4, ja que en l'expressió de pIM4 podria haver interferència de l'expressió d'un possible gen truncat entre *glcB* i la diana *HindIII*.

Aquest patró d'expressió ens va permetre assignar la proteïna de 56 kDa al gen *glcD*, la proteïna de 40 kDa al gen *glcE* i la proteïna de 13 kDa al gen *glcG*. L'assignació de la proteïna de 82 kDa al gen *glcB* ja havia estat demostrada (Molina *et al*, 1994) per altres mecanismes. Les masses moleculars dels productes d'expressió dels gens són concordants amb les deduïdes de la corresponent seqüència nucleotídica.

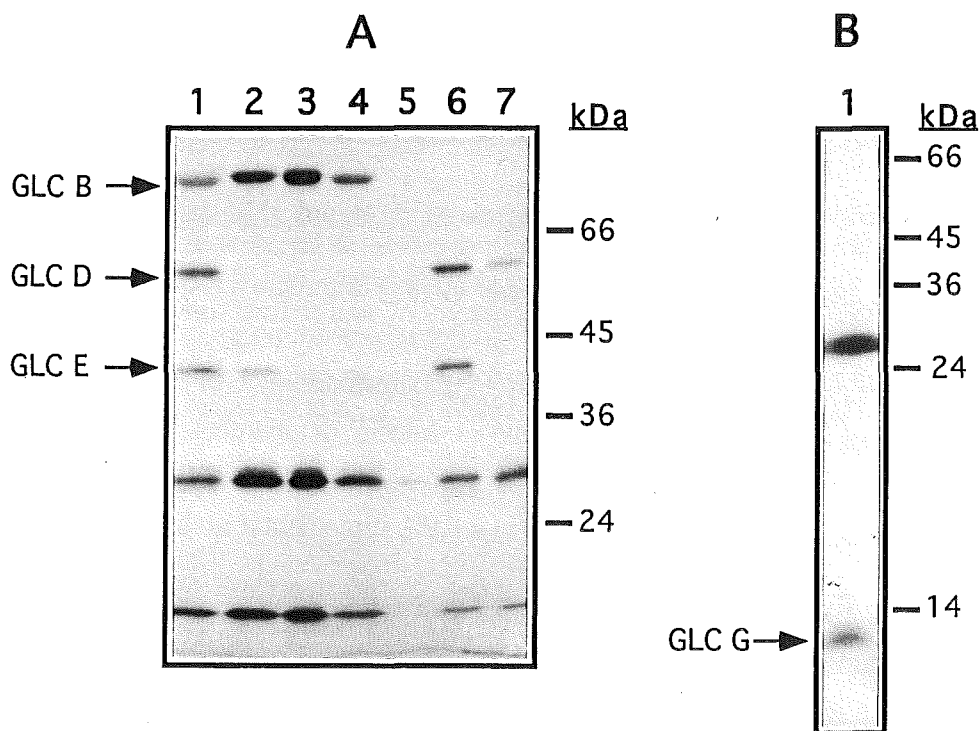


Figura 9. Autorradiografies de l'expressió dels gens del sistema *glc* mitjançant maxicèl·lules. Panell A. Autorradiografia dels productes gènics codificats en plasmidis recombinants que contenen diferents parts del locus *glc* i expressades en el sistema de maxicèl·lules usant la soca JA120 com a hoste. Les proteïnes es van separar en un gel d' SDS-poliacrilamida al 10 %. Carril 1: plasmidi pTP21; carril 2: plasmidi pTP22; carril 3: plasmidi pTP23; carril 4: plasmidi pIM4; carril 5: pBluescript; carril 6: plasmidi pTP25; carril 7: plasmidi pTP26. Panell B: Autorradiografia dels productes gènics codificats en el plasmidi pTP28 i separats en un gel d' SDS-poliacrilamida al 13.5 % (carril 1). Els productes gènics estan indicats a la part esquerra de cada panell. Els marcadors de massa molecular indicats a la dreta dels panells són albúmina sero-bovina (66 kDa), ovoalbúmina (45 kDa), gliceraldehid-3-fosfat deshidrogenasa (36 kDa), tripsinogen (24 kDa) i lactalbúmina (14 kDa).

És important assenyalar que la seqüència nucleotídica que ha estat establerta en aquest treball, i que difereix en un nucleòtid de la que ha estat descrita per Blattner *et al* (1997) dins el projecte de seqüenciació dels genoma d'*Escherichia coli*, sembla ser la correcta, ja que els marcs oberts de lectura que se'n deriven són coherents amb el resultat de l'expressió en maxicèl.lules. Segons establerta per Blattner, hauria d'expressar-se una cadena polipeptídica d'aproximadament 84 kDa de massa molecular (anomenat per aquests autors com a ORF_761) que correspondria a la fusió dels gens *glcE* i *glcF*. El resultat de l'expressió dels gens demostra que aquest ORF_761 no s'expressa.

El mètode de les maxicèl.lules per estudiar l'expressió dels gens del sistema *glc* no va permetre identificar els productes de *glcC* ni de *glcF*.

2.2. Localització subcel.lular de les subunitats de l'enzim glicolat oxidasa

Experiments duts a terme per Sallal i Nimer (1989) demostraven que la glicolat oxidasa és un enzim associat a la membrana citoplasmàtica. Per aquest motiu es va procedir a obtenir la fracció de membrana de la soca JA151 (Δ *glc recA*) transformada amb el plasmidi pTP21. Com a control de l'experiment es van obtenir també la fracció de membrana de la soca JA151 transformada amb el vector pBluescript i de la soca salvatge MC4100. Els cultius van ser obtinguts per creixement en hidrolitzat de caseïna a l'1% suplementat amb glicolat.

En la Figura 10 es mostra la separació electroforètica de les mostres obtingudes en un gel d' SDS-poliacrilamida al 10 % i visualitzades per tinció amb Coomassie Brilliant Blue. En el carril corresponent a les membranes de la soca JA151 transformada amb el plasmidi pTP21 es poden apreciar les proteïnes de 56 kDa i 40 kDa que ja havien estat detectades per l'expressió en maxicèl.lules, i a més es pot apreciar una banda proteica addicional de 45 kDa, que podria correspondre al producte gènic de *glcF*. L'alta expressió dels gens clonats en pTP21 ha facilitat la detecció d'aquestes bandes, les quals són poc visibles en la fracció de membrana de la soca tipus salvatge MC4100, d'acord amb el baix nivell d'activitat glicolat oxidasa detectat (Resultats 6.1.2.).

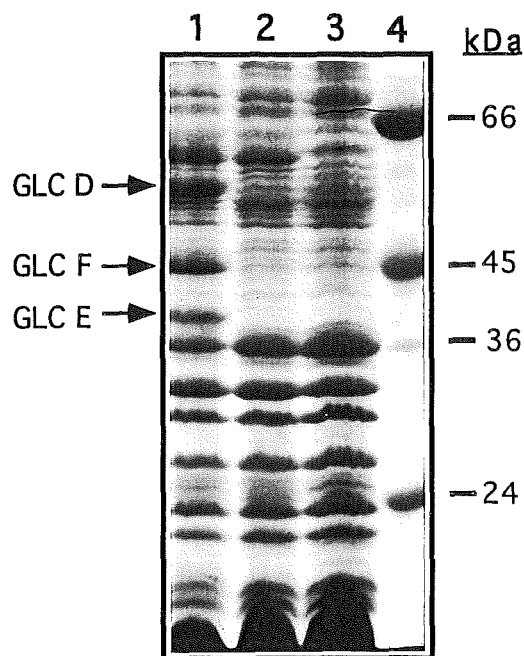


Figura 10. Electroforesi de la fracció de membrana de diferents soques. Electroforesi en gel d' SDS-poliacrilamida al 10 % de la fracció de membrana de diferents soques crescudes en hidrolitzat de caseïna 1 % en presència de glicolat. Cada carril conté 150 µg de la fracció de membrana, i estan visualitzats per tinció amb Coomassie Brilliant Blue. Carril 1: soca JA151 transformada amb el plasmidi pTP21; carril 2: soca JA151 transformada amb pBluescript; carril 3: soca salvatge MC4100. Els productes gènics estan indicats a la part esquerra de la figura. Els marcadors de massa molecular indicats a la dreta de la figura són albúmina serbovina (66 kDa), ovoalbúmina (45 kDa), gliceraldehid-3-fosfat deshidrogenasa (36 kDa), tripsinogen (24 kDa).

2.3 Anàlisi funcional dels productes gènics del locus *glc* per disrupció dels gens del sistema

Per tal de poder assignar una funció als gens seqüenciats, es va procedir a anul·lar l'expressió de tots ells individualment i veure d'aquesta manera quin efecte podria tenir sobre la utilització del glicolat en la soca MC4100. La mutagènesi dirigida per inserció d'un casset CAT19 (Fuqua, 1992), que codifica l'enzim cloramfenicol acetil transferasa i confereix resistència a cloramfenicol, va ser el mètode elegit per poder mutar específicament (Winans *et al.*, 1985) cadascun dels gens que són necessaris per complementar l'activitat GOX així com el gen *glcB* i *glcT*.

Per construir el mutants de cadascun del gens del sistema *glc*, el procediment seguit va ser el descrit anteriorment (Mètodes 1.4.2.1.). El plasmidi pTP51, usat per mutar el gen *glcC*, va ser construït partint del clon pTP26 (fragment *Bam*HI de 3.1 kb clonat en pBluescript). Aquest va ser digerit amb l'endonucleasa de restricció *Nru*I, i en ella es va clonar el casset *Sma*I-Cm^f. Seguidament el fragment mutat es va subclonar en el vector pBR322.

El plasmidi pTP52, usat per mutageneitzar el gen *glcD*, va ser construït partint del clon pTP26 (fragment *Bam*HI de 3.2 kb). Aquest clon es va digerir amb l'enzim *Mlu*I, diana en la qual es va clonar el casset *Sma*I-Cm^f, i llavors l'insert mutat va ser subclonat en el vector pBR322.

El plasmidi pTP53 va ser usat per mutar el gen *glcE*. Es va partir del clon pTP22 (fragment *Hind*III-*Sal*I de 6.3 kb clonat en pBluescript). Es va digerir amb l'enzim *Bam*HI i va ser en aquesta diana on es va introduir el casset *Bam*HI-Cm^f. L'insert mutat va ser subclonat en pBR322.

El plasmidi pTP22 va ser també el clon de partida per construir pTP54, usat per mutar el gen *glcF*. El clon pTP22 va ser digerit amb *Eco*RV, diana en la qual es va insertar el casset *Sma*I-Cm^R i l'insert mutat va ser posteriorment subclonat al vector pBR322.

Les mutacions contingudes en els plasmidis pTP51, pTP52, pTP53 i pTP54 van ser transferides al genoma de la soca JC7623 per recombinació homòloga, prèvia linialització dels plasmidis amb l'enzim de restricció *Cla*I.

Per mutar el gen *glcG*, es van fer dues construccions. Una d'elles tenia el casset CAT19 a la diana de restricció *Nru*I (plasmidi pTP55). La interrupció de la proteïna es produeix a només 9 aminoàcids de l'extrem C-terminal. El plasmidi pTP55 va ser construït a partir del clon pTP28. L'insert *Cla*I-*Kpn*I de 2.8 kb del plasmidi pTP28 va ser subclonat al vector pBR322. Aquest plasmidi recombinant va ser digerit amb l'enzim *Nru*I, on es va insertar el casset *Sma*I-Cm^f. L'altra construcció per mutar el gen *glcG* va ser el clon pTP01, que es va construir partint d'un plasmidi que conté el fragment *Cla*I-*Sal*I de 4.4 kb clonat en pBR322. Aquest es va digerir amb l'enzim de restricció *Age*I i en aquesta diana es va insertar el casset *Sma*I-Cm^f. En aquesta construcció el casset CAT19 interromp el marc de lectura de *GlcG* a 35 aminoàcids de l'extrem C-terminal. Posteriorment, els clons pTP55 i pTP01 van ser digerits amb *Cla*I i introduïts al genoma de la soca JC7623.

Per dur a terme la mutagènesi de *glcB* es va partir del clon pTP21. El seu insert es va subclonar en el vector pBR322. El subclon derivat es va digerir amb *Cla*I, on es va insertar el casset *Sma*I-Cm^f. La digestió amb un enzim del polylinker del vector de clonatge va permetre linialitzar el clon pTP59 i transferir-lo així a genoma.

Finalment, la construcció del mutant *glcT* va dur-se a terme mitjançant el clon pTP02. Aquest va ser construït partint del clon pTP31, que conté tota la seqüència codificant de *glcT*. Aquest clon es va digerir amb *HindIII*, on es va insertar el casset *SmaI*-Cm^r. La digestió amb *ClaI* va permetre transferir-lo a genoma.

Cadascuna d'aquestes mutacions presents al genoma de la soca JC7623, va ser transferida al genoma de la soca salvatge MC4100 per transducció generalitzada amb bacteriòfag P1. Les soques mutants derivades, seleccionades per l'adquisició de resistència a cloramfenicol, van anomenar-se: JA154 (*glcC::cat*), JA155 (*glcD::cat*), JA156 (*glcE::cat*), JA157 (*glcF::cat*), JA158 (*glcG::cat* a *NruI*), TP01 (*glcG::cat* a *AgeI*), JA159 (*glcB::cat*) i TP02 (*glcT::cat*). La correcta inserció dels cassets en el genoma es va comprovar per transducció generalitzada amb bacteriòfag P1 i/o per PCR. A la Figura 11 es mostra la localització dels cassets CAT19 que interrompen els diferents gens del sistema *glc* en el genoma de la soca MC4100.

Per comprovar que el fenotip de les soques mutants per inserció del casset CAT19 era causat per la interrupció d'aquell gen i no a un efecte polar del casset sobre els gens posteriors, es va procedir a transformar les soques mutants amb plasmidis que contenien exclusivament la subunitat en qüestió. Totes les activitats abolides eren restaurades amb els plasmidis que contenen el gen mutat intacte.

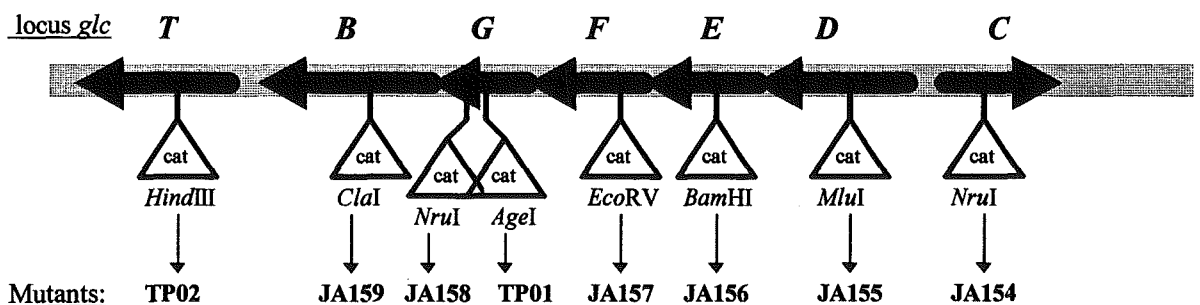


Figura 11. Mapa físic de la mutagènesi dirigida dels gens del locus *glc* per inserció del casset Cm^r en el genoma de la soca MC4100. Les dianes de restricció usades per la disrupció dels gens s'indiquen per cada inserció així com el nom de la soca mutant que se'n deriva.

Els resultats de la mutagènesi dirigida (Taula 1) van permetre veure que el mutant del gen *glcC*, la soca JA154, perdia les dues activitats codificades en el locus *glc*, la malat sintasa G

i la glicolat oxidasa, i mostrava un fenotip glicolat negatiu. Aquests resultats semblen indicar que aquest gen codifica la proteïna activadora del sistema *glc*.

Els resultats de la interrupció dels gens *glcD*, *glcE*, *glcF* mostraven que els mutants havien perdut l'activitat GOX i eren incapaços de créixer en glicolat com a única font de carboni i energia . En els extractes de la soca JA159 (*glcB::cat*) no es podia detectar activitat MSG, mentre que l'expressió de GOX era normal. La deficiència en MSG no impedia a la soca JA154 créixer en glicolat com a única font de carboni i energia ja que la seva funció era suplida per l'isoenzim malat sintasa A (Resultats 6.1.1.).

Taula 1. Característiques de les soques mutants del sistema *glc* obtingudes per inserció d'un casset de resistència a Cm.

Soca	Gen mutat	Fenotip glicolat	Activitats abolides	Complementada per
JA154	<i>glcC</i>	-	MSG, GOX	pTP30
JA155	<i>glcD</i>	-	GOX	pTP26
JA156	<i>glcE</i>	-	GOX	pTP25
JA157	<i>glcF</i>	-	GOX	pTP28
JA158	<i>glcG (NruI)</i>	+	cap	-
TP01	<i>glcG(AgeI)</i>	+	cap	-
JA159	<i>glcB</i>	+	MSG	pIM4
TP02	<i>glcT</i>	-/+	cap	-

En els extractes cel.lulars de la soca JA158 (*glcG* mutat per inserció del casset CAT en la diana *NruI*) es van poder determinar les activitats GOX i MSG, indicant que el producte d'aquest gen, GlcG, no sembla formar part de cap de les dues activitats enzimàtiques codificades en el locus *glc*. A més el fenotip en glicolat era clarament positiu. Com que la inserció en la diana *NruI* només interromp la proteïna en els 9 aminoàcids de l'extrem carboxiterminal, es va pensar que potser aquesta mutació no inactivava la proteïna, i per tant donava lloc a una proteïna encara funcional. Per aquest motiu es va procedir a la construcció de la soca TP01. El gen mutat en aquest cas és també *glcG*, però en la diana *AgeI*. Aquesta inserció provoca la pèrdua de 35 aminoàcids de l'extrem carboxiterminal de la proteïna corresponent. En aquest cas el comportament de la soca va ser idèntic al de la soca JA158, la qual cosa confirmava que GlcG no sembla formar part de l'enzim GOX.

La inserció del casset CAT a *glcT* no aboleix cap activitat enzimàtica però enlenteix el creixement en glicolat.

3. ORGANITZACIÓ GÈNICA DEL LOCUS *glc*

3.1. Anàlisi de l'organització transcripcional mitjançant Northern Blot

Per tal de poder estudiar tant el patró d'expressió com la longitud dels transcrits derivats del sistema *glc*, es va obtenir RNA total de la soca salvatge MC4100 tal i com s'ha descrit anteriorment (Mètodes 5.1.) a partir de creixements en diferents condicions d'inducció.

Els creixements de la soca salvatge van dur-se a terme en hidrolitzat de caseïna al 0.5%, hidrolitzat de caseïna al 0.5% suplementat amb glicolat 30mM, hidrolitzat de caseïna al 0.5% suplementat amb glioxilat 30mM i en medi mineral SM amb glicolat 30mM com a única font de carboni i energia.

Com a sondes per poder identificar els transcrits es van usar fragments de DNA interns a cada marc obert de lectura contingut en el locus *glc*. Aquestes sondes van ser marcades radiactivament pel mètode de "random primed" (Mètodes 3.3.2.1.).

Per identificar el transcrit de *glcC* es va utilitzar la sonda S₁, de 500 pb de longitud i interna a *glcC*, que es va obtenir per digestió del clon pTP26 amb els enzims de restricció *EcoRV* i *NruI*. Les sondes usades per estudiar l'expressió dels gens estructurals són les que es descriuen a continuació. La sonda S₂ és un fragment de 800 pb intern a *glcD* obtingut per digestió amb l'enzim *SalI*. La sonda S₃, interna a *glcE*, té 420 pb i és el resultat de la digestió amb els enzims *AvaII* i *EcoRI* del clon pTP27. La sonda S₄ és interna a *glcF*; té 750 pb de longitud i es va obtenir per digestió dels clon pTP28 amb *EcoRV* i *NdeI*. La sonda S₅, interna a *glcB*, és un fragment de 400 pb obtingut a partir del clon pTP21 per digestió amb *AvaII* i *ClaI*. La sonda S₆, interna a *glcT*, és un fragment de 1100 pb obtingut en la digestió del clon pTP31 amb *BstXI* i *PstI*.

El resultat de la hibridació amb les diferents sondes de les preparacions d'RNA total, immobilitzades en filtres de nitrocel·lulosa, obtingudes a partir de la soca salvatge MC4100 creixent en diverses fonts de carboni va permetre observar dos patrons d'expressió clarament

diferenciats (Figura 12). D'una banda, tots els gens estructurals del sistema *glc* van donar un patró d'inducció molt clar: mentre que el transcrit era difícilment detectable en creixements en hidrolitzat de caseïna, aquest era molt aparent quan les cèl.lules havien crescut en hidrolitzat de caseïna més glioxilat, més glicolat o en medi mínim més glicolat. En aquestes condicions d'inducció, els tamany dels transcrits detectats per cada gen era igual en totes elles, tot i que la intensitat (els RNAs estan normalitzats) era molt superior en el cas dels creixements on el glicolat era l'única font de carboni. El nivell d'expressió dels gens estructurals assolit en presència de glioxilat com a inductor era més baix que el produït per glicolat. Aquest resultat concorda amb els nivells d'activitats enzimàtiques MSG i GOX determinats en els extractes cel.lulars de la soca MC4100 crescuda en presència de glicolat o de glioxilat (Resultats 6.1.2.).

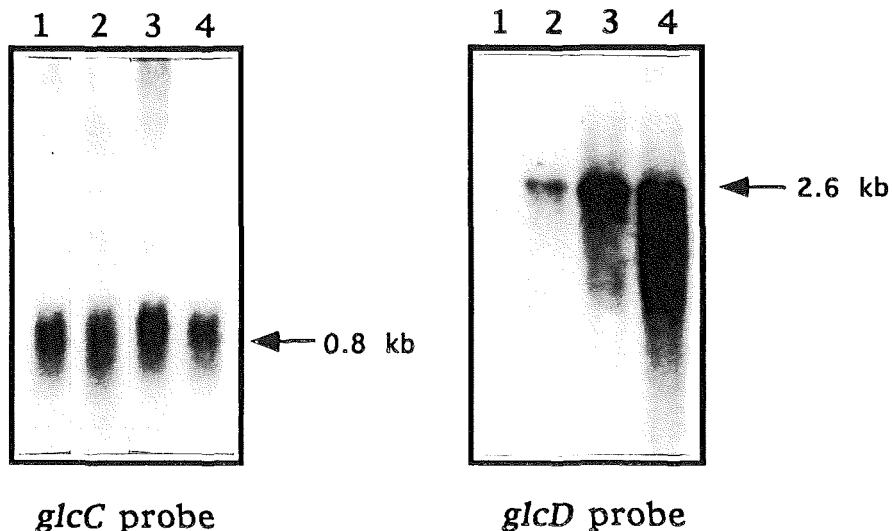


Figura 12. Northern Blots d'RNA total de la soca MC4100 creixent en diferents condicions amb les sondes S_1 i S_2 . Es van sembrar 9 μ g d'RNA obtingut per creixement de la soca MC4100 en: hidrolitzat de caseïna (carrils 1), hidrolitzat de caseïna més glioxilat 30 mM (carrils 2), hidrolitzat de caseïna més glicolat (carrils 3) i medi mínim SM més glicolat (carrils 4). Les hibridacions es van dur a terme amb la sonda S_1 , corresponent al gen *glcC* (*glcC* probe) i com a mostra representativa de l'expressió dels gens estructurals, la sonda S_2 , interna a *glcD* (*glcD* probe).

El tamany dels transcrits de *glcD*, *glcE*, *glcF*, *glcG*, *glcB* i *glcT* no corresponia al tamany individual deduït de la seqüència nucleotídica de cada gen. El tamany observat era més gran que el del transcrit independent, però menor que el d'un transcrit policistrònic, únic per tots els gens estructurals. Per aquest motiu la informació obtinguda a partir del tamany dels transcrits no permetia saber quina és l'estructura gènica del sistema *glc*, i únicament mostrava el patró d'inducció.

Pel que fa al transcrit de *glcC*, aquest mostrava un patró d'inducció diferent al dels gens estructurals. Els RNAs obtinguts de cèl·lules creixent en presència d'hidrolitzat de caseïna mostraven un transcrit més intens que en els altres tres creixements, tot i que en els darrers casos sempre era perceptible un transcrit del mateix tamany. El tamany del transcrit concorda amb el del tamany de l'mRNA deduït de la seva seqüència de nucleòtids. Aquest resultat mostrava que el regulador del sistema *glc* s'expressa com a únic transcrit, independent de la resta de gens estructurals del locus i que la seva expressió pot ser considerada com a basal i no induïble.

3.2. Anàlisi de l'organització transcripcional mitjançant fusions de promotor al gen *lacZ*. Identificació de promotors funcionals

L'estudi de l'organització gènica del sistema *glc* mitjançant Northern Blot únicament havia reportat informació sobre el patró d'inducció, però no havia permès saber si els gens estructurals d'aquest sistema estan organitzats en diverses unitats transcripcionals o bé estan sota el control d'un únic promotor funcional que regula la seva expressió, constituint un operó.

Per tal de poder saber si cadascun dels gens del sistema *glc* té el seu propi promotor, si alguns d'ells comparteixen promotor amb d'altres o si tots ells estan sota el control d'un únic promotor que els regula globalment a tots ells, es va procedir a construir fusions transcripcionals de la seqüència situada en posició 5' en cadascun dels gens. Aquestes seqüències es van fusionar al gen marcador *lacZ*, que codifica l'enzim β -galactosidasa.

El protocol seguit per construir les diferents fusions transcripcionals va ser el següent. Després d'obtenir-se els fragments dels quals es volia assajar l'activitat promotora, aquests van fer-se roms. D'aquesta manera van ser clonats en el vector pBluescript, en la seva diana *SmaI*. Un cop comprovada l'orientació de l'insert respecte el vector (mitjançant seqüenciació del clon recombinant) aquest va ser alliberat per digestió amb *EcoRI* i *BamHI* i posteriorment subclonat en les mateixes dianes del vector pRS550 o pRS551 segons la seva orientació. La incorporació de les fusions obtingudes al genoma, primerament de la soca TE2680, i posteriorment de la soca MC4100 es va fer seguint els protocols descrits anteriorment (Mètodes 4.1. i 1.2. respectivament).

Per construir la fusió de *glcC* es va usar el fragment *SaII* de 400 pb, que conté els 296 nucleòtids per davant el codó ATG. El mateix fragment, clonat en la direcció oposada, va ser fusionat al gen *lacZ* per tal d'assajar el promotor de *glcD*. En aquesta direcció, hi ha 360 pb

per davant de l'ATG que obra el marc de lectura corresponent al primer gen estructural (Figura 13).

El fragment de 990 pb comprès entre les dianes *AvaII* i *SalI* va ser usat per determinar la possible activitat promotora del gen *glcE*, i conté 640 nucleòtids per davant de l'ATG. Per construir la fusió del promotor de *glcF*, es va obtenir el fragment *EcoRV-BamHI* de 900 pb, que conté 422 nucleòtids per davant del codó d'inici de la traducció (Figura 13).

La construcció de la fusió de la regió 5' de *glcG* es va fer amb el fragment *NdeI-AgeI* de 480 pb, que deixa 189 nucleòtids per davant de l'ATG. El promotor de *glcB* es va construir amb el fragment *ClaI-DdeI* de 930 nucleòtids. Aquest fragment conté 215 nucleòtids per davant de l'ATG inicial. Per últim, el promotor de *glcT* va construir-se per clonació del fragment *BstXI-NsiI* de 1400 pb. Aquest comprèn els 235 nucleòtids davant l'ATG (Figura 13).

Un cop es van tenir totes les fusions de promotor clonades i transferides en els fons genètic de la soca MC4100, es va procedir a assajar l'activitat promotora per mesura de l'activitat β -galactosidasa en cultius d'hidrolitzat de caseïna i en el mateix medi complementat amb l'inductor glicolat.

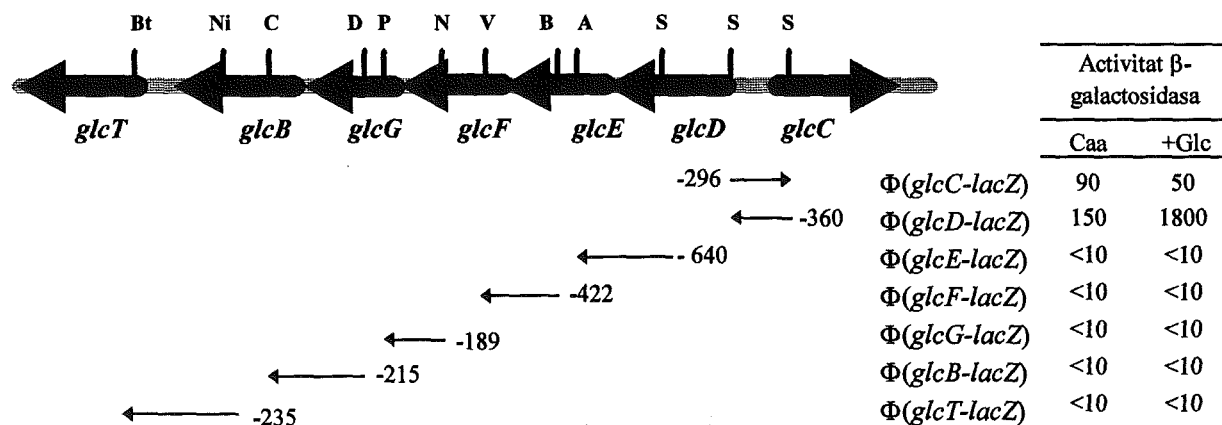


Figura 13. Fragments usats en la construcció de les fusions de promotor dels gens del sistema *glc*. Es pot observar el mapa físic de la regió cromosòmica que conté els gens del sistema *glc* i les dianes de restricció usades per obtenir els possibles promotors. A costat de cada fragment s'indica la longitud en nucleòtids per davant de l'ATG. La taula, a la dreta del dibuix, mostra els valors d'activitat β -galactosidasa, expressats en unitats Miller, obtinguts a partir de creixements de la soca MC4100 contenint cadascuna de les fusions, en hidrolitzat de caseïna al 0.5 % (Caa) i també suplementada amb glicolat 30mM (+Glc). A: *AvaII*; B: *BamHI*; Bt: *BstXI*; C: *ClaI*; D: *DdeI*; N: *NruI*; Ni: *NsiI*; P: *PstI*; S: *SalI*; V: *EcoRV*.

El resultat de l'assaig de les activitats β -galactosidasa de les diferents fusions transcripcionals va permetre observar que:

i) *glcC* disposa d'un promotor feble, que s'expressa tant en hidrolitzat de caseïna com en presència de glicolat. Quan el medi està complementat amb glicolat, l'expressió de *glcC* disminueix a valors aproximadament la meitat dels observats en hidrolitzat de caseïna. Aquest tipus d'expressió baixa i poc induïble és típica de promotors de gens reguladors (Figura 13), i

ii) només existeix un promotor funcional induïble per tots els gens estructurals del sistema *glc*. Aquest està situat en posició 5' de gen *glcD*. Aquest promotor és induïble per l'addició de glicolat en el medi, tot i que en hidrolitzat de caseïna es pot observar una expressió aproximadament 10 vegades inferior (Figura 13).

D'aquestes dades es pot deduir que el sistema *glc* està organitzat en dues unitats transcripcionals, una que correspon al gen regulador i l'altra als gens estructurals, els quals s'estructuren en forma d'operó (*glcDEFGBT*). Un únic promotor regula l'expressió tant dels gens que codifiquen la glicolat oxidasa, com de la malat sintasa G, com del possible gen del transport.

3.3. Identificació de l'inici de transcripció de les dues unitats transcripcionals del sistema *glc*

Un cop determinada l'existència de dues unitats transcripcionals es va procedir a trobar els seus inicis de transcripció, per poder localitzar els elements reguladors en la seqüència del promotor.

La tècnica usada va ser la de "primer extension". El protocol usat per dur a terme la determinació de l'inici de transcripció o tsp és el que ha estat especificat anteriorment (Mètodes 5.3.).

Es va obtenir RNA total de la soca MC4100 crescuda en hidrolitzat de caseïna en presència de glicolat. Aquest RNA va usar-se tant en la reacció de "primer extension" per determinar el tsp de *glcD* com de *glcC*.

A

TTCATCTTTTCATTCCCGGACCTCGTGCACAGGTAGGACCAATTTTTCTGTTCTTAACTCGCAAAACACGC -138
 AAGTAGAAAGTAAGGGCCTGGAGCACGTGTCCATCCTGGTTAAAAAGACAAGAATTGAGCGTTTTGTGCG

ACATCACGTAAGTGATTATGTTACATCAATTTAACATTGAGTTAACCAAGACAAGGTCACAGAGCTGGAA -68
 TGTAGTGCATTACTAATACAATGTAGTTAAATTGTAACTTCAATTGGTTCTGTTCCAGTGTCTCGACCTT
 ← *glcC* mRNA -10 -35 *glcDEFGBT*mRNA →

AAAAATGGTCTGACCGGTAGCTAAAGAGATAGACGAAAACGAAAAGCCCGCTTAATAACTGTTACAGAA +2
 TTTTTTACCAGACTGGCCATCGATTTCTCTATCTGCTTTTGGCTTTTCGGGCGAATTATTGACAAGTGTCT

AGCAGCGCGCAAAAATCAGCTGCCACACAACAACAAGCGAAGCCTACTCATGAG +59
 TCGTCGCGGTTTTTAGTCGACGGTGTGTTGTGTTGTTTCGCTTCGGATGAGTACTC

B

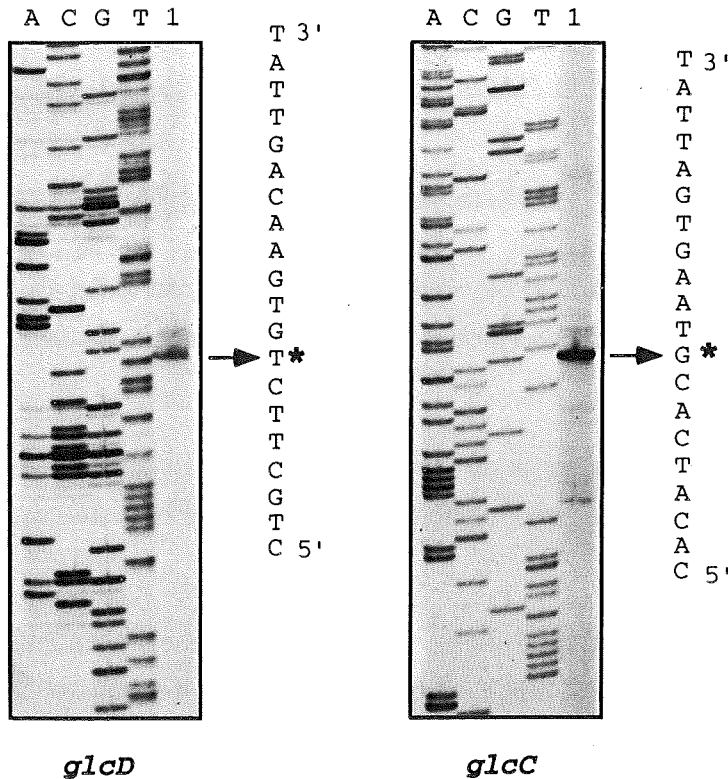


Figura 14. Seqüència de la regió que comprèn els promotors divergents de *glcD* i *glcC* i mapeig de l'extrem 5' dels mRNAs derivats. Panell A. Es mostra la seqüència dels promotors de *glcD* i *glcC*. Estan indicades les caixes -10 i -35 de cada gen (subratllades i en negra), el tsp de cada unitat transcripcional

(indicats per triangles negres), els ATG inicials (subratllats) i les seqüències Shine-Dalgarno (en un rectangle). Les posicions de la seqüència estan indicades respecte el tsp de *glcD*. Panell B. Es mostra l'autoradiografia dels productes de la "primer extension". En els carrils 1 es mostren els productes de l'extensió. Els carrils A, C, G, i T mostren la seqüència obtinguda sobre el clon pTP26, usant com a primer el mateix oligonucleòtid emprat en cada extensió sobre l'mRNA.

Per la determinació de l'inici de transcripció de *glcD* es va utilitzar el primer (5'-GGTCGACATCGGGTAAAGC-3') complementari a una regió codificant del gen (posicions + 52 a + 33). El tsp es va localitzar a 55 pb per davant del codó ATG d'inici de la traducció. Aquesta posició va ser designada com a + 1, i a partir d'ara totes les posicions es donaran en referència a aquesta. L'inici de transcripció està precedit per unes seqüències promotores que són semblants al consens descrit per la unió de la subunitat σ^{70} de l'RNA polimerasa. L'hexàmer en posició -10 té una seqüència TAATAA i el situat a -35, TAGACG. Ambdós hexàmers estan separats per 17 pb (Figura 14).

El tsp de *glcC* es va determinar també pel mateix mètode, usant un primer complementari a una regió codificant del gen (5'-TGACCGACCTTCAGTACGCCG-3'). Inclou 21 pb situats a 62 pb de l'ATG en direcció 3' a *glcC*. El tsp es va localitzar a 66 pb per davant del codó ATG d'inici de la traducció. Les caixes -10 i -35 estan relativament conservades respecte el consens, com correspon a una baixa activitat promotora de gens reguladors. L'hexàmer -10 té una seqüència CATAAT i el -35 TTAACT. Estan separats per 19 pb (Figura 14).

3.4. Anàlisi funcional del promotor de l'operó *glc* per deleccions de l'extrem 5'

Un cop identificat que l'únic promotor funcional dels gens estructurals del locus *glc* és el localitzat en posició 5' a *glcD*, es va procedir a fer una anàlisi d'aquest promotor mitjançant deleccions del mateix. Amb aquest estudi es pretenia conèixer els elements en *cis* necessaris per la regulació de l'expressió d'aquest promotor.

Les deleccions es van construir de manera que la zona 3' dels fragments fós constant i només es van anar eliminant seqüències de l'extrem 5'. Per dur a terme aquestes construccions es va utilitzar la tècnica de PCR. El DNA motllo en totes les deleccions era el clon pTP26. El "primer" de la zona 3' del promotor de totes les construccions, era: 5'-GGGGAATTCTAAAGCGCCATCAAGACG-3', que és complementari a les posicions + 77 a + 94 de *glcD*. A l'extrem 5' d'aquest oligonucleòtid existeix una diana de restricció *EcoRI* (subratllada), la qual permetia el clonatge dirigit en el vector de fusió transcripcional pRS550.

Els diferents oligonucleòtids de l'extrem 5' estaven dissenyats amb un 100 % d'homologia respecte a la seqüència a la qual primaven i també tenien en el seu extrem 5' una diana de restricció. En aquest cas, la diana utilitzada era la diana *Bam*HI, dissenyada amb el mateix propòsit de permetre el clonatge dirigit en el vector pRS550. Es van construir d'aquesta manera 10 delecions del promotor de *glcD*, amb oligonucleòtids de 18 pb més la cua de l'extrem 5' amb la diana *Bam*HI (subratllada). Els primers usats per construir cadascuna de les delecions tenien tots l'estructura: 5'-NNGGATCC...18 pb...-3'. Cada delecio s'ha anomenat indicant la posició del seu extrem 5' en relació al tsp del gen *glcD* (p.e. delecio f-247: l'extrem 5' de la delecio es troba a la posicio -247 del gen *glcD*). Les diferents delecions van ser introduïdes en el cromosoma de la soca salvatge MC4100, on es van assajar les activitats β -galactosidasa derivades.

En la Figura 15 es mostren les diferents delecions del promotor i l'activitat β -galactosidasa assajada per cadascuna d'elles en creixements en hidrolitzat de caseïna i en presència de glicolat.

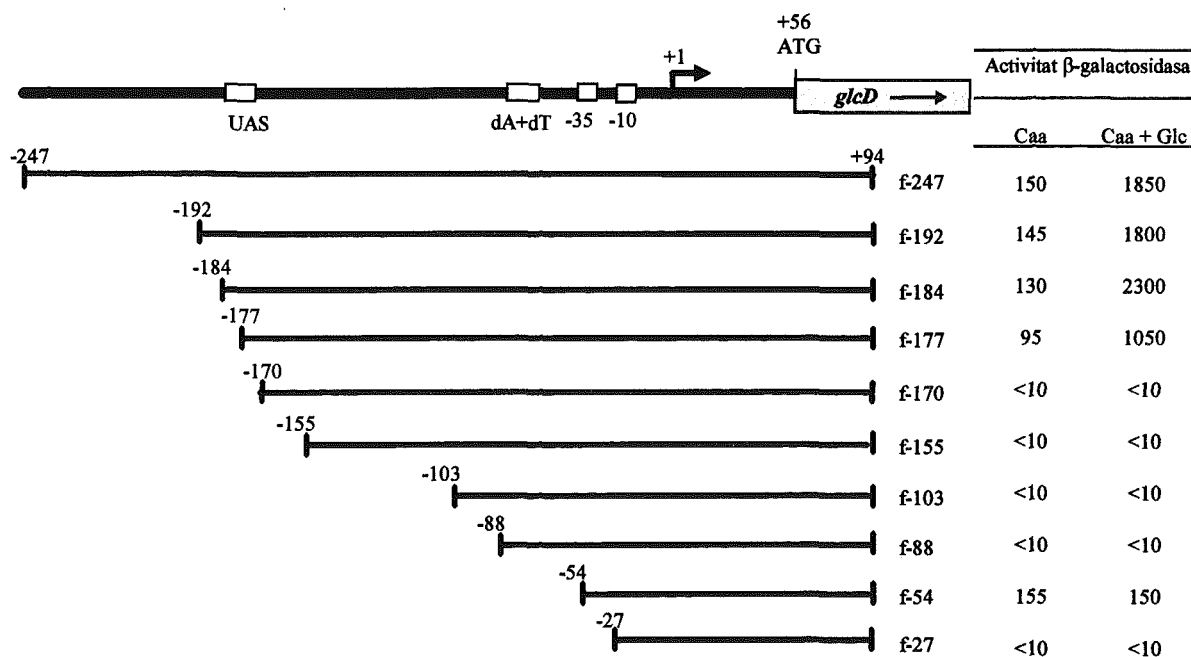


Figura 15. Delecions del promotor de *glcD*. En la part superior de la figura es mostren les seqüències promotores -10 i -35 així com els elements reguladors de la transcripció deduïts a partir de l'anàlisi de les delecions. A sota es mostren els fragments corresponents a cada delecio junt amb el nom de cada construccio. La taula de la figura mostra les activitats β -galactosidasa (expressades en unitats Miller) de la soca salvatge MC4100, que conté les delecions fusionades al gen *lacZ* en forma d'unicopia en el genoma, assajades en cultius d'hidrolitzat de caseïna (Caa) i també en presència de glicolat 30 mM (Caa +Glc). UAS: upstream activator site.

De l'anàlisi de les diferents deleccions del promotor es va observar que per aconseguir una inducció total de l'operó eren necessàries les seqüències promotores compreses fins la posició -184. La construcció f-177 mostrava uns nivells d'activitat reduïts en un 40%, mentre que la construcció f-170 abolia completament l'expressió de *glcD*. Sembla, doncs, que entre les posicions -184 i -177 existeix una possible seqüència activadora, que es designa com UAS (upstream activator site). La localització d'aquesta UAS entre el tsp i el codó ATG d'inici de la traducció del gen regulador *glcC* pot explicar l'observació que l'expressió de *glcC* està reduïda en condicions d'inducció en presència de glicolat. Sembla que quan GlcC està activant l'expressió dels gens estructurals, alhora està exercint una autorepressió de la seva pròpia expressió.

La construcció f-54 mostra una expressió basal no induïble. Aquesta apareix quan s'elimina una zona en el promotor rica en dA+dT. La construcció f-27, que ja ha perdut la caixa -35 no expressa activitat β -galactosidasa, demostrant la funcionalitat de l'hexàmer en l'inici de la transcripció.

4. FACTORS TRANSCRIPCIONALS QUE REGULEN L'EXPRESSIÓ DEL SISTEMA GÈNIC *glc*

4.1. Factors de transcripció específics del sistema *glc*. Paper de la proteïna GlcC

El fet que el producte del gen *glcC* tingui una alta similitud amb proteïnes reguladores de procariotes, i que mutacions en aquest gen anul·lin l'expressió de les dues activitats enzimàtiques codificades en aquest locus (glicolat oxidasa i malat sintasaG) fa pensar que la proteïna GlcC pot ser el regulador transcripcional de l'operó *glc*.

Per tal de corroborar la funció d'aquesta proteïna es van dur a terme estudis de l'expressió de la fusió transcripcional $\Phi(glcD-lacZ)$ en mutants *glcC*, estudis de l'efecte dosi de la proteïna i retardacions de la mobilitat electroforètica del promotor amb extractes cel·lulars de diferents soques.

4.1.1. Expressió en mutants *glcC*

El mutant *glcC* usat en aquest treball és la soca JA154. Aquest mutant, isogènic amb la soca tipus salvatge MC4100, conté una disrupció en el gen *glcC* per inserció d'un casset CAT (Resultats 2.3.). La seva mutació pot ser transferida a qualsevol altre fons genètic mitjançant transducció amb P1 i és seleccionable per la resistència a cloramfenicol.

Sobre la soca JA154 es van transferir les dues fusions de promotor actives de l'operó *glc*: les fusions $\Phi(\textit{glcD-lacZ})$ i $\Phi(\textit{glcC-lacZ})$. Per tal d'analitzar l'efecte de la mutació *glcC* en l'expressió dels gens estructurals o sobre la seva pròpia expressió, es va procedir a analitzar aquestes fusions per mesura de l'activitat β -galactosidasa en creixements en hidrolitzat de caseïna o en presència de glicolat. L'anàlisi de les mateixes fusions en el fons genètic de la soca parental MC4100 es va realitzar en paral·lel com a control. Els resultats dels assajos es mostren a la Taula 2.

Taula 2. Efecte de mutacions en *glcC* en l'expressió de les fusions $\Phi(\textit{glcD-lacZ})$ i $\Phi(\textit{glcC-lacZ})$. Els creixements van ser fets en hidrolitzat de caseïna al 0.5 % (Caa) i en hidrolitzat de caseïna al 0.5 % en presència de glicolat 30 mM (Caa + Glc). Les activitats estan expressades en unitats Miller.

Soca	Fusió	Activitat β -galactosidasa	
		Caa	Caa + Glc
MC4100	$\Phi(\textit{glcD-lacZ})$	150	1800
JA154	$\Phi(\textit{glcD-lacZ})$	<10	<10
MC4100	$\Phi(\textit{glcC-lacZ})$	90	50
JA154	$\Phi(\textit{glcC-lacZ})$	110	100

L'expressió de la fusió $\Phi(\textit{glcD-lacZ})$ és totalment dependent del producte gènic GlcC. Es pot observar que els nivells d'activitat β -galactosidasa expressats eren indetectables tant en condicions no inductores com inductores. Aquesta deficiència era revertida quan la soca JA154 que portava la fusió $\Phi(\textit{glcD-lacZ})$ era transformada amb el plasmidi pTP26 (inclou el gen *glcC*) (Resultats 1.1.), demostrant que l'absència d'expressió era causada per una manca de GlcC. Sembla doncs que GlcC és un activador transcripcional dels gens estructurals de l'operó *glc*. La inducció específica del sistema, observada en condicions de creixement en presència de glicolat, sembla ser conduïda per GlcC.

L'efecte de mutacions en el gen *glcC* sobre l'expressió de la fusió $\Phi(\textit{glcC-lacZ})$ és contrari. Els mutants en la proteïna reguladora mostren una activitat del promotor *glcC* elevada. Els nivells d'activitat determinats són del mateix ordre que els observats en la soca salvatge en absència de glicolat. Aquest resultat sembla indicar que GlcC és un repressor de la seva pròpia expressió.

4.1.2. Efecte dosi de la proteïna GlcC sobre l'expressió de $\Phi(\textit{glcD-lacZ})$

Per analitzar l'efecte dosi de la proteïna GlcC sobre l'expressió dels gens estructurals de l'operó *glc* es van fer dues sèries d'experiments. D'una banda, es va introduir en la soca MC4100 $\Phi(\textit{glcD-lacZ})$ el plasmidi pTP26 que té clonat el gen *glcC* (Resultats 1.1.) i que expressa la proteïna GlcC en un gran nombre de còpies. Aquest experiment també es va fer sobre la soca MC4100 $\Phi(\textit{glcC-lacZ})$. Els creixements es van dur a terme en hidrolitzat de caseïna a l'1 % i en hidrolitzat de caseïna a l'1 % més glicolat 30 mM. Es va poder comprovar que l'expressió de la fusió $\Phi(\textit{glcD-lacZ})$, quan la soca està transformada amb pTP26, està augmentada. Aquesta activació de la transcripció causada pel gran nombre de molècules de GlcC és independent de la presència de glicolat, ja que en hidrolitzat de caseïna l'activitat β -galactosidasa de la soca MC4100 $\Phi(\textit{glcD-lacZ})$ transformada assoleix els mateixos valors d'inducció per glicolat que la soca sense transformar. Aquesta observació confirma el paper de la proteïna GlcC com a activador transcripcional del promotor de *glcD*.

L'efecte de la sobreexpressió de GlcC en l'activitat β -galactosidasa de la soca MC4100 $\Phi(\textit{glcC-lacZ})$ era contrària que per la fusió dels gens estructurals. S'observa una disminució del 40 % de l'expressió quan la soca salvatge que duu la fusió $\Phi(\textit{glcC-lacZ})$ es transforma amb pTP26.

L'altra manera d'abordar l'estudi de l'efecte dosi de la proteïna GlcC van ser els estudis de competició, on es va assajar la variació de l'activitat β -galactosidasa en la soca MC4100 $\Phi(\textit{glcD-lacZ})$ quan aquesta era transformada amb diferents plasmidis, que contenien fragments que abarquen diferents regions dels promotor de *glcD*. Tots ells van clonar-se en el vector multicòpia pBluescript, per aconseguir tenir un gran nombre de còpies dels fragments en el citosol cel.lular. Els fragments usats comprenen les posicions -247 a +94 (fragment G), de -177 a +94 (fragment G1), de -170 a +94 (fragment G2), de -88 a +94 (fragment G3), de -54 a +94 (fragment G4), de -247 a -134 (fragment P) i de -155 a +94 (fragment M). La soca MC4100 $\Phi(\textit{glcD-lacZ})$ va ser transformada amb aquests plasmidis (i amb el pBluescript com a

control) i es va assajar l'activitat β -galactosidasa en creixements d'hidrolitzat de caseïna a l'1 % en presència de glicolat (Taula 3).

L'activitat β -galactosidasa disminuïa a nivells del 20% respecte el control quan la soca assajada era transformada amb el fragment G i G1. Quan la soca contenia els fragments G2, G3, G4, P i M, els nivells d'activitat genòmica eren del 70-80 % dels valors control.

Aquesta observació està d'acord amb els resultats de l'anàlisi de delecions i indica que en la zona compresa entre les posicions -247 i -170 es localitzen els elements en *cis* necessaris per l'activació de *glcD* (UAS), i que pot ser la zona on s'uneix la proteïna GlcC. La presència d'aquesta seqüència en alt nombre de còpies segresta proteïna GlcC, fet que es tradueix en una devallada de l'expressió de la fusió genòmica $\Phi(glcD-lacZ)$. El fet que el fragment P no sembla segrestar proteïna tot i contenir la possible zona UAS, indica que les seqüències situades a 3' de la posició -134 són necessàries per la unió correcta de la proteïna GlcC.

Taula 3. Efecte del nombre de còpies de fragments del promotor de *glcD* en l'expressió de la fusió $\Phi(glcD-lacZ)$ en el fons genètic de la soca MC4100. Els creixements es van dur a terme en hidrolitzat de caseïna a l'1% en presència de glicolat 30 mM a 30°C. Les activitats es mostren en % respecte l'activitat de la soca MC4100 $\Phi(glcD-lacZ)$ sense transformar. pBS: pBluescript. β -gal: β -galactosidasa.

Fragment	-	pBS	G	G1	G2	G3	G4	P	M
Posicions	-	-	-247/+94	-177/+94	-170/+94	-88/+94	-54/+94	-247/-134	-155/+94
Activitat									
β -gal (%)	100	100	21	18.5	71	79	85	81	79

4.1.3. Estudis de retardació de la mobilitat electroforètica amb extractes crus de la soca salvatge i de mutants *glcC*

Es va procedir a estudiar la unió de GlcC al promotor de *glcD* mitjançant retardacions de la mobilitat electroforètica en gel del promotor amb extractes crus de la soca MC4100 i del mutant JA154 (*glcC*) creixent en diferents condicions. Es van fer els assajos de retardació seguint els protocols descrits anteriorment (Mètodes 4.2.1.1. i 4.4.2.4.1.). En alguns casos es va afegir glicolat a les mescles de reacció per tal d'estudiar el seu paper com a molècula efectora de l'acció de GlcC.

El fragment de DNA escollit per dur a terme aquests experiments va ser el fragment P. Aquest fragment abarca les posicions -247 a -134 i conté la seqüència UAS (-184 a -177)

identificada a partir de l'anàlisi de les delecions de promotor (Resultats 3.4.). La construcció del fragment P es va fer per PCR, utilitzant el primers GLCBam (5'-CGCGGATCCGTTTCGATACTCTCTGCAACC-3') i GLCP (5'-GGGGAATTCGTGCGTGTTTTGCGAG-3'). Les respectives dianes BamHI i EcoRI estan subratllades. Un cop purificat el fragment es va marcar radioactivament amb T4 polinucleòtid kinasa (Mètodes 3.3.2.2.).

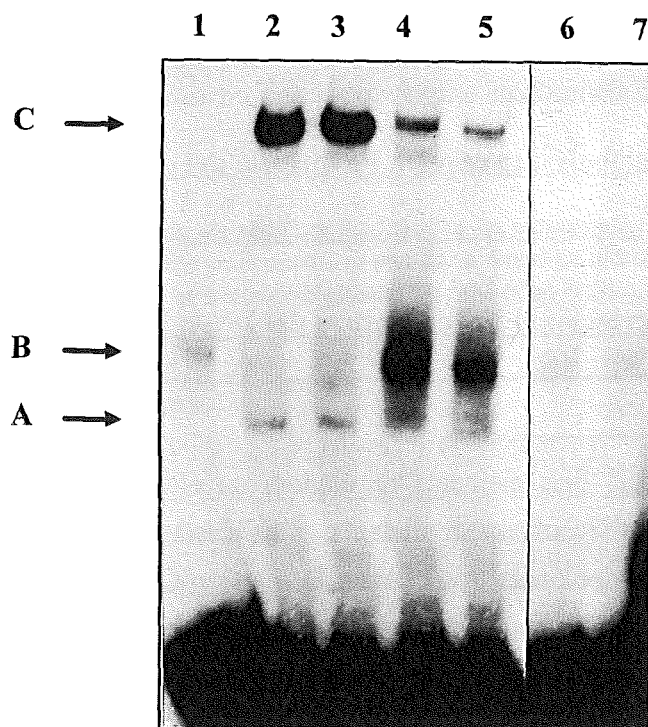


Figura 16. Assaig de retardació de la mobilitat electroforètica en gel del fragment P amb extractes crus de la soca MC4100 i del mutant *glcC* JA154. L'electroforesi es va realitzar en un gel de poliacrilamida al 5 %. Tots els carrils contenien 25 nmols de fragment P marcat amb ^{32}P . En el carril 1 no existia proteïna afegida. Els carrils 2 i 3 contenien 4 μg d'extracte cru de la soca MC4100 crescuda en hidrolitzat de caseïna 0.5 % més glicolat 30 mM. Els carrils 4 i 5 contenien 4 μg d'extracte cru de la soca MC4100 crescuda en hidrolitzat de caseïna 0.5 %. El carril 6 contenia 4 μg d'extracte cru de la soca JA154 crescuda en hidrolitzat de caseïna 0.5 %. El carril 6 contenia 4 μg d'extracte cru de la soca JA154 crescuda en hidrolitzat de caseïna 0.5 % més glicolat. Els carrils 3 i 5 contenien també glicolat 15 mM, el qual va ser afegit a la mescla d'incubació. Totes les mescles de reacció contenien un excés 500 molar de poli(dI.dC). A: banda A; B: banda B; C: banda C.

Els extractes crus de la soca MC4100 obtinguts a partir de creixements en hidrolitzat de caseïna més glicolat mostraven un patró diferent als provinents del creixement en caseïna (Figura 16). En els extractes induïts (carrils 1 i 3) s'observava una clara banda majoritària molt retardada (banda C) i una altra banda amb major mobilitat electroforètica (banda A), molt més

feble que la banda C. L'addició de glicolat a les mescles de reacció no feia variar aquest patró. En els extractes crus de la soca MC4100 crescuda en hidrolitzat de caseïna (carrils 4 i 5) es podia observar com al banda C estava molt disminuïda, la banda A lleument augmentada respecte els cultius induïts i també l'aparició d'una nova banda (banda B) de mobilitat intermitja entre la banda A i la banda C.

En les retardacions amb extractes de la soca JA154 (*glcC*) crescuda en hidrolitzat de caseïna es podia observar com les bandes C i A pràcticament havien desaparegut i només era lleugerament visible la banda B. L'addició de glicolat a les mescles no feia variar el patró de bandes en cap cas.

4.2. Factors de transcripció globals que regulen l'expressió del sistema *glc*

4.2.1. Activació transcripcional per curvatura del DNA. Paper de la proteïna IHF

L'estudi de la distribució dels elements reguladors en el promotor de *glcD* permet observar que tots ells es troben relativament distants. S'ha descrit que promotors d'aquest tipus requereixen factors transcripcionals que canvïen la topologia del DNA per permetre el contacte entre les diferents proteïnes reguladores. Un d'aquests factors és la proteïna IHF, i per aquest motiu es va decidir analitzar si aquest factor regulava l'expressió del sistema gènic *glc*.

4.2.1.1. Expressió en mutants IHF

Per conèixer si IHF regula l'expressió del sistema *glc* es van construir mutants de la soca MC4100 en cadascun dels dos gens que codifiquen les dues subunitats d'IHF. Les mutacions van ser transferides des de les soques HN1491 (Granston i Nash, 1993) (mutació en el gen *himA*, que codifica la subunitat α) o K2704 (Mozola i Friedman, 1985) (mutació a *himD3*, que codifica la subunitat β) a la soca MC4100 mitjançant transducció generalitzada amb el bacteriòfag P1, i selecció dels transductants per resistència a cloramfenicol. Els mutants van anomenar-se respectivament JA162 i JA163.

Sobre JA162 i JA163 es van transduir les fusions de promotor $\Phi(glcD-lacZ)$ i $\Phi(glcC-lacZ)$. Les soques resultants es van fer créixer en hidrolitzat de caseïna i també en presència de glicolat i es va assajar l'activitat β -galactosidasa en cada cultiu. En la Taula 4 es motren les activitats β -galactosidasa d'aquests cultius.

El resultat de l'assaig mostra que la fusió $\Phi(glcD-lacZ)$ no s'expressa en el fons genètic dels mutants JA162 i JA163, la qual cosa demostra que l'expressió dels gens estructurals de l'operó *glc* és totalment dependent d'IHF. Per contra, l'expressió de la fusió $\Phi(glcC-lacZ)$ no es veu modificada per les mutacions *himA* o *himD3*, la qual cosa indica que IHF no controla l'expressió del gen regulador del sistema *glc*. Cal remarcar, però, que en aquests mutants els nivells d'expressió de $\Phi(glcC-lacZ)$ en cultius d'hidrolitzat de caseïna més glicolat estan augmentats i són del mateix ordre que els obtinguts en absència de glicolat.

Taula 4. Efecte de mutacions en els gens que codifiquen les subunitats d'IHF en l'expressió de les fusions $\Phi(glcD-lacZ)$ i $\Phi(glcC-lacZ)$. Els creixements es van dur a terme en hidrolitzat de caseïna al 0.5 % (Caa) i en hidrolitzat de caseïna 0.5 % en presència de glicolat 30 mM (Caa + Glc). Les activitats β -galactosidasa estan expressades en unitats Miller.

Soca	Fusió	Activitat β -galactosidasa	
		Caa	Caa + Glc
MC4100	$\Phi(glcD-lacZ)$	150	1800
JA162	$\Phi(glcD-lacZ)$	<10	<10
JA163	$\Phi(glcD-lacZ)$	<10	<10
MC4100	$\Phi(glcC-lacZ)$	90	50
JA162	$\Phi(glcC-lacZ)$	87	85
JA163	$\Phi(glcC-lacZ)$	89	87

4.2.1.2. Estudis de retardació de la mobilitat electroforètica amb extractes crus de la soca salvatge i de mutants IHF

Aquests estudis de la mobilitat electroforètica es van dur a terme amb un fragment de DNA de 360 pb (fragment G) que abarcava des de les posicions -247 a + 94 del promotor de *glcD*. Aquest es va obtenir per PCR amb els primers GLCBam (5'-CGCGGATCCGTTTCGATACTCTCTGCAACC-3') i GLCEco (5'-GCGGGAATTCTAAAGCGCCATCAAGACG-3') usant com a motillo el plasmidi pTP26.

Les dianes de restricció *Bam*HI i *Eco*RI afegides a l'extrem 5' de cada primer estan subratllades.

Aquest fragment, marcat radioactivament amb la T4 PNK (Mètodes 3.3.2.2.), es va combinar amb extractes crus de la soca salvatge MC4100 i del mutant IHF JA162 (*himA*). Els extractes crus es van obtenir a partir de creixements en hidrolitzat de caseïna, així com a partir de cultius induïts per l'addició de glicolat 30 mM, tal i com es detalla anteriorment (Mètodes 4.2.1.1.).

A la Figura 17 es pot observar que amb els extractes cel.lulars de la soca salvatge MC4100 apareixen dues bandes de retardació (bandes X i Y), tant en els extractes crus crescuts en hidrolitzat de caseïna com en hidrolitzat de caseïna més glicolat. Els extractes crus de la soca JA162 crescuda en hidrolitzat de caseïna no provocaven l'aparició de cap banda de retardació en la mobilitat electroforètica del fragment G, mentre que amb els extractes del mutant crescut en presència de glicolat apareixia únicament la banda Y (Figura 17). L'addició de glicolat a les mescles de reacció no feia variar el patró de bandes.

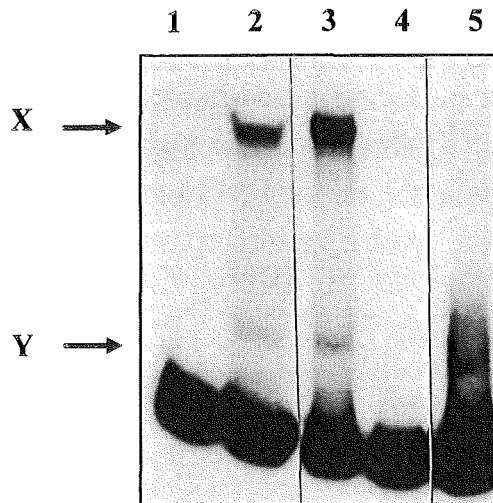


Figura 17. Assaig de retardació de la mobilitat electroforètica en gel del fragment G amb extractes crus de la soca MC4100 i del mutant *himA* JA162. L'electroforesi es va realitzar en un gel de poliacrilamida al 5 %. Tots els carrils contenien 25 nmols de fragment G marcat amb ^{32}P . En el carril 1 no existia proteïna afegida. El carril 2 contenia 4 μg d'extracte cru de la soca MC4100 crescuda en hidrolitzat de caseïna 0.5 %. El carril 3 contenia 4 μg d'extracte cru de la soca MC4100 crescuda en hidrolitzat de caseïna 0.5 % més glicolat 30 mM. El carril 4 contenia 4 μg d'extracte cru de la soca JA162 crescuda en hidrolitzat de caseïna 0.5%. El carril 5 contenia 4 μg d'extracte cru de la soca JA162 crescuda en hidrolitzat de caseïna 0.5 % més glicolat 30 mM. Totes les mescles contenien un excés 500 molar de poli(dI.dC). X: banda X; Y: banda Y.

4.2.1.3. Localització de la possible caixa d'unió d'IHF al promotor de *glcD*

Per tal de localitzar la possible zona d'unió d'IHF al promotor de *glcD* es va usar un programa informàtic anomenat MacTargsearch. Aquest programa genèric serveix per buscar i predir l'activitat de llocs d'unió de proteïnes que s'uneixen al DNA. En aquest cas concret, el programa utilitzat s'ha dissenyat per trobar els possibles llocs d'unió d'IHF. La informació per crear la seqüència consens per la unió d'IHF es va obtenir de l'aliniament de 27 llocs d'unió per IHF descrits en els promotors d'*Escherichia coli* (Goodrich *et al*, 1990). Dins el promotor de *glcD*, el programa va trobar dues regions amb un alta probabilitat de ser llocs d'unió per IHF: entre les posicions -128 a -93 en la cadena codificant de *glcD*, i entre les posicions -142 a -95 en la cadena complementària (Figura 18).

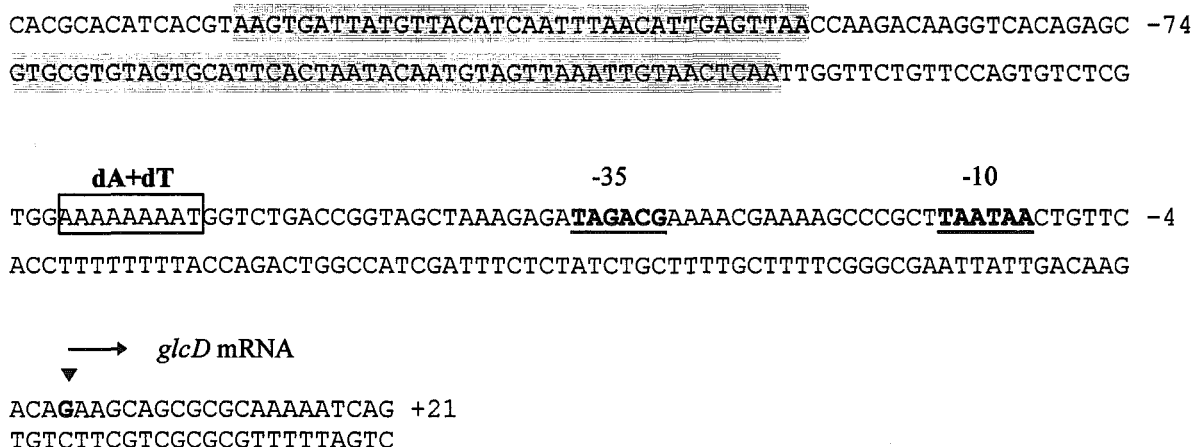


Figura 18. Possibles seqüències d'unió d'IHF al promotor de *glcD*. Es mostra part del promotor de *glcD* i en sombrejat s'indiquen les seqüències amb alta probabilitat de ser llocs d'unió per IHF segons el programa MacTargsearch (Goodrich *et al*, 1990). S'indiquen les caixes -10 i -35 (subratllades i en negreta), el tsp (triangle negre) i la zona dA+dT (enquadrada en un rectangle).

4.2.2. Repressió catabòlica

Com que el sistema *glc* està implicat en el metabolisme de compostos que poden ser usats com a font de carboni (glicolat i glioxilat), es va pensar que podia estar sota repressió catabòlica. Per analitzar si *glc* està regulat per repressió catabòlica, es va fer créixer la soca MC4100Φ(*glcD-lacZ*) en els següents medis: 1) medi mínim mineral SM-glucosa 1 mM, 2) SM-glucosa 1 mM + glicolat 30 mM, 3) SM-glucosa 1 mM + glicolat 30 mM + cAMP 5 mM,

i 4) SM-succinat 20 mM + glicolat 30 mM. Aquests cultius es van inocular amb preinòculs de la soca creixent en glicolat (induits) o sense glicolat (no induïts). Es prenen alíquotes al cap de 2 hores, 3 hores 30 min, 5 hores i 6 hores 30 min, i s'assajaven les activitats β -galactosidasa.

Es va poder observar que els cultius d'SM-glucosa expressaven un baix nivell d'activitat, que es mantenia constant al llarg de tots els assajos. Els cultius d'SM-glucosa-glicolat tenien uns nivells d'activitat baixos en les dues primeres mesures de l'activitat, però aquesta augmentava fins a nivells d'inducció normals en l'últim temps, moment en el qual s'havia consumit la glucosa present. Els cultius d'SM-glucosa-glicolat-cAMP mostraven els mateixos nivells d'activitat que els d'SM-glucosa-glicolat i no s'aconseguia una reversió de la repressió exercida per la glucosa en els primers temps del creixement. Sembla doncs que la glucosa reprimeix l'expressió de $\Phi(glcD-lacZ)$ però que aquesta repressió no és revertida per la presència de cAMP en el medi. Per últim, el cultiu control d'SM-succinat-glicolat va mostrar inducció en els primers temps, essent màxima a les 3 hores 30 min, i aquesta es va mantenir en les altres determinacions.

Cal esmentar que no es localitza cap seqüència consens per la unió de CRP (TGTGA(N₆)TCACA) en el promotor de *glcD*, per la qual cosa es pot assumir que la repressió exercida per la glucosa no està mediada per l'acció directa del complex CRP-cAMP sobre el promotor de *glcD*. Altres mecanismes, com són l'exclusió de l'inductor o altres sistemes de control per catabolit independents de cAMP, poden ser els responsables d'aquesta repressió.

4.2.3 Regulació aeròbica/anaeròbica del sistema *glc*

En l'anàlisi de la seqüència nucleotídica del promotor del sistema *glc* s'observen seqüències similars al consens per la unió d'ArcA definit per Lynch i Lin (1996a). Es va creure interessant comprovar si els gens del sistema *glc* estan sota repressió anaeròbica i en cas afirmatiu, estudiar quin és el sistema global de regulació de la transcripció que controla aquest procés. Per aquest motiu es van dur a terme experiments tant *in vivo* com *in vitro* per comprovar la implicació de la disponibilitat d'oxigen en la regulació de l'expressió de l'operó *glc*.

4.2.3.1 Expressió dels promotors de *glcD* i *glcC* en condicions d'anaerobiosi

En primer lloc es van realitzar experiments *in vivo*. Per tal de comprovar si l'expressió dels gens estructurals de l'operó *glc* està regulada per la disponibilitat d'oxigen es va fer créixer la soca MC4100 Φ (*glcD-lacZ*) en condicions anaeròbiques, tant en presència com en absència d'acceptors finals d'electrons, tals com el nitrat, fumarat o DMSO. Quan en el medi no es va afegir cap acceptor d'electrons alternatiu a l'oxigen, *Escherichia coli* duia a terme una fermentació mixte, mentre que en presència dels acceptors era capaç de respirar anaeròbicament. El medi de creixement era hidrolitzat de caseïna a l'1 % en presència o absència de glicolat 60 mM.

Quan els cultius anaeròbics no contenien cap acceptor final d'electrons el creixement cel·lular assolía nivells de densitat òptica a 600 nm (O. D.₆₀₀) de 0.1, de 0.2 en el cas del fumarat (40 mM) i de DMSO (70 mM) i quan s'afegia nitrat (20 mM) s'assolia una O. D.₆₀₀ de 0.5. Aquests rendiments cel·lulars estan en concordància amb el potencial redox de cadascun dels acceptors d'electrons afegits.

En la Taula 5 es mostren les activitats β -galactosidasa de la soca MC4100 Φ (*glcD-lacZ*) en cultius d'hidrolitzat de caseïna i en el mateix medi en presència de glicolat, quan es varia la disponibilitat d'oxigen i/o de nitrat com a acceptor d'electrons alternatiu. Alguns cultius contenien piruvat que és un estimulador de la transducció de senyals del sistema ArcB/ArcA (Iuchi, 1993). Cal esmentar que cap dels acceptors finals d'electrons tenien efecte sobre l'expressió aeròbica de Φ (*glcD-lacZ*).

Es pot observar que l'expressió dels gens estructurals és dependent de la presència d'oxigen. L'expressió de la fusió Φ (*glcD-lacZ*) en condicions anaeròbiques sense cap acceptor d'electrons (condicions fermentatives) en cultius induïts en glicolat és el 3.6 % dels valors en condicions aeròbiques. La regulació més global (en aquest cas la disponibilitat d'oxigen) prevaleix sobre l'efecte més específic (la inducció per presència de glicolat). En presència de nitrat, la transcripció de *glcD* és la mateixa que en condicions aeròbiques. La repressió anaeròbica és revertida pràcticament per complet per l'addició de nitrat 20 mM, que actua com a acceptor final d'electrons alternatiu. Aquesta reversió s'observa tant en els cultius d'hidrolitzat de caseïna com en cultius en presència de glicolat, que assoleixen els nivells dels d'expressió del creixements aeròbics.

La repressió anaeròbica dels cultius realitzats en presència de glicolat es veu incrementada per l'addició de piruvat. L'efecte repressor del piruvat és conseqüència de l'estimulació de la fosforilació d'ArcB que aquest compost promou. Aquest efecte de repressió s'observa també quan s'addiciona al medi nitrat com a acceptor final d'electrons. En aquestes

condicions, la reversió de la repressió provocada per l'addició de nitrat és només parcial. Aquest fet indica que nitrat i piruvat estan exercint efectes antagònics en la regulació anaeròbica de l'operó *glc*.

Taula 5. Efecte de la disponibilitat d'oxigen, de nitrat i de piruvat sobre l'expressió de $\Phi(glcD-lacZ)$ en la soca MC4100. La soca MC4100 $\Phi(glcD-lacZ)$ es va fer créixer en condicions d'anaerobiosi (-O₂) o d'aerobiosi (+O₂). Caa: hidrolitzat de caseïna 1%; Caa + Glc: hidrolitzat de caseïna més glicolat 30 mM en condicions aeròbiques i 60 mM en condicions anaeròbiques. L'activitat β -galactosidasa s'expressa en unitats Miller.

Condicions de creixement	Activitat β -galactosidasa	
	-O ₂	+O ₂
Caa	<20	150
Caa + Nitrat 20 mM	140	155
Caa + Piruvat 20 mM	<20	135
Caa + Piruvat 20 mM + Nitrat 20 mM	50	140
Caa + Glc	65	1800
Caa + Glc + Nitrat 20 mM	1600	1850
Caa + Glc + Piruvat 20 mM	40	1650
Caa + Glc + Piruvat 20 mM + Nitrat 20 mM	500	1700

Aquests resultats demostren que l'expressió dels gens estructurals de l'operó *glc* està reprimida en condicions d'absència d'oxigen. El que es va fer a continuació és esbrinar si aquest efecte era indirecte, causat per un mecanisme de repressió de l'expressió del gen regulador *glcC*. Com que el producte gènic de *glcC* és l'activador de la transcripció dels gens estructurals del sistema *glc* (Resultats 4.1.), una disminució en la producció anaeròbica de GlcC es traduiria en una manca d'expressió de l'operó *glc*.

Per tal de determinar aquest efecte es van fer creixements de la soca MC4100 $\Phi(glcC-lacZ)$ en hidrolitzat de caseïna i hidrolitzat de caseïna més glicolat en condicions aeròbiques i anaeròbiques. No es va veure cap canvi en el patró d'expressió de $\Phi(glcC-lacZ)$. Aquest resultat indica que la regulació anaeròbica produïda sobre l'expressió dels gens estructurals no és causada per una disminució de la taxa de transcripció del gen regulador *glcC*.

4.2.3.2 Expressió del sistema *glc* en mutants *arcA*

Els resultats de l'apartat anterior indiquen que l'expressió de la fusió $\Phi(glcD-lacZ)$ està controlada per la disponibilitat d'oxigen. El fet que la repressió anaeròbica estigui estimulada per piruvat i reprimida per nitrat sembla indicar que aquest control és exercit pel sistema de dos components ArcB/ArcA. Per tal de comprovar aquesta hipòtesi es va construir un mutant *arcA* derivat de la soca MC4100 mitjançant transducció generalitzada a partir d'un lisat P1 de la soca ECL618 (Iuchi *et al.*, 1989). Aquesta soca té un Tn10 proper al gen *arcA* mutat, de manera que el percentatge de cotransducció de la mutació *arcA* i del marcador Tn10 és del 50 %. Per això els mutants es van seleccionar per resistència a tetraciclina i es va comprovar posteriorment la seva sensibilitat al blau de toluidina. La soca mutant va anomenar-se JA164.

Sobre la soca JA164 es va transduir la fusió $\Phi(glcD-lacZ)$, i es va analitzar la seva expressió en cultius induïts en hidrolitzat de caseïna més glicolat creixent en condicions aeròbiques i anaeròbiques, en presència o absència de nitrat. En la Taula 6 es mostren les activitats β -galactosidasa obtingudes en aquests assajos. Es pot observar com la introducció de la mutació en el gen que codifica ArcA aboleix completament la repressió anaeròbica, indicant el paper de la proteïna en la regulació aeròbica/anaeròbica del sistema *glc*.

La introducció de la mutació *arcA* modifica lleugerament l'expressió aeròbica del sistema: s'observa un augment d'aproximadament el 14 % en l'expressió de JA164 $\Phi(glcD-lacZ)$ respecte a MC4100 $\Phi(glcD-lacZ)$ en condicions aeròbiques.

Taula 6. Efecte de la mutació *arcA2* en l'expressió de la fusió $\Phi(glcD-lacZ)$ en diferents condicions de creixement. Les soques es van fer créixer en hidrolitzat de caseïna (Caa) 1 % en presència de glicolat (Glc) 60 mM en els cultius anaeròbics i de 30 mM en els aeròbics. El nitrat s'addicionava a 20 mM. Les activitats s'expressen en unitats Miller.

Soca	Fusió	Activitat β -galactosidasa			
		-O ₂		+O ₂	
		Caa + Glc	Caa + Nitrat + Glc	Caa + Glc	Caa + Nitrat + Glc
MC4100	$\Phi(glcD-lacZ)$	65	1600	1800	1850
JA164	$\Phi(glcD-lacZ)$	1700	1750	2050	2100

4.2.3.3. Estudis d'unió de la proteïna ArcA al promotor de *glcD*

Aquests estudis *in vitro* es van dur a terme al laboratori del Professor E. C. C. Lin al Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, USA.

A l'apartat anterior s'ha demostrat *in vivo* la implicació del sistema Arc en la repressió anaeròbica dels gens estructurals de l'operó *glc*. Aquest apartat estudia les interaccions moleculars entre el promotor de *glcD* i ArcA.

4.2.3.3.1. Estudis de retardació de la mobilitat electroforètica del promotor de *glcD* amb proteïna Hys₆-ArcA purificada

La purificació d'ArcA es va dur a terme segons havia estat descrita anteriorment (Lynch i Lin, 1996a). Per facilitar l'obtenció de la proteïna pura es va usar el sistema en el qual s'addiciona una cua de sis histidines a l'extrem N-terminal del polipèptid (Mètodes 2.5.). D'aquesta manera s'aconsegueix augmentar el rendiment i disminuir els passos de purificació. S'ha demostrat que l'addició d'aquest hexapèptid d'histidines no modifica la capacitat d'unió al DNA d'alguns reguladors que formen part de sistemes de dos components (Hoch i Silhavy, 1995). D'aquí en endavant la proteïna Hys₆-ArcA s'anomenarà ArcA.

La unió d'ArcA als seus promotors diana es veu afavorida per la fosforilació. Així, la capacitat d'unió al DNA és de 10 vegades superior quan la proteïna està fosforilada que quan no ho està (Drapal i Sawers, 1995). La fosforilació d'ArcA (50 µg/ml) es va dur a terme en presència de carbamoïl fosfat a 50µM durant 1h a 30°C. Per incubació directe amb donadors de grups fosfat de baix pes molecular, p.e. acetil fosfat o carbamoïl fosfat, s'ha demostrat que moltes proteïnes d'unió al DNA procariotes assoleixen una forma activa (Lukat *et al*, 1992).

Per a dur a terme els estudis de la retardació de la mobilitat electroforètica es va usar un fragment de DNA pertanyent al promotor de *glcD*, que inclou les seqüències compreses entre les posicions -247 i +94. Aquest es va obtenir per PCR amb els oligonucleòtids GLCBam (5'-CGCGGATCCGTTTCGATACTCTCTGCAACC-3') i GLCEco (5'-GCGGGAATTCTAAAGCGCCATCAAGACG - 3') usant com a motillo el plasmidi pTP26. Les dianes de restricció (*Bam*HI i *Eco*RI) afegides a l'extrem 5' de cada primer estan subratllades. El resultat va ser un fragment de 360 pb.

Els assajos de retardació amb ArcA purificada es van dur a terme en les condicions descrites anteriorment (Mètodes 4.2.). Les mescleres de reacció contenien diferents quantitats de proteïna no fosforilada (ArcA) o fosforilada (ArcA-P).

Els assajos de retardació es van dur a terme en un gel de poliacrilamida al 5 % i el resultat es mostra a la Figura 19. En ells s'observa que la retardació de la mobilitat electroforètica del DNA es veu incrementada per la fosforilació d'ArcA. A mesura que s'augmenta la concentració de proteïna ArcA-P en les mescles de reacció s'observen complexos més retardats. Aquest fet implica que en el promotor de *glcD* existeixen múltiples zones d'unió a ArcA-P o bé que la proteïna unida al promotor s'oligomeritza.

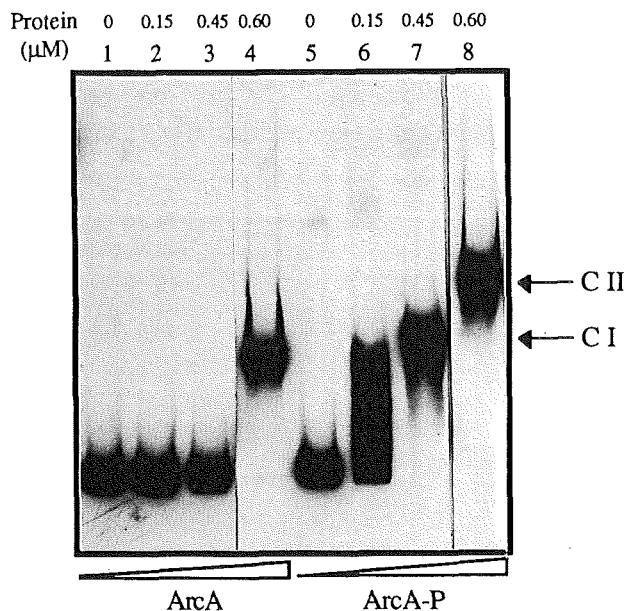


Figura 19. Assaig de retardació en gel de la regió promotora de *glcD* amb ArcA i ArcA-P. L'assaig es va dur a terme en un gel de poliacrilamida al 5%. El fragment de 360 pb, marcat radioactivament amb 32 P es va addicionar a 2.5 nM en les mescles de reacció, que contien diferents concentracions d'ArcA o ArcA-P. Les concentracions de proteïna es mostren a sobre de cada carril. S'indica la migració dels dos complexos DNA-ArcA-P (C I: complexe I; C II: complexe II). En cada reacció es va incloure un excés 500 molar de DNA d'esperma de salmó. Els triangles de la figura mostren concentracions creixents d'ArcA o ArcA-P.

4.2.3.3.2 Localització de les caixes d'unió a ArcA-P en el promotor de *glcD*

Per tal de localitzar la posició de les caixes d'unió d'ArcA-P al promotor de *glcD* es van dur a terme experiments de protecció a la DNasa I. Aquests es van realitzar tant amb ArcA com amb ArcA-P i utilitzant com a DNA substrat el mateix fragment de 360 pb que havia esta usat en els experiments de retardació de la mobilitat electroforètica. Després del marcatge radioactiu, el fragment va ser separat en dues alíquotes i digerit en cada cas amb *Bam*HI o *Eco*RI. Per tal d'eliminar les curtes cues i els nucleòtids marcats lliures es van fer extraccions fenol/cloroform de les digestions. La relació entre les concentracions de DNA i proteïna eren equivalents a les utilitzades en els experiments de retardació en gel.

No es van observar regions discretes de protecció quan en la mescla hi havia ArcA, mentre que en presència d'ArcA-P eren visibles àmplies zones de protecció (Figura 20). Tot i que aquestes zones de protecció eren força exteses, es podien observar segments altament protegits a la digestió de la DNasa I. Les zones altament protegides per ArcA-P es localitzaven des de la posició -45 a +31 (ArcA site I) i des de -106 a -161 (ArcA site II). Les posicions s'indiquen respecte al tsp del gen *glcD* (Figura 20). Al llarg de les dues regions d'unió a ArcA-P es podien observar posicions amb una augmentada sensibilitat a l'acció de la DNasa I. Les dues zones protegides per ArcA-P es podien observar en les dues cadenes de DNA, si bé a la Figura 18 només es mostra una cadena en cada cas.

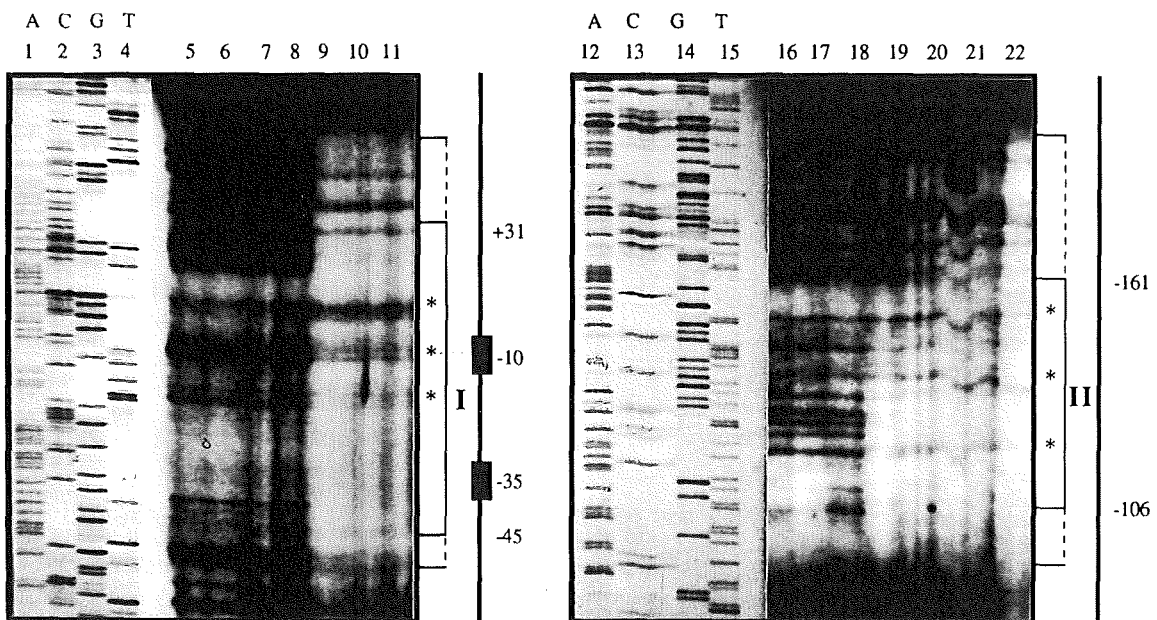


Figura 20. Assaig de protecció a la DNasa I del promotor de *glcD* per ArcA i ArcA-P. Els substrats marcats radiactivament únicament per un extrem van ser generats per digestió del fragment de 360 pb amb *EcoRI* (carrils 5-11) i *BamHI* (carril 16-22). El DNA estava present a 2nM en reaccions que a més contenien 0.6 μM ArcA (carril 6), 1.5 μM ArcA (carrils 7 i 17), 0.2 μM ArcA-P (carrils 9 i 19), 0.6 μM ArcA-P (carrils 10 i 20), 1.0 μM ArcA-P (carril 21) i 1.5 μM ArcA-P (carrils 11 i 22). El carrils 5,8,16 i 18 no contenien proteïna addicionada. Les reaccions de seqüència es van generar per extensió dels primers *GLCBam* (carrils 1-4) o *GLCEco* (carrils 12-15) marcats radiactivament amb ³²P usant el plasmidi pTP25 com a motllo. I: ArcA-P site I; II: ArcA-P site II. Amb una línia contínua es determina la zona més fortament protegida, i amb línia discontinua la zona de protecció més laxa. Les bases amb una sensibilitat augmentada a l'acció de la DNasa I dins les regions protegides per ArcA-P estan assenyalades per asteriscs (*).

La localització de les caixes d'unió d'ArcA-P al promotor de *glcD*, juntament amb els altres elements reguladors, es mostren a la Figura 21.

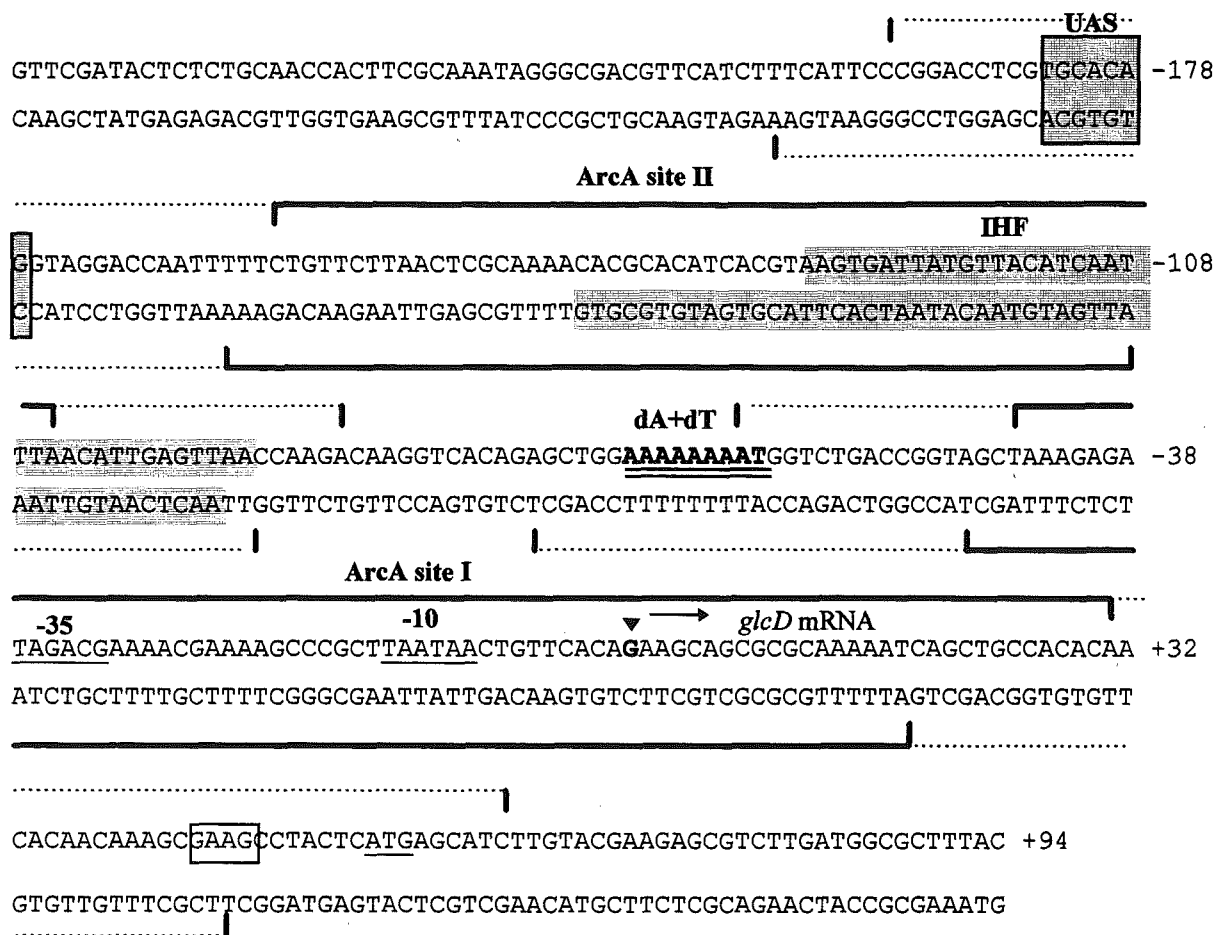


Figura 21. Localització dels elements reguladors del promotor *glcD*. La seqüència abarca des de la posició -247 a la +94 del promotor de *glcD* (correspon als fragments usats en els assajos de retardació de la mobilitat electroforètica i als de protecció a la DNasa I). En la figura es mostren: les caixes ArcA site I i ArcA site II, amb les zones més protegides assenyalades en línia contínua i les de protecció més laxa en línia discontinua; les posicions -10 i -35 (subratllades), el tsp (indicat amb un triangle negre), la seqüència Shine-Dalgarno (amb un rectangle); l'ATG iniciador (subratllat); la seqüència dA+dT (en negreta i doblement subratllada); les possibles zones d'unió d'IHF (sombrejades amb gris) i la zona UAS (en un rectangle i sombrejada amb gris).

5. VALIDACIÓ DE LA SEQÜÈNCIA CONSENS D'ArcA-P USANT COM A MODEL EL PROMOTOR D'*aldA*

Els resultats inclosos en aquest apartat i referents al sistema *aldA* van estar realitzats al laboratori del Professor E. C. C. Lin al Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, USA. Els experiments inclosos en aquest apartat es van fer

de manera paral·lela en el mateix període durant el qual es van realitzar els del sistema *glc* (Resultats 4.2.3.3.). Els estudis allà realitzats tenien com a finalitat demostrar la funcionalitat tant *in vitro* com *in vivo* de la seqüència consens proposada per Lynch i Lin (1996a) per la unió d'ArcA als seus promotors diana.

El gen *aldA* està regulat a nivell transcripcional per la disponibilitat d'oxigen. Existeix una clara repressió del sistema en condicions anaeròbiques. Aquesta repressió està mediada pel sistema de dos components ArcB/ArcA (Limon *et al*, 1997).

L'elecció del sistema *aldA* com a font d'estudi per aquests experiments va basar-se en dos motius:

i) el gen *aldA* ha estat font d'estudi per ambdós laboratoris durant els darrers anys i per tant era interessant aprofundir en l'anàlisi dels mecanismes moleculars que regulen la seva expressió aeròbica/aneròbica, i

ii) la possible seqüència consens d'unio d'ArcA al promotor d'*aldA* té una similitud del 100 % amb la proposada per Lynch i Lin (1996a).

Cal esmentar que la relació entre el gen *aldA* i el sistema *glc*, objecte d'estudi en aquest treball, s'estableix pel fet que el producte gènic d'*aldA* és una aldehyd deshidrogenasa que genera glicolat a partir de glicolaldehyd en els metabolismes d'etilenglicol, L-lixosa i D-arabiosa a *Escherichia coli*.

5.1. Estudi de la retardació de la mobilitat electroforètica del promotor salvatge amb proteïna Hys₆-ArcA purificada

Per a comprovar que la proteïna ArcA reprimeix la transcripció d'*aldA* per unió al seu promotor, es va procedir a la purificació de la mateixa i a la posterior anàlisi de la unió entre promotor i factor de transcripció.

El fragment de DNA que va usar-se per a dur a terme les retardacions en gel era de 363 pb. Va obtenir-se mitjançant PCR amb els primers ALDBam (5'-GGGGATCCGGCAGAAGTCATCCATGCATGG- 3') i ALDEco (5'-GGGGAATTCCATGCGTCTCCACGCCAGG- 3'), usant com a motllo el plasmidi pALD1.2 (Hidalgo *et al*, 1996). En la seqüència dels primers estan subratllades, respectivament, les dianes de restricció BamHI i EcoRI. El fragment incloïa una seqüència de la zona promotora d'*aldA* compresa entre les posicions -236 a +110.

Els assajos de retardació amb ArcA purificada es van dur a terme en les condicions descrites anteriorment (Mètodes 4.2.). Les mescles de reacció contenien diferents quantitats de proteïna no fosforilada (ArcA) o fosforilada (ArcA-P), preparada tal i com descriuen Lynch i Lin (1996a).

La unió d'ArcA al promotor d'*aldA* sembla ser relativament inespecífica, ja que és contrarestatada per l'addició de DNA competitiu inespecífic en les mescles de reacció (Figura 22, comparar carrils 1-3 amb 8-10). Per contra, el patró de retardació del fragment amb ArcA-P sembla específic, ja que tot i que existeix una disminució de la mobilitat per addició de DNA competitiu no marcat, aquesta és molt menor que en el cas de les mescles amb ArcA (Figura 22, comparar carrils 4-6 amb 11-14).

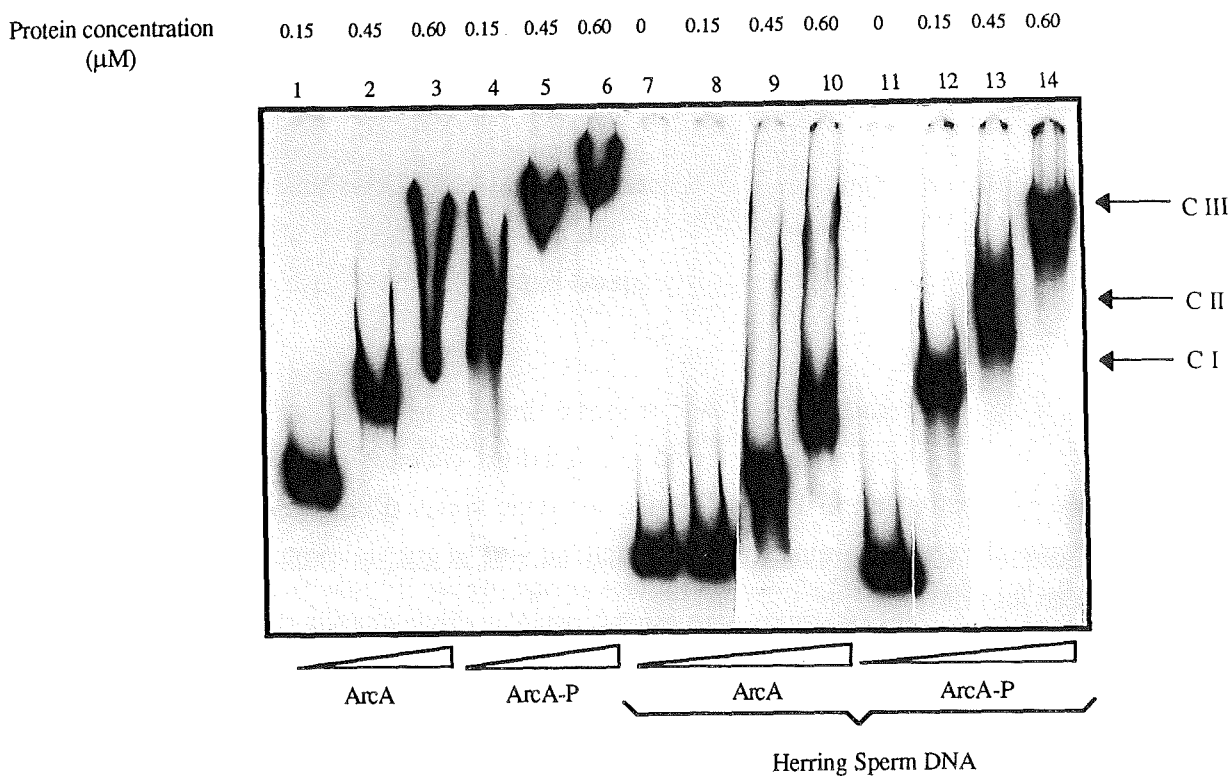


Figura 22. Assaig de retardació en gel de la regió promotora tipus salvatge d'*aldA* amb ArcA i ArcA-P. L'electroforesi es va realitzar en un gel de poliacrilamida al 5%. El fragment de 363 pb, marcat radiactivament amb ^{32}P , es va mesclar a una concentració 2.5 nM amb diferents quantitats de proteïna no fosforilada (ArcA) o fosforilada (ArcA-P). Les concentracions de proteïna en cada mescla es mostren a sobre de cada carril. S'indica la migració dels tres complexes (C I: complexe I; C II: complexe II; C III: complexe III). En els carrils 6-18 les mescles de reacció contenien en excés 500 molar de DNA d'esperma de salmó. Els triangles de la figura representen quantitats creixents d'ArcA o d'ArcA-P.

A mesura que s'incrementa la quantitat d'ArcA-P en les mescles de reacció, en presència de DNA competidor, apareixen complexos més retardats. Aquest fet implica la presència de múltiples llocs d'unió per ArcA en el promotor d'*aldA* i/o l'existència d'oligomerització de la proteïna fosforilada unida. Es poden observar tres complexos retardats: complexe I (C I), complexe II (C II) i complexe III (C III). Concentracions d'ArcA-P en les mescles de reacció superiors a les presentades a la Figura 22. no provocaven l'aparició de complexos més retardats.

5.2. Localització de les caixes d'unió a ArcA-P en el promotor d'*aldA*

Limon *et al* (1997) havien descrit la possible localització de la caixa d'unió d'ArcA al promotor d'*aldA* per comparació amb la seqüència consens proposada per Lynch i Lin (1996a) i per l'estudi de diferents delecions del promotor.

Per poder localitzar exactament la situació de la(es) caixa(es) d'unió d'ArcA a la regió 5' d'*aldA*, es va procedir a estudiar la protecció que la proteïna exerceix sobre el promotor de gen *aldA* enfront la digestió de la DNasa I. En aquest experiment es va usar el mateix fragment de 363 pb que en els assajos de retardació en gel.

La protecció a la DNasa I va dur-se a terme des dels dos extrems del fragment, de manera que es pogués observar la protecció en ambdues cadenes del DNA. Els resultats d'aquest assaig es mostren en la Figura 23. De les dues zones d'unió d'ArcA al promotor només es mostra el resultat obtingut en una cadena, tot i que la protecció va aparèixer paral·lelament en les dues.

El resultat de la protecció amb DNasa I indica que existeixen dues regions en el fragment assajat que estan protegides a la digestió per aquesta nucleasa (Figura 23). Aquestes dues zones s'extenen des de les posicions -18 a +29 (ArcA site I) i de -171 a -117 (ArcA site II). En la regió anomenada ArcA site I es troben dues seqüències que tenen un 10/10 de similitud amb el consens proposat per Lynch i Lin (1996a) (posicions -13 a -4 en la cadena codificant i -2 a -11 en la no codificant). En la regió ArcA site II existeix una seqüència que té 7/10 posicions conservades (posicions -138 a -129 en la cadena codificant).

Les zones de protecció no són discretes, sinó que abarquen moltes bases de la zona promotora. Aquest tipus de protecció s'ha descrit per diferents membres del moduló Arc (Lynch i Lin, 1996a). La protecció no s'incrementa a l'augmentar la concentració de proteïna.

Quan la concentració de la proteïna ArcA-P disminueix per sota de 0.15 μM , només s'observava protecció a ArcA site I.

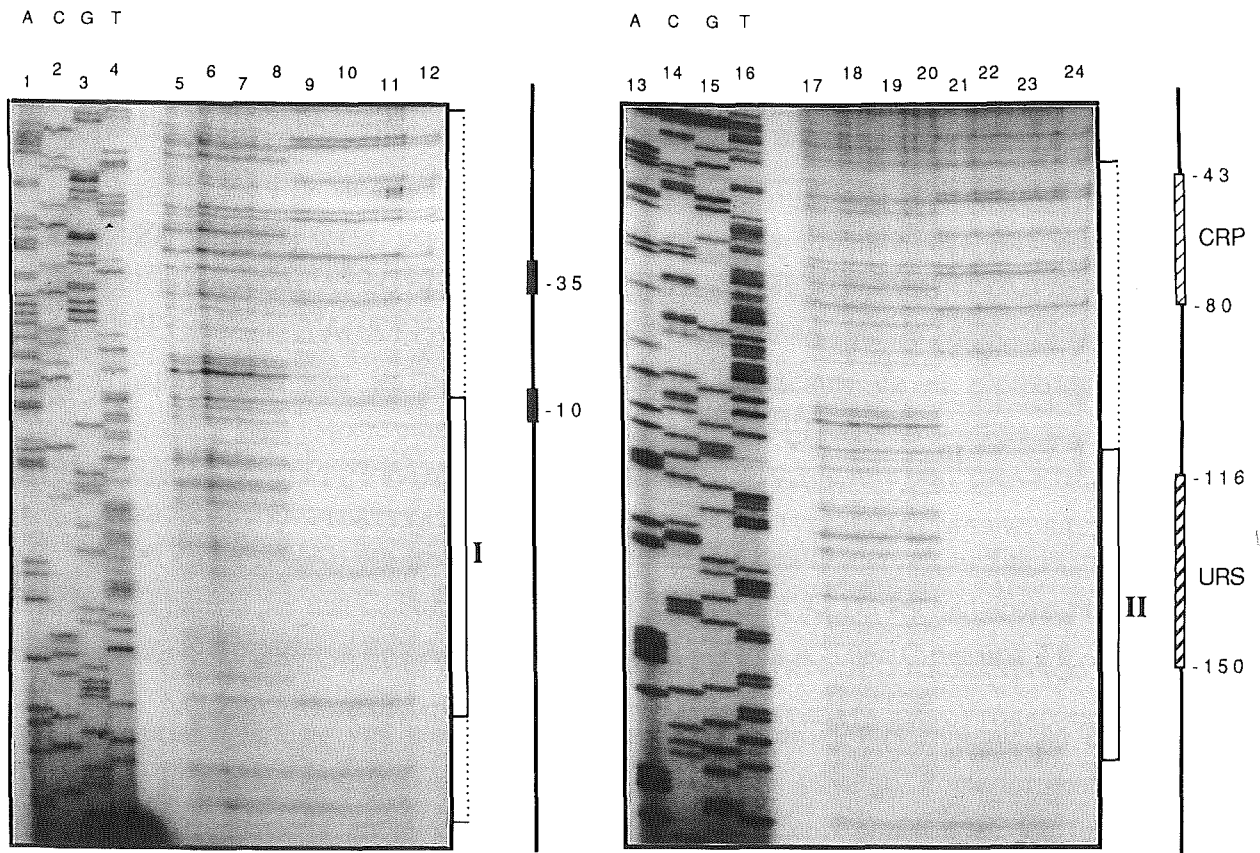


Figura 23. Assaig de protecció a la DNasa I del promotor d'*aldA* per ArcA i ArcA-P. Els fragments marcats radioactivament amb ^{32}P van ser digerits amb *Bam*HI (carrils 5-12) i *Eco*RI (carrils 12-24) i es van addicionar a 2 nM. Els diferents carrils contenen: 0.6 μM ArcA (carrils 6 i 18); 1.5 μM ArcA (carrils 7 i 19); 0.2 μM ArcA-P (carrils 9 i 21); 0.6 μM ArcA-P (carrils 10 i 22); 1.0 μM ArcA-P (carrils 11 i 23) i 1.5 μM ArcA-P (carrils 12 i 24). Els carrils 5, 8, 17 i 20 no contenen proteïna addicionada. Les reaccions de seqüència van ser generades per extensió dels primers marcats radioactivament ALDEco (carrils 1-4) i ALDBam (carril 13-16). I: ArcA-P site I; II: ArcA-P site II; CRP: proteïna reguladora per catabolít; URS: upstream regulatory sequence.

La regió promotora continguda en el fragment de DNA usat en els assajos d'unió d'ArcA-DNA es mostra a la Figura 24. A més es mostren els elements dels promotors que havien estat descrits anteriorment (Limon *et al.*, 1997) i les zones d'unió d'ArcA al promotor d'*aldA* identificades en aquest treball.

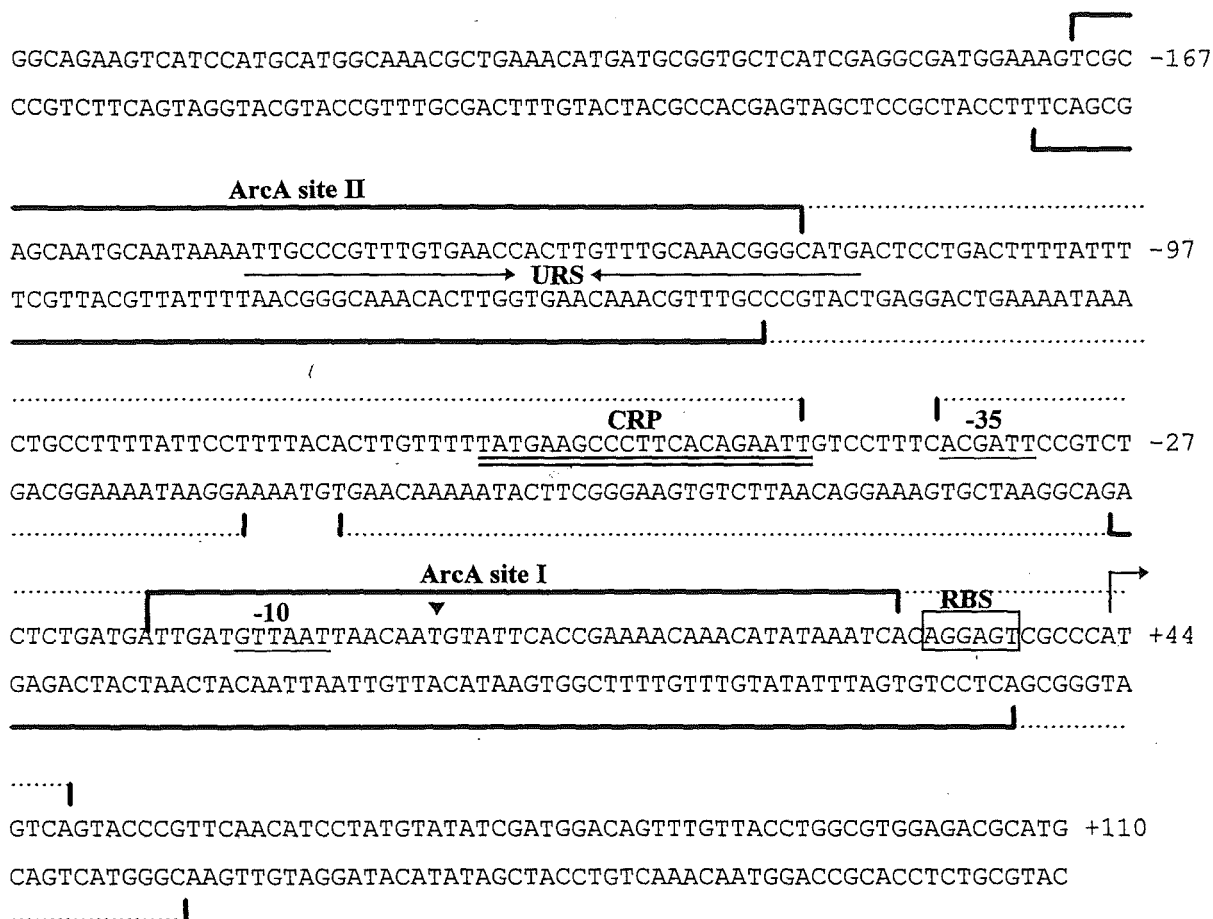


Figura 24. Organització dels elements reguladors en la regió promotora d'aldA. Fragment de 363 pb de la zona promotora d'aldA (que inclou la seqüència entre les posicions -236 a +110) utilitzat en els assajos de retardació en gel i de protecció a la DNasa I. Les dues regions d'unió d'ArcA-P (ArcA-P site I i II) que es mostren han estat determinades mitjançant la protecció a la DNasa I: amb una línia contínua es determina la zona més fortament protegida, i amb línia discontinua la zona de protecció més laxa. Els elements prèviament identificats per Limon *et al* (1997) són: URS (upstream regulatory sequence), indicada amb fletxes invertides; CRP (catabolite repression protein), indicada amb una doble línia; el tsp (transcriptional start point) està marcat amb un triangle negre; la seqüència d'unió a ribosoma (RBS) està insertada en un rectangle. Les seqüències -10 i -35 estan subtilades.

5.3. Obtenció de mutants en la caixa ArcA-P situada a la regió -10 del promotor d'aldA mitjançant mutagènesi dirigida

Per obtenir mutants en la caixes ArcA (posicions -13 a -4 en la cadena codificant i -2 a -11 en la no codificant) situades en la regió ArcA site I, es va usar un mètode de mutagènesi

dirigida anomenat Sculpture™ (Amersham Life Science Inc.). Es va decidir mutar les caixes situades en ArcA site I per dues raons:

i) les caixes que en ella es troben tenen un 10/10 de similitut amb la seqüència consens proposada per Lynch i Lin (1996a) i,

ii) a baixes concentracions de proteïna, ArcA-P només s'uneix a aquesta zona i no a ArcA site II.

La mutagènesi es va desenvolupar seguint el protocol descrit anteriorment (Mètodes 1.4.2.2.). Els primers usats per construir cadascun dels mutants van ser: ALD2 (5'-GTTTTCGGTGAATACATTCTTAATTAAGATCAATCATCAGAGAGACGG-3') i ALD3 (5'-GTTTTCGGTGAATACATTATTAATTAATATCAATCATCAGAGAGACGG-3'). Les posicions subratllades són les que introdueixen les mutacions en el promotor d'*aldA*. D'aquesta manera la seqüència consens tipus salvatge 5'-TGTTAATTAC-3' passa a ser en el cas del primer ALD2 5'-TCTTAATTAAG-3' i en els cas d'ALD3 5'-TATTAATTAAT-3'.

En el cas d'ALD2, les mutacions introduïdes no modifiquen el contingut relatiu de bases (una G canvia per una C, i una C per una G). En el cas d'ALD3 les mutacions provoquen un increment en el contingut de dA+dT (una G canvia per una A, i una C per una T).

En primer lloc es van clonar les seqüències dels promotors mutants en el vector pBluescript, juntament amb la seqüència del promotor tipus salvatge. Posteriorment, es van subclonar els tres inserts en el vector fusió pRS550. Un cop construïdes les fusions de promotor al gen *lacZ* en el vector pRS550, aquestes van ser transferides al genoma de la soca salvatge RYC1000 en forma de lisògens a través del fag λ RS45 (Mètodes 1.3.). Posteriorment, amb un lisat d'aquests lisògens es van transferir les fusions al fons genètic de la soca salvatge MC4100.

Les soques obtingudes de la introducció de les fusions dels promotors en el fons MC4100 es van anomenar: MC4100 Φ (*ald1-lacZ*) (promotor tipus salvatge); MC4100 Φ (*ald2-lacZ*) (promotor mutat obtingut amb el primer ALD2) i MC4100 Φ (*ald3-lacZ*) (promotor mutat obtingut amb el primer ALD2).

Sobre la soca MC4100 Φ (*ald1-lacZ*) es va transferir la mutació *arcA2* present en la soca ECL618 (Iuchi *et al*, 1989), de manera que es disposava d'una soca amb el promotor salvatge però sense proteïna ArcA activa. La soca resultant va anomenar-se MC4100 Φ (*ald1-lacZ*) *arcA2*.

5.4. Anàlisi dels mutants en la caixa ArcA-P (ArcA site I) del promotor d'*aldA*

5.4.1. Anàlisi mitjançant fusions de promotor al gen *lacZ*

L'efecte de la introducció de les mutacions en el promotor d'*aldA* va ser analitzat assajant l'activitat β -galactosidasa que cada construcció expressava. Per a dur a terme els creixements de la soca MC4100 amb les diferents fusions es va utilitzar l'hidrolitzat de caseïna, que és un medi inductor de l'operó *aldA* (Baldomà i Aguilar, 1987). Els creixements es van realitzar en condicions aeròbiques i també en condicions anaeròbiques. Les activitats β -galactosidasa van ser assajades de la manera descrita per Miller (1972). En la Taula 7 es mostren les activitats β -galactosidasa d'aquestes fusions en diferents fons genètics, i també el quocient anaeròbic/aeròbic per cada soca i construcció.

La soca MC4100 Φ (*ald1-lacZ*) en condicions anaeròbiques mostrava una disminució del 87 % de l'activitat β -galactosidasa respecte els nivells aeròbics. La soca MC4100 Φ (*ald1-lacZ*)*arcA2*, que no té proteïna ArcA funcional, tenia una reducció únicament del 12 % respecte els nivells en presència d'oxigen. Es pot comprovar com la introducció de la mutació en el gen *arcA* evita la repressió anaeròbica que existeix en la soca salvatge (Taula 7).

La soca MC4100 Φ (*ald2-lacZ*) mostrava un nivell d'expressió en creixements en absència d'oxigen del 79 % respecte els nivells aeròbics, i la soca MC4100 Φ (*ald3-lacZ*) un 81 %. Es pot concloure doncs que la introducció de les mutacions en les caixes consens per ArcA en el promotor d'*aldA* provoca en gran mesura la desaparició de la repressió anaeròbica produïda pel sistema Arc.

Taula 7. Activitats β -galactosidasa de les fusions tipus salvatge i mutants en diferents fons genètics. Les activitats estan expressades en unitats Miller. Els creixements es van realitzar en hidrolitzat de caseïna a l'1 %.

Soca	Fusió	Activitat β -galactosidasa		Quocient anaeròbic/aeròbic
		Aerobiosi	Anaerobiosi	
MC4100	Φ (<i>ald1-lacZ</i>)	4500	600	0.13
MC4100 <i>arcA2</i>	Φ (<i>ald1-lacZ</i>)	3300	2900	0.88
MC4100	Φ (<i>ald2-lacZ</i>)	3900	3100	0.79
MC4100	Φ (<i>ald3-lacZ</i>)	3450	2800	0.81

5.4.2. Estudi de retardació de la mobilitat electroforètica

L'efecte de la introducció de les mutacions en el promotor d'*aldA* va ser també analitzat mitjançant experiments de retardació de la mobilitat electroforètica en gel. Aquests assajos van dur-se a terme de la mateixa manera que anteriorment (Resultats 5.1.). Com a control de l'experiment es va usar el mateix fragment de 363 pb, pertanyent al promotor d'*aldA* tipus salvatge (anomenat ald1). Els fragments mutats en les caixes d'ArcA van anomenar-se ald2 (obtingut en la mutagènesi dirigida amb el primer ALD2) i ald3 (obtingut en la mutagènesi dirigida amb el primer ALD3). Aquests fragments també tenien un tamany de 363 pb, de manera que les possibles diferències en la mobilitat electroforètica del DNA només siguessin causades per la introducció de les mutacions. Els resultats es mostren a la Figura 25.

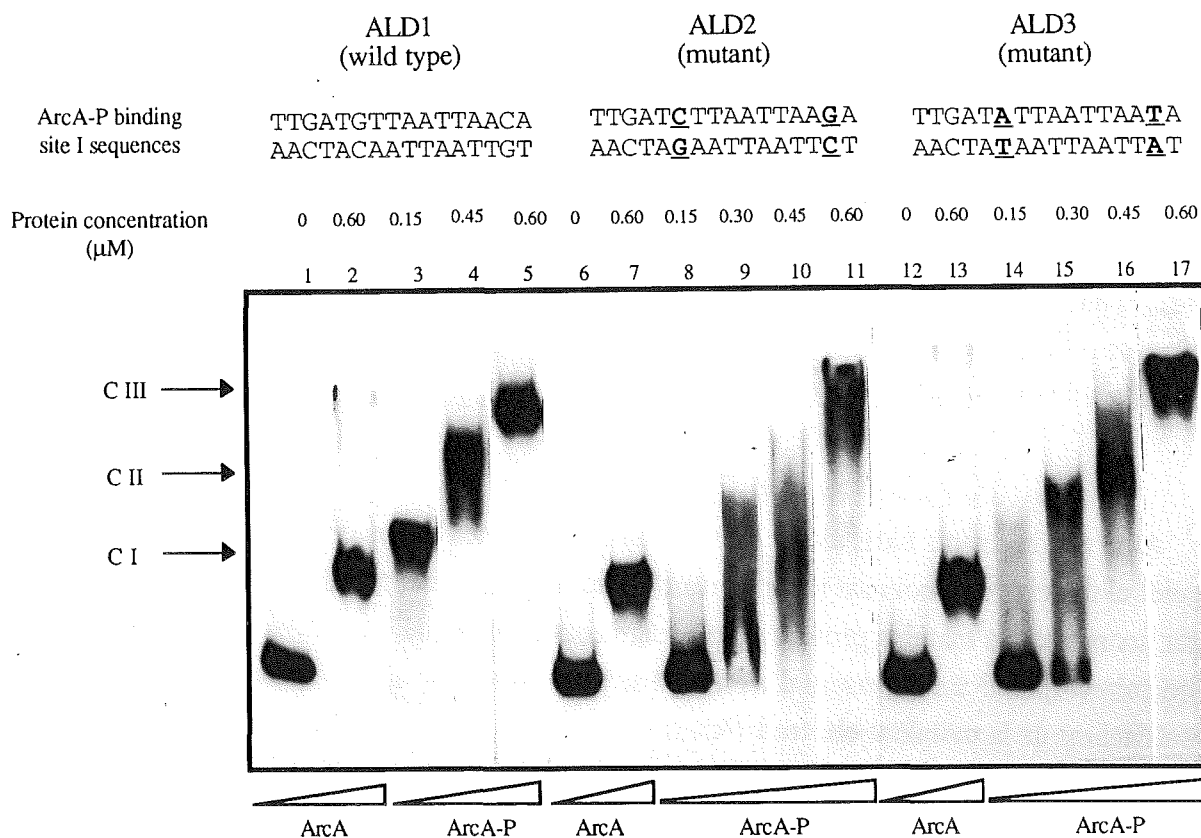


Figura 25. Assaig de retardació en gel del promotor d'*aldA* tipus salvatge i dels promotors mutats amb les proteïnes ArcA i ArcA-P. La retardació es va realitzar en un gel de poliacrilamida al 5%. El DNA marcat amb ³²P estava a una concentració de 2.5 nM. Cada reacció contenia un excés 500 molar de DNA d'esperma de salmó. Les mescleres de reacció contenien diferents quantitats de proteïna (ArcA o ArcA-P), que s'indiquen a sobre de cada carril. Es mostren les seqüències dels promotors, i en els promotors mutats estan subratllades les posicions modificades. La migració dels tres complexos DNA-ArcA-P es mostra com C I (complexe I), C II (complexe II) i C III (complexe III). Els triangles a la part inferior de la figura mostren quantitats creixents d'ArcA o d'ArcA-P. Carrils 1-5: promotor ald1 (tipus salvatge); carrils 6-11: promotor ald2 (mutant) i carrils 12-17: promotor ald3 (mutant).

La unió relativament inespecífica d'ArcA no fosforilada al promotor d'*aldA* es manté en els dos promotors mutats. Contràriament, la retardació de la mobilitat electroforètica dels fragments mutats de DNA requereix una concentració d'ArcA-P més elevada que en el cas del promotor tipus salvatge. A concentracions de 0.15 μM només s'observa retardació en el promotor salvatge, però no en els mutats.

Els complexos de retardació en el fragment *ald1* es formen a concentracions de 0.15 μM (complexe I), 0.45 μM (complexe II) i 0.60 μM (complexe III). El fragment *ald2* forma un complexe similar al complexe I a una concentració de 0.30 μM ArcA-P. Estructures semblants als complexos II i III es formen a concentracions superiors. És remarcable, però, que en tots els casos els complexos tenen un aspecte més heterogeni que els corresponents complexos formats amb el fragment *ald1*.

El mutant *ald3* també mostra una afinitat per ArcA-P disminuïda, tot i que amb un patró diferent a *ald2*. Les diferències entre els dos fragments mutats poden ser conseqüència del diferent contingut en dA+dT.

5.4.3. Estudis de protecció a la DNasa I

Amb els fragments *ald1*, *ald2* i *ald3* es van dur a terme assajos de protecció a la DNasa I, seguint el mateix protocol usat anteriorment (Resultats 5.2.). Els resultats obtinguts mostren que per aconseguir la mateixa protecció dels fragments mutats que del promotor tipus salvatge es requereixen quantitats superiors d'ArcA-P (de manera similar al que succeeix en el cas de les retardacions de la mobilitat electroforètica).

La zona protegida de la digestió amb la DNasa I és la mateixa per tots ells, indicant que pot existir una disminució de l'afinitat dels promotors mutats però que l'especificitat per la zona d'unió al DNA és la mateixa.

6. EXPRESSIÓ DEL SISTEMA *glc* EN FUNCIÓ DE MODULACIONS DEL METABOLISME CENTRAL CONFLUENT EN L'INTERMEDIARI GLIOXILAT

6.1. Inducció creuada dels sistemes *glc* i *ace*

Els sistemes *glc* i *ace*, mobilitzats respectivament en creixements en glicolat i acetat, tenen en comú la formació del metabolit glioxilat. Aquest compost de dos carbonis és transformat per tots dos sistemes en malat, mitjançant l'activitat enzimàtica malat sintasa. *Escherichia coli* conté dos isoenzims de malat sintasa:

1) malat sintasa A (MSA), codificada pel gen *aceB* del sistema *ace*, i induïda per creixements en acetat, i

2) malat sintasa G (MSG), codificada pel gen *glcB*, pertanyent al sistema *glc* i induïda en creixements en glicolat.

Pel creixement en acetat és indispensable l'existència de l'activitat malat sintasa. L'observació que mutants d'MSA són capaços de créixer en acetat, sempre i quan el locus *glc* no estigui mutat, fa pensar que l'MSG pot suplir la deficiència d'MSA en creixements en acetat.

6.1.1. Anàlisi genètica

Per poder analitzar les implicacions de les dues malat sintases en els metabolismes del glicolat i de l'acetat, es van construir diferents mutants deficients en un o en ambdós isoenzims i es van analitzar els seus fenotips creixent en glicolat i en acetat. Com a soca salvatge es va utilitzar la soca MC4100.

Per construir els mutants d'MSA i el seus derivats es va partir de la soca DV21A05 (Vanderwinkel i DeVlieghere, 1968). Aquesta soca és un mutant puntual d'MSA i té delecionat el locus *glc* (*aceB* i Δ *glc*). Per transducció generalitzada amb bacteriòfag P1 (Miller, 1972) a partir de la soca MC4100 es va seleccionar un derivat, la soca JA166, que havia recuperat el locus *glc* i per tant era capaç de créixer en glicolat. Aquest mutant és només deficient en MSA (*aceB*). Sobre aquesta soca JA166 es va transferir la mutació *glcB* per

transducció generalitzada a partir d'un lisat P1 de la soca JA159 (mutant d'MSG, construït per inserció d'un cassette que confereix resistència a cloramfenicol en la zona codificant del gen *glcB* en el genoma de la soca salvatge MC4100 (Resultats 2.3.)). D'aquesta manera es disposava d'un doble mutant de malat sintasa (*aceB* i *glcB*), anomenat soca JA168, però que mantenia intactes les altres activitats enzimàtiques codificades tant en el locus *glc* com en el locus *ace*.

D'altra banda, es va partir de la soca DV21A01, mutant puntual d'isocitrat liasa que té deletionat el locus *glc* (*aceA* i Δglc) (Vanderwinkel i DeVlieghere, 1968). Sobre aquest es va transferir el locus *glc* a partir de la soca MC4100 per transducció generalitzada amb el bacteriòfag P1. Aquesta transducció va donar lloc a la soca JA165, que únicament té mutat el gen que codifica la isocitrat liasa (*aceA*). Amb un lisat P1 de la soca JA159 es va transduïr la mutació *glcB*, donant lloc a la soca JA167. Aquesta soca és un mutant d'isocitrat liasa i d'MSG (*aceA* i *glcB*), però té la resta de gens dels sistemes *ace* i *glc* intactes.

Amb tots aquests mutants es van dur a terme cultius en medi mínim sòlid per comprovar quin era el seu fenotip. Per cada soca es va estudiar la capacitat de creixement en plaques d'acetat i de glicolat. El control del creixement es va fer en plaques de medi ric LB. Els resultats es mostren a la Taula 8.

Taula 8. Evidències genètiques de la inducció creuada dels sistemes *glc* i *ace*

Soques	Fenotip relevant	Origen	Activitats enzimàtiques deficientes	Fenotip	
				Glicolat	Acetat
MC41000	soca salvatge	Casadaban, 1976	-	+	+
JA159	<i>glcB</i>	Aquest estudi	MSG	+	+
DV21A05	<i>aceB</i> Δglc	Vanderwinkel i DeVlieghere, 1968	MSA, MSG, GOX	-	-
JA166	<i>aceB</i>	(MC4100) P1 x DV21A05	MSA	+	+
JA168	<i>aceB glcB</i>	(JA159) P1 x JA166	MSA, MSG	-	-
DV21A01	Δglc , <i>aceA</i>	Vanderwinkel i DeVlieghere, 1968	ICL, MSG, GOX	-	-
JA165	<i>aceA</i>	(MC4100) P1 x DV21A01	ICL	+	-
JA167	<i>aceA glcB</i>	(JA159) P1 x JA165	ICL, MSG	+	-

La soca salvatge MC4100 és capaç de créixer en ambdues fonts de carboni. El mateix succeeix amb els mutants senzills d'MSA (soca JA166) o d'MSG (soca JA159), que tot i tenir una de les malat sintases mutada són capaços de créixer tant en glicolat com en acetat, possiblement mobilitzant l'altre isoenzim (Taula 8).

Tantmateix, el doble mutant MSA/MSG, la soca JA168, és incapaç de créixer en glicolat i acetat com a fonts de carboni.

El mutant DV21A01, deficient en isocitrat liasa i delecionat en el locus *glc*, és incapaç de créixer en glicolat, per manca del sistema *glc*, i en acetat, per manca del primer enzim del cicle del glioxilat. Tots els mutants derivats d'ell són incapços de créixer en acetat. Quan es fa positiu en el locus *glc*, soca JA165, recupera la capacitat de créixer en glicolat. La soca JA167, que a més de la mutació isocitrat liasa és malat sintasa G negativa, continua mantenint la capacitat de créixer a expenses de glicolat, presumiblement usant l'MSA.

D'aquesta anàlisi genètica es pot deduir que tot i que cada isoenzim de malat sintasa està especialment dissenyat per mobilitzar-se en les condicions de creixement que li són específiques, en absència d'un isoenzim, l'altre és capaç de suplir la funció del primer, la qual cosa suggereix la inducció creuada dels sistemes *glc* i *ace*. Aquest fenomen és recíproc: MSA es mobilitza en mutants *glcB* creixent en acetat i glicolat (soca JA159) i MSG actua en mutants *aceB* creixent tant en acetat com en glicolat (soca JA166).

6.1.2. Anàlisi enzimàtica

Com s'ha comprovat en l'anàlisi genètica (Resultats 6.1.1.), els dos isoenzims de malat sintasa són capaços de complementar recíprocament les seves funcions. Per aquest motiu es va creure adient mesurar les activitats enzimàtiques malat sintasa A, malat sintasa G i glicolat oxidasa en creixements en diferents fonts de carboni.

Està descrit que el sistema *glc* està induït per creixement en glicolat. El glicolat és oxidat a glioxilat per la glicolat oxidasa. A causa de la convergència dels metabolismes de glicolat i acetat en la formació de glioxilat i l'existència a *Escherichia coli* d'una activitat glioxilat reductasa (GXR) constitutiva que transforma glioxilat en glicolat, es va creure d'interès analitzar la inducció del sistema *glc* per creixement de la soca tipus salvatge MC4100 en aquestes fonts de carboni relacionades, és a dir, en creixements en acetat i en glioxilat. Paral·lelament a aquests estudis es van realitzar creixements en medi basal (hidrolitzat de caseïna al 0.5 %) i en el mateix medi en presència de l'inductor glicolat.

En els extractes cel.lulars corresponents a cada cultiu es van determinar les activitats enzimàtiques GOX i malat sintasa. L'activitat malat sintasa determinada directament en els extractes cel.lulars correspon a l'activitat total, que és la suma de les activitats dels dos isoenzims que hi són presents. En aquest treball, la quantificació específica de l'isoenzim MSG es va dur a terme per mesura de l'activitat malat sintasa després de la inactivació tèrmica de l'MSA i també per electroimmunodifusió amb anticossos específics contra l'isoenzim MSG.

Els resultats obtinguts es mostren a la Taula 9.

Taula 9. Activitats enzimàtiques de la soca salvatge MC4100 creixent en diferents fonts de carboni. Els valors d'activitat s'expressen en mU/mg de proteïna. MST: malat sintasa total; MSA: malat sintasa A; MSG: malat sintasa G i GOX: glicolat oxidasa.

Creixement	MST (MSA + MSG)	MSG	GOX
Caa 0.5%	180	30	<5
+ Glicolat 30mM	450	300	40
+Glioxilat 30mM	300	150	18
+ Acetat 30 mM	220	50	<5

Els resultats mostren que la major inducció del sistema *glc* s'observa quan la soca salvatge creix en presència de glicolat. Els creixements en presència de glioxilat són, en menor grau, condicions inductores del sistema *glc*. El glioxilat provoca una inducció que és aproximadament la meitat que la provocada pel creixement en presència de glicolat.

Es pot observar que l'acetat indueix el sistema, però només a nivell de l'MSG. L'addició d'acetat al medi de cultiu sembla només induir una de les activitats enzimàtiques codificades en el locus *glc*. Aquest fet podria suggerir l'existència d'una regulació independent de la glicolat oxidasa i de la malat sintasa G, almenys respecte a la inducció en presència d'acetat. Cal tenir en compte que pel creixement en acetat no es necessita la participació de la GOX. Una altra interpretació d'aquest resultat és que com l'activitat específica GOX és molt baixa, inclús en creixements en glicolat, en condicions de menor inducció, com és el cas de l'acetat, el mètode de determinació no és prou sensible per detectar tant poques unitats d'activitat enzimàtica.

6.1.3. Expressió del sistema *glc* per acetat mitjançant fusions transcripcionals

Els resultats de les activitats enzimàtiques glicolat oxidasa i malat sintasa G obtinguts anteriorment (Resultats 6.1.2.) semblen indicar que la mobilització de la malat sintasa G és independent a la de la glicolat oxidasa en creixements en acetat, indicant la possible regulació independent dels gens que codifiquen aquests enzims. Tantmateix, la manca d'expressió de β -galactosidasa a partir de la fusió corresponent a la regió 5' del gen *glcB* ($\Phi(\textit{glcB-lacZ})$) en cultius de la soca MC4100 crescuts en presència de glicolat semblaven no recolzar aquesta hipòtesi (Resultats 3.2.). Per tal d'esbrinar si existeix un promotor alternatiu per l'expressió de *glcB* induïda per acetat, es va plantejar la construcció d'una nova fusió de la zona 5' del gen *glcB* que abarqués una regió més àmplia, incluint tota la seqüència codificant del gen anterior (*glcG*). Aquesta nova fusió, que va ser anomenada $\Phi(\textit{glcBA-lacZ})$, va transferir-se al fons genètic de la soca MC4100 i la seva expressió es va assajar en cultius induïts tant per la presència de glicolat com d'acetat. Els resultats presentats a la Taula 10 mostren que la fusió $\Phi(\textit{glcBA-lacZ})$ no expressa activitat β -galactosidasa en cap cas.

Aquests resultats descarten definitivament l'expressió diferencial de *glcB* respecte els gens que codifiquen la glicolat oxidasa. Per tant, la inducció creuada del sistema *glc* recau sobre el promotor de *glcD*.

Taula 10. Anàlisi de la fusió $\Phi(\textit{glcBA-lacZ})$, corresponent a la regió 5' del gen *glcB*, mitjançant mesura de l'activitat β -galactosidasa en cultius de la soca MC4100 induïts per acetat o glicolat. Caa: hidrolitzat de caseïna 0.5 %; Caa + Ace: hidrolitzat de caseïna 0.5 % més acetat 30 mM; Caa + Glc: hidrolitzat de caseïna 0.5 % més glicolat 30 mM. Les activitats s'expressen en unitats Miller.

Soca	Fusió	Activitat β -galactosidasa		
		Caa	Caa + Ace	Caa + Glc
MC4100	$\Phi(\textit{glcBA-lacZ})$	<10	<10	<10
MC4100	$\Phi(\textit{glcD-lacZ})$	150	375	1800

Amb la finalitat d'esbrinar si la inducció de l'operó *glc* en presència d'acetat està controlada pel mateix acetat o per la formació de glioxilat, es va transduir la fusió $\Phi(\textit{glcD-lacZ})$ al fons genètic de diferents mutants del sistema *ace*: soca JA165 (deficient en isocitrat liasa), JA166 (deficient en malat sintasa A) i JA168 (deficient en malat sintasa A i G). Aquestes soques es van fer créixer en cultius d'hidrolitzat de caseïna al 0.05 %, d'hidrolitzat de caseïna

al 0.05 % més acetat i d'hidrolitzat de caseïna al 0.05 % més glicolat. Es van assajar les activitats β -galactosidasa d'aquests cultius i en els seus extractes crus corresponents es va dur a terme la detecció immunològica de la malat sintasa G (Figura 26).

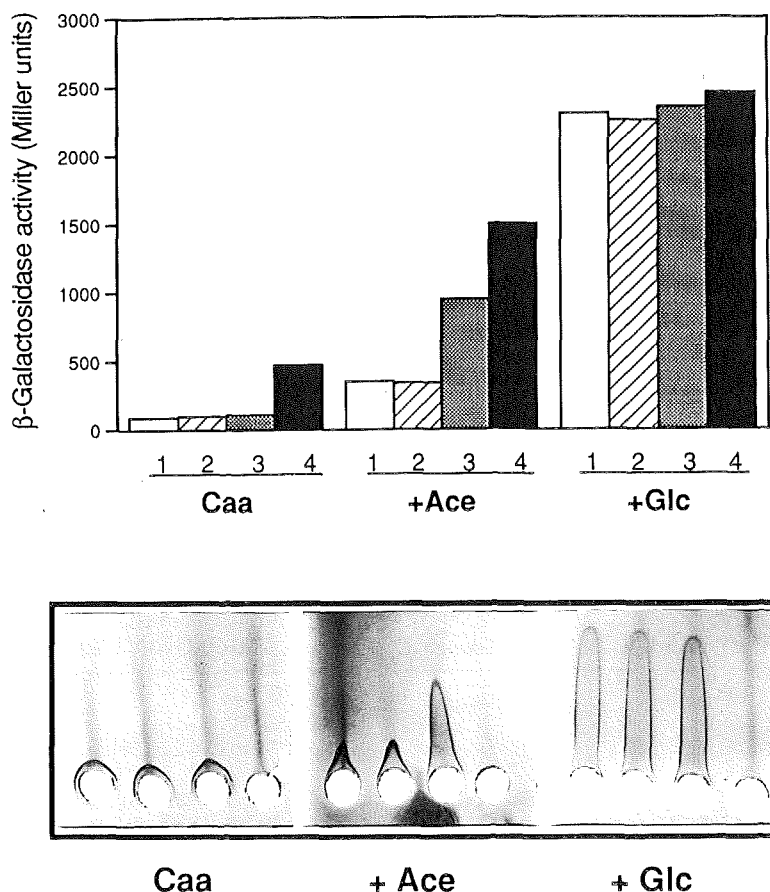


Figura 26. Anàlisi de la inducció de l'operó *glc* per acetat en diferents mutants *ace*. Els nivells d'activitat β -galactosidasa de les soques MC4100 (1), JA165 (2), JA166 (3), i JA168 (4) contenen la fusió $\Phi(glcD-lacZ)$ creixent en Caa 0.05% únicament (Caa) o en aquest medi en presència d'acetat (Ace) o glicolat (+Glc) es mostren en el diagrama de barres. Les activitats estan expressades en unitats Miller. La quantificació de l'immunoprecipitat d'MSG, duta a terme en els corresponents extractes cel.lulars, es mostra a la part inferior de la figura. Les plaques d'electroimmunodifusió contenen un 1.3 % de la fracció γ -globulina anti-malat sintasa G; a cada carril es van aplicar 65 μ g dels extractes cel.lulars.

En la soca JA165 la fusió $\Phi(glcD-lacZ)$ té el mateix patró d'expressió en qualsevol condició de creixement que en la soca MC4100. La mutació en la isocitrat liasa no afecta l'expressió de l'operó *glc*, la qual cosa indica que o bé l'acetat o bé un intermediari metabòlic format fins a la síntesi d'isocitrat és l'inductor de *glcD*.

En la soca JA166 Φ (*glcD-lacZ*), on només MSG està present, el contingut intracel·lular de glioxilat està augmentat quan el bacteri creix en presència d'acetat. Aquesta soca expressa 2.5 vegades més d'activitat β -galactosidasa en presència d'acetat, ja sigui per efecte del glioxilat acumulat o pel glicolat format per acció de la glioxilat reductasa (Introducció 2.1.).

La soca JA168 Φ (*glcD-lacZ*) té un patró d'inducció diferent a la soca tipus salvatge, no només en acetat sinó també en hidrolitzat de caseïna. En aquest mutant deficient en ambdós isoenzims de malat sintasa la quantitat de glioxilat no metabolitzat és encara major. Això es reflecteix en un augment de 5 vegades de la inducció de l'operó en presència d'acetat. L'activació de Φ (*glcD-lacZ*) en hidrolitzat de caseïna suggereix la formació de glioxilat en la degradació d'aminoàcids.

En tots els casos la detecció immunològica de la malat sintasa G estava d'acord amb els nivells d'expressió transcripcional de Φ (*glcD-lacZ*) (Figura 26).

6.2. Efecte de diferents fonts de carboni sobre l'expressió de l'operó *glc*

S'ha observat que el sistema *glc* està induït per fonts de carboni que generen glioxilat com a intermediari (glicolat, glioxilat, acetat) (Resultats 6.1.2.) i reprimat per altres (glucosa) (Resultats 4.2.2.). En aquest punt del treball es va creure interessant analitzar l'efecte d'altres fonts de carboni en l'expressió dels gens de l'operó *glc*. Els compostos assajats inclouen fonts de carboni que no provoquen repressió per catabolisme, com és el cas del succinat o de la D-xilosa, així com d'altres compostos estructuralment relacionats amb el glicolat, com són els α -hidroxiàcids L-lactat, D-lactat i glicerat.

Es van fer creixements de la soca MC4100 Φ (*glcD-lacZ*) en glucosa, D-xilosa, succinat, glicolat, glioxilat, L-lactat, hidrolitzat de caseïna, acetat, D-lactat i glicerat. En tots ells es va assajar l'activitat β -galactosidasa i es va quantificar immunològicament la malat sintasa G present als extractes corresponents.

Els resultats es mostren a la Figura 27. La màxima expressió de Φ (*glcD-lacZ*) s'observa en glicolat i glioxilat, tot i que la inducció per aquest darrer compost és la meitat que l'exercida pel glicolat. La glucosa és la font de carboni que dona uns nivells més baixos d'expressió (20 unitats Miller d'activitat β -galactosidasa), mentre que en fonts no repressores (D-xilosa i succinat) els nivells d'expressió són lleugerament superiors en el cas del succinat i similars als nivells observats en cultius d'hidrolitzat de caseïna.

Entre els compostos estructuralment relacionats amb el glicolat, el D-lactat i el glicerat provoquen una inducció del sistema que és, respectivament, el 30 % i el 40 % dels nivells assolits en creixements en glicolat. L'L-lactat no indueix l'expressió de l'operó, la qual cosa sembla indicar que la conformació espacial dels compostos influeix en la seva capacitat d'activació de la transcripció. L'acetat, àcid monocarboxílic de dos carbonis, tot i ser un inductor de l'expressió, només aconsegueix uns nivells sis vegades menors als que produeix el glicolat. La inducció provocada per qualsevol font de carboni era totalment abolida en el fons genètic de la soca JA154 (*glcC*). La detecció de l'MSG per electroimmunodifusió en tots els extractes està d'acord amb els nivells d'expressió de $\Phi(glcD-lacZ)$ (Figura 27).

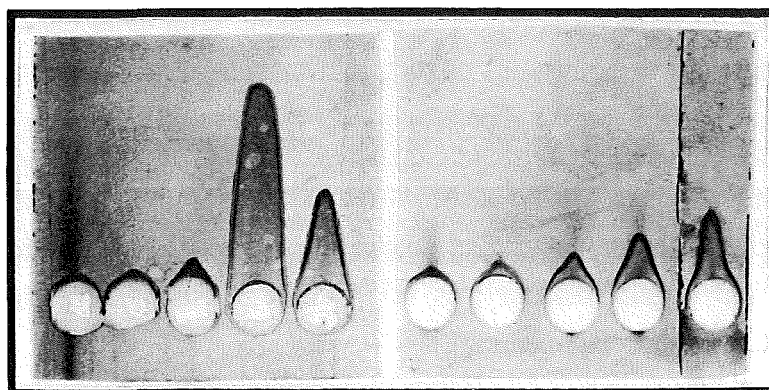
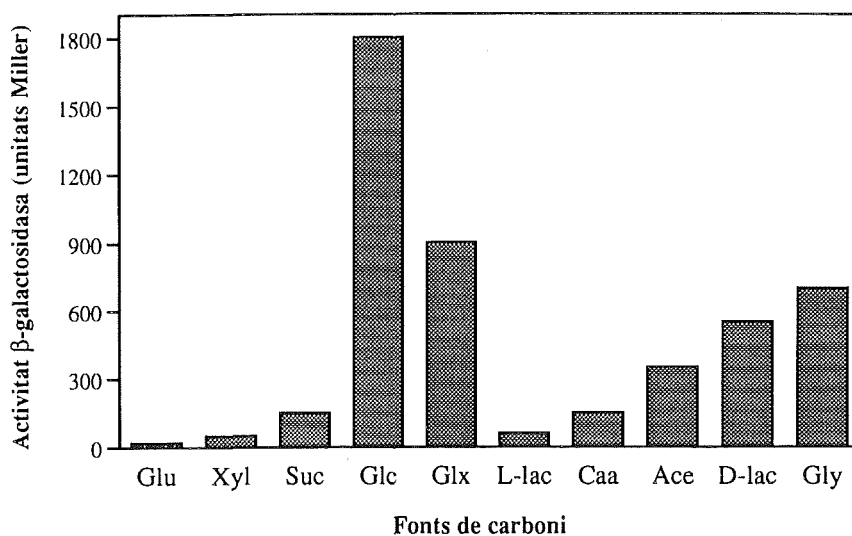


Figura 27. Anàlisi de la inducció de l'operó *glc* per diferents fonts de carboni en la soca MC4100. Els nivells d'activitat β -galactosidasa de la soca MC4100 $\Phi(glcD-lacZ)$ creixent en glucosa (Glu), xilosa (Xyl), succinat (Suc), glicolat (Glc), glioxilat (Glx), L-lactat (L-lac), hidrolitzat de caseïna (Caa), acetat (Ace), D-lactat (D-lac) i glicerat (Gly) es mostren en el diagrama de barres. Les activitats estan expressades en unitats Miller. La quantificació de l'immunoprecipitat d'MSG, duta a terme en els corresponents extractes cel·lulars, es mostra a la part inferior de la figura. Les plaques d'electroimmunodifusió contenien un 1.3 % de la fracció γ -globulina anti-malat sintasa G; a cada carril es van aplicar 65 μ g dels extractes cel·lulars.

6.3. Efecte de mutacions en els gens estructurals del sistema *glc* sobre el patró d'expressió de l'operó *glcDEFGBT*

El glicolat, i en menor mesura el glioxilat, són els principals inductors de l'expressió de l'operó *glc*. Ja que tots dos compostos són interconvertibles dins la cèl.lula per acció dels enzims glicolat oxidasa i glioxilat reductasa (Introducció 2.1.), és difícil saber quin dels dos compostos, o bé tots dos, és la molècula efectora de l'activació de la proteïna GlcC. Una primera aproximació a aquest estudi es va fer analitzant l'expressió de la fusió $\Phi(\textit{glcD-lacZ})$ en el fons genètic de mutants en les diferents subunitats de la glicolat oxidasa; aquests mutants són incapaços de transformar glicolat en glioxilat. La fusió transcripcional es va transferir a les soques JA155 (*glcD::cat*), JA156 (*glcE::cat*), i JA157 (*glcF::cat*). Es van assajar les activitats β -galactosidasa en creixements en hidrolitzat de caseïna, hidrolitzat de caseïna més glicolat i hidrolitzat de caseïna més glioxilat

Els resultats obtinguts van ser idèntics en els tres mutants de glicolat oxidasa (Taula 11). Sorprenentment en creixements en hidrolitzat de caseïna, l'expressió de $\Phi(\textit{glcD-lacZ})$ s'elevava als mateixos nivells assolits en condicions de màxima inducció, és a dir, en creixements en presència de glicolat. En presència de glioxilat les soques mutants tenien un nivell d'expressió menor que la soca salvatge.

Per tal de comprovar que l'augment de la inducció del sistema no és conseqüència de l'efecte de la presència del casset CAT en els gens mutats, es va assajar l'activitat β -galactosidasa de la fusió $\Phi(\textit{glcD-lacZ})$ en la soca JA159 (*glcB::cat*). Es va comprovar que els nivells d'activitat eren els mateixos que en la soca salvatge MC4100. El mateix experiment es va dur a terme amb mutants d'altres gens de l'operó *glc* no implicats en l'activitat glicolat oxidasa, com són la soca JA158 (*glcG::cat*) i TP02 (*glcT::cat*) i els resultats van ser els mateixos. Es pot concloure, doncs, que mutacions en els gens estructurals que codifiquen la glicolat oxidasa comporten un augment de l'expressió en hidrolitzat de caseïna.

Per tal de comprovar si l'increment en l'expressió de $\Phi(\textit{glcD-lacZ})$ en aquests mutants era causat per un efecte dosi de la proteïna GlcC (Resultats 4.1.2.) es va transferir la fusió del gen regulador $\Phi(\textit{glcC-lacZ})$ al fons genètic de les soques JA155, JA156 i JA157. Es van assajar les activitats β -galactosidasa en creixements en hidrolitzat de caseïna, hidrolitzat de caseïna més glicolat i hidrolitzat de caseïna més glioxilat. Es va poder comprovar que els nivells d'activitat eren els mateixos que en la soca salvatge MC4100, per tant l'efecte observat no era causat per un augment en la taxa de transcripció de *glcC* (Taula 11).

Aquests resultats semblaven indicar que el metabolisme de l'hidrolitzat de caseïna genera glicolat, el qual s'acumularia intracel.lularment en els mutants glicolat oxidasa. L'increment en la transcripció observat seria provocat per un increment en els nivells intracel.lulars de la molècula efectora. Es coneixen dues vies per la formació intracel.lular de glicolat: a partir de l'oxidació de glicolaldehid per acció de l'aldehid deshidrogenasa, producte del gen *aldA*, o bé a partir de la reducció de glioxilat catalitzada per l'enzim constitutiu glioxilat reductasa. La utilització de mutants en aquest gens permetia estudiar quina d'elles és la responsable de la formació i acúmulo de glicolat en els mutants GOX.

Pel que fa a la glioxilat reductasa, el gen que la codifica no ha estat ni localitzat ni descrit, i per tant no es disposa de mutants que permetin analitzar les conseqüències de la introducció d'una mutació en aquest gen. Aquest aspecte, per tant, no s'ha abordat en aquest treball, tot i que actualment és motiu d'estudi en el nostre laboratori.

Per esbrinar si la formació i acumulació de glicolat en els mutants glicolat oxidasa prové de de l'oxidació de glicolaldehid es va transferir la mutació *aldA* al fons genètic de la soca MC4100 i de les soques mutants GOX (JA155, JA156 i JA157) per transducció generalitzada a partir d'un lisat de P1 de la soca ECL493. Els resultats obtinguts descarten la formació de glicolat a partir de glicolaldehid, ja que el mutant aldehid deshidrogenasa té el mateix patró d'expressió que la soca salvatge i el doble mutant no varia l'expressió de la soca GOX negativa (Taula 11).

Taula 11. Efecte de mutacions en els gens *glcD* i *aldA* en l'expressió de les fusions transcripcionals del sistema *glc*. Els creixements es van realitzar en hidrolitzat de caseïna 0.5 % (Caa), hidrolitzat de caseïna més glicolat (Caa + Glc) i hidrolitzat de caseïna més glioxilat (Caa + Glx). Les activitats s'expressen en unitats Miller.

Soca	Fusions	Activitat β -galactosidasa		
		Caa	Caa + Glc	Caa + Glx
MC4100	$\Phi(glcD-lacZ)$	150	1800	900
JA155	$\Phi(glcD-lacZ)$	1830	1800	450
MC4100	$\Phi(glcC-lacZ)$	90	50	60
JA155	$\Phi(glcC-lacZ)$	92	50	65
MC4100 <i>aldA</i>	$\Phi(glcD-lacZ)$	160	1800	890
JA155 <i>aldA</i>	$\Phi(glcD-lacZ)$	1850	1800	450

DISCUSSIÓ

V. DISCUSSIÓ

1. LA GLICOLAT OXIDASA D'*Escherichia coli* ÉS UN ENZIM HETEROLIGOMÈRIC LLIGAT A LA MEMBRANA CITOPLASMÀTICA

La glicolat oxidasa d'*Escherichia coli* és un enzim associat a la membrana plasmàtica (Sallal i Nimer, 1989) que catalitza l'oxidació de glicolat a glioxilat. El nom d'oxidasa pot induir a error en quant al seu mecanisme d'acció. Tot i anomenar-se com a oxidasa, l'enzim no transfereix els electrons directament a l'oxigen (Lord, 1972). El potencial reductor dels electrons alliberats en la reacció és transferit a algun component, encara per determinar, de la cadena de transport electrònic. Per aquest motiu el nom més adequat per l'enzim podria ser glicolat deshidrogenasa. L'estructura de la glicolat oxidasa d'*Escherichia coli* no ha estat descrita encara.

S'han descrit, caracteritzat, purificat i cristal·litzat glicolat oxidases de plantes (Sandalova i Lindqvist, 1993) (Nishimura *et al*, 1983) i mamífers (Schwam *et al*, 1979). En aquests organismes superiors aquests enzims són flavoenzims peroxisomals que catalitzen l'oxidació d' α -hidroxiàcids als corresponents α -cetoàcids. Estructuralment són oligòmers que en solució formen tetràmers o octàmers de subunitats idèntiques de 43000 Da. En plantes verdes participen en la fotorespiració, que és un procés que disminueix la taxa neta de fototsíntesi. En animals participen en la producció d'oxalat, per la qual cosa la seva inhibició sembla ser beneficiosa en el tractament de malalties causades per una sobreproducció d'oxalat càlcic. Els mecanismes d'acció de les glicolat oxidases d'organismes superiors són totalment

coneguts i els seus centres catalítics han estat localitzats i estudiats amb profunditat (Macheroux *et al*, 1993). El nom d'oxidasa és apropiat en aquest cas ja que s'ha comprovat els electrons alliberats en la reacció són transferits directament a l'oxigen.

En aquest treball es presenta la primera descripció de l'estructura de la glicolat oxidasa d'*Escherichia coli*. S'ha comprovat que l'estructura d'aquesta oxidasa procariota és molt diferent de la que presenten les glicolat oxidases d'organismes superiors. L'enzim és un oligòmer format per tres subunitats diferents. Els productes gènics de *glcD*, *glcE* i *glcF* formen part de la proteïna glicolat oxidasa, ja que l'activitat determinada en extractes crus de la soca salvatge MC4100 s'aboleix quan s'interromp algun d'aquests tres gens. Aquestes mutacions independents en cadascun dels gens impedeixen el creixement en glicolat com a única font de carboni i energia.

L'existència dels productes gènics de *glcD* i *glcE* s'ha comprovat per expressió dels mateixos en un sistema de maxicèl.lules. Aquest sistema no ha permès detectar el producte de *glcF*. La proteïna GlcF sí que s'ha pogut detectar, junt amb GlcD i GlcE, a la fracció de membrana de la soca MC4100 crescuda en presència de glicolat i de la soca JA151 transformada amb el plasmidi pTP21. Els pesos moleculars de totes elles concorden amb els deduïts de la seqüència d'aminoàcids.

En l'obtenció de la fracció de membrana de la soca MC4100 crescuda en glicolat es va poder detectar el 90 % de l'activitat glicolat oxidasa en aquesta fracció. En la soca JA151 transformada amb el plasmidi pTP21, on existeix una sobreexpressió dels gens que codifiquen les subunitats de l'enzim, l'activitat glicolat oxidasa està repartida al 50 % entre el sobrenedant i les membranes. Aquest fet indica que tot i tractar-se d'un enzim associat a membrana, les seves subunitats són catalíticament actives en el citosol cel.lular. Aquesta observació permet pensar que podria existir una proteïna que fós l'anclatge de les subunitats catalítiques a la membrana plasmàtica, que no participaria en la catàlisi però que afavoriria el mecanisme de transferència d'electrons a través de la cadena de transport electrònic. Un model alternatiu seria la unió a la membrana mitjançant interaccions amb fosfolípids.

Per tal de poder intuir el mecanisme d'acció de la glicolat oxidasa, es va procedir a estudiar les característiques individuals de les seves subunitats.

GlcD i GlcE, de pesos moleculars determinats experimentalment 56000 i 40000 respectivament, són proteïnes amb una alta similitud a diferents deshidrogenases de varis organismes, sobretot amb la D-lactat deshidrogenasa d'*Escherichia coli*. Les seves freqüències d'utilització de codó no són gaire elevades (0.49 per *glcD* i 0.60 per *glcE*) tot i que tant en els experiments d'expressió en maxicèl.lules com en l'obtenció de la fracció de membrana s'expressen sense cap dificultat i són fàcilment detectables. El seu perfil hidropàtic no mostra

cap estructura que pugui ser compatible amb una localització transmembranal. Tenen el perfil de proteïnes solubles.

En l'estructura primària de la proteïna GlcD s'ha pogut detectar un possible lloc d'unió a nucleòtids de flavina, segons decriuen Branden i Tooze (1991). Aquesta seqüència d'unió a aquests nucleòtids també es troba present en altres deshidrogenases, com són la glicerol-3-fosfat deshidrogenasa anaeròbica, la succinat deshidrogenasa o la lipoamida deshidrogenasa. Tot i que s'ha descrit que enzims que poden tenir com a cofactors els nucleòtids de flavina FAD i FMN veuen augmentada la seva activitat per addició en el medi de reacció d'aquests dos compostos, p.e. la glicerol-3-fosfat deshidrogenasa anaeròbica d'*Escherichia coli* (Schryvers i Weiner, 1981), aquest fenomen no es va observar en l'assaig de glicolat oxidasa. Un comportament similar al de la glicolat oxidasa s'ha descrit per altres enzims, p.e. fumarat reductasa d'*Escherichia coli* (Miki i Lin, 1973). Aquest fet però no descarta que GlcD estigui unit a FAD o FMN. Pot ser que la unió entre l'apoenzim i el cofactor sigui covalent i que la constant d'afinitat entre ells sigui molt alta. En aquestes condicions l'enzim sempre estaria saturat de cofactor, i per tant l'addició del mateix en la cuveta d'assaig no variaria l'activitat enzimàtica. La fumarat reductasa està unida covalentment al cofactor FAD (Weiner i Dickie, 1979). Per conèixer la naturalesa de la unió entre FAD o FMN i GlcD s'hauria d'analitzar l'espectre d'absorció de l'enzim purificat en condicions no denaturants, després d'un tractament suau amb detergents i després d'un tractament dràstic.

També s'ha pogut identificar en GlcD i GlcE una seqüència altament similar a la dels centres actius de varies òxidoreductases, entre elles la glicerol-3-fosfat deshidrogenasa anaeròbica, la succinat deshidrogenasa i la fumarat reductasa d'*Escherichia coli*. Els dos darrers enzims són, igual que la glicolat oxidasa, enzims localitzats a la membrana d'*Escherichia coli* i estan implicats en processos respiratoris.

Està descrit que la glicolat oxidasa pot utilitzar D-lactat com a substrat alternatiu al glicolat (Lord, 1972). El D-lactat i el glicolat són estructuralment similars i és possible que la glicolat oxidasa i la D-lactat deshidrogenasa evolucionessin d'un ancestre comú que catalitzés indistintament l'oxidació dels dos compostos. En aquest sentit, un mutant D-lactat deshidrogenasa podria ser complementat amb la glicolat oxidasa. La similitud de GlcD i GlcE amb la D-lactat deshidrogenasa possiblement es basa en que comparteixen el lloc de reconeixement pel D-lactat. Per aquest motiu, i pel fet que s'han identificat seqüències altament conservades en òxidoreductases i centres d'unió a flavines, el dímer GlcDE sembla formar el centre catalític de la glicolat oxidasa, tot i que per sí sol no és capaç de catalitzar l'oxidació de glicolat a glioxilat. Perquè això sigui possible es requereix la presència de GlcF.

El producte gènic de *glcF* no va poder analitzar-se mitjançant experiments d'expressió en el sistema de maxicèl·lules; tantmateix aquesta proteïna va poder ser detectada en la fracció

de membrana de la soca JA151 a partir de la sobreexpressió del clon pTP21. El fet que l'experiment d'expressió en maxicèl.lules no permetés la identificació de GlcF pot explicar-se en base a la seva freqüència en l'ús del codó. El "codon usage" de *glcF* és major que els de *glcD* i *glcE*, tot i que si s'analitza en els diferents segments de la longitud dels gen, es pot observar com l'extrem 5' té un "codon usage" parcial més baix. Això pot repercutir en la velocitat de síntesi de la proteïna, ja que l'inici de la traducció es veu enlentit a causa dels pocs tRNAs disponibles. Per tant la inducció d'una hora en els experiments en maxicèl.lules pot no ser suficient per la producció d'una quantitat de proteïna detectable per aquest mètode. La quantitat de metionines presents en la proteïna (que són els aminoàcids que provoquen el marcatge radioactiu en aquest experiment) no sembla ser limitant, ja que encara que estiguin presents en un baix % (1.13 %) haurien de ser suficients per poder veure l'aparició de la proteïna marcada (després de setmanes de contactació no s'observa l'aparició de la banda corresponent a l'expressió de *glcF*). En canvi, en l'obtenció de la fracció de membrana el temps no és un factor limitant ja que les cèl.lules es recullen al final de la fase exponencial de creixement i per tant la proteïna GlcF és detectable.

És característic de GlcF el gran contingut en cisteïnes: té un total de 17 residus cisteïnics agrupats en diferents centres. La comparació de la seqüència d'aminoàcids de GlcF amb les proteïnes del banc de dades GenBank mostra que aquesta té una alta similitud amb proteïnes implicades en processos respiratoris, sobretot en les regions riques en cisteïnes. Aquestes formen centres ferrosulfurats no hèmics. A més de la disposició de les cisteïnes en aquestes zones també estan conservades altres posicions. És remarcable l'alta similitud que presenta aquesta subunitat de glicolat oxidasa amb la subunitat ferrosulfurada GlpC de la glicerol-3-fosfat deshidrogenasa d'*Escherichia coli*. El perfil hidropàtic de GlcF no correspon tampoc amb el d'una proteïna integral de membrana.

Per la seva estructura primària, GlcF sembla ser l'acceptor dels electrons alliberats en l'oxidació del glicolat i no estar implicat pròpiament en la catàlisi, que semblaria estar produïda per les subunitats GlcD i GlcE. El baix nivell d'expressió de *glcF* en comparació a *glcD* i *glcE* pot ser un mecanisme d'ajustament de les concentracions relatives de les diferents subunitats, i indicaria que varies unitats de GlcD i GlcE s'assossien a una única molècula de GlcF. L'observació que GlcF és imprescindible per la catàlisi pot estar indicant que la subunitat ferrosulfurada diposa correctament en l'espai a GlcD i GlcE perquè puguin ser catalíticament actives.

De la comparació de les tres subunitats de la glicolat oxidasa amb els bancs de dades de seqüències proteïques s'han identificat tres proteïnes que per la seva estructura primària i per la funció que duen a terme podrien tenir una estructura similar a la glicolat oxidasa. Aquestes són

la D-lactat deshidrogenasa, la glicerol-3-fosfat deshidrogenasa anaeròbica i la fumarat reductasa. Tots tres enzims són flavoproteïnes unides a membrana.

La D-lactat deshidrogenasa d'*Escherichia coli* és una de les primeres deshidrogenases de la cadena respiratòria (Ingledew i Poole, 1984). És una proteïna monomèrica de 571 residus aminoacídics. Conté FAD unit no covalentment i està unida a la membrana citoplasmàtica per interaccions amb els fosfolípids, els quals l'activen.

La glicerol-3-fosfat deshidrogenasa anaeròbica i la fumarat reductasa estan induïdes anaeròbicament per glicerol-3-fosfat i fumarat respectivament. Els dos enzims formen un complex funcional que pot catalitzar la deshidrogenació del glicerol-3-fosfat per formar dihidroxicetona fosfat a expenses de fumarat sense cap altre cofactor addicional (Miki i Lin, 1973). La fumarat reductasa està composta per una subunitat catalítica de 69 kDa unida covalentment al cofactor FAD (Weiner i Dickie, 1979), una subunitat de 27 kDa que conté un centre ferrosulfurat (Cole *et al.*, 1982) i dues petites subunitats molt hidrofòbiques de 15 i 13 kDa que serveixen d'anclatge de les subunitats de 69 i 27 kDa a la membrana citoplasmàtica (Lemire *et al.*, 1982).

La glicerol-3-fosfat deshidrogenasa anaeròbica està composta per tres subunitats. Les dues subunitats catalítiques són GlpA, de 62 kDa, que està unida a FAD, i GlpB, de 43 kDa, unida a FMN. La tercera subunitat, anomenada GlpC, té un alt contingut en cisteïnes (17 residus cisteïnics) que es localitzen en dues regions on es troben disposades amb una estructura de centres ferrosulfurats. Aquesta subunitat està fortament lligada a la membrana citoplasmàtica i és l'anclatge a la membrana del dímer actiu (Cole *et al.*, 1988). L'associació del dímer GlpAB amb GlpC no és quantitativa: la relació molecular entre el dímer catalític i la subunitat ferrosulfurada no és equimolar. L'associació entre els diferents components de l'enzim depèn del seu potencial redox (Kuritzkes *et al.*, 1984).

L'estructura de la glicolat oxidasa i de la glicerol-3-fosfat deshidrogenasa anaeròbica s'assemblen molt. Totes dues presenten dues subunitats catalítiques unides a flavines i una subunitat ferrosulfurada. Entre elles les subunitats tenen un pes molecular molt similar ($Pm_{GlcD}/Pm_{GlpA} = 1.11$; $Pm_{GlcE}/Pm_{GlpB} = 1.07$; $Pm_{GlcF}/Pm_{GlpC} = 1.02$). Els electrons alliberats en la catàlisi són transferits a la subunitat GlcF (glicolat oxidasa) o GlpC (glicerol-3-fosfat deshidrogenasa anaeròbica). La diferència que existeix entre elles és que GlcF és necessària per la catàlisi mentre que el dímer GlpAB és actiu per sí sol i GlpC únicament reb els electrons i els transfereix a la cadena de transport electrònic o al fumarat (en condicions anaeròbiques on el fumarat és l'acceptor d'electrons present). A més GlpC és una proteïna de membrana, amb un perfil hidropàtic molt hidrofòbic, que serveix d'anclatge del dímer actiu a la membrana.

La unió de la glicolat oxidasa a la membrana citoplasmàtica podria establir-se amb fosfolípids de la bicapa lipídica (Figura 1), de la mateixa manera que ho fa la D-lactat deshidrogenasa. En aquest sentit, el fet de no trobar cap subunitat catalítica amb un perfil hidropàtic de proteïnes de membrana fa possible aquesta hipòtesi. La similitud estructural entre l'oxidasa i la deshidrogenasa faria possible que ambdues estiguessin de la mateixa manera unides a la membrana. Actualment s'estan duent a terme experiments en el laboratori per tal d'establir si la unió de la glicolat oxidasa a la membrana citoplasmàtica està mediada per interaccions amb fosfolípids.

Respecte a la similitud de la glicolat oxidasa amb la fumarat reductasa, aquests enzims comparteixen la propietat d'estar units covalentment al cofactor FAD, mostren una alta similitud entre les seves subunitats catalítiques i les seves subunitats ferrosulfurades no estan integrades a la membrana. La unió de la fumarat reductasa a la membrana s'estableix per dos petits polipèptids de 15 i 13 kDa que es localitzen a l'interior de la bicapa lipídica i serveixen d'anclatge de les subunitats catalítiques a la membrana. La reductasa és funcional sense la presència dels dos polipèptids, tot i que aquests modulen l'activitat i l'estabilitat de les subunitats catalítiques.

En aquest sentit, la unió de la glicolat oxidasa a la membrana podria produir-se gràcies a la participació d'un petit polipèptid, que actuaria d'anclatge de la mateixa manera que els dos polipèptids de 15 i 13 kDa ho fan en el cas de la fumarat reductasa. El producte del gen *glcG* és una proteïna amb una massa molecular de 13 kDa. La seva seqüència d'aminoàcids no mostra similitud amb cap grup conegut de proteïnes. GlcG no és necessària per l'activitat glicolat oxidasa, de manera similar al que succeeix amb els petits pèptids de la fumarat reductasa. El seu perfil hidropàtic no és altament hidrofòbic en la seva totalitat, encara que és remarcable una zona d'alta hidrofobicitat a l'extrem carboxiterminal. Aquesta regió podria estar immersa en la membrana plasmàtica i servir d'anclatge per les subunitats catalítiques (Figura 1). La funció de GlcG com a molècula d'anclatge podria estudiar-se utilitzant un mutant deletionat en GlcG i comprovant la distribució subcel·lular de les tres subunitats catalítiques i també purificant GlcG, immobilitzant-la en una columna cromatogràfica i comprovant si les tres subunitats de glicolat oxidasa s'hi addereixen. No és descartable que el pèptid actui en forma de dímer dins la membrana, formant una via de canalització dels electrons. La seva alta freqüència en l'ús del codó podria permetre una síntesi proteica més ràpida i eficient respecte les subunitats de glicolat oxidasa.

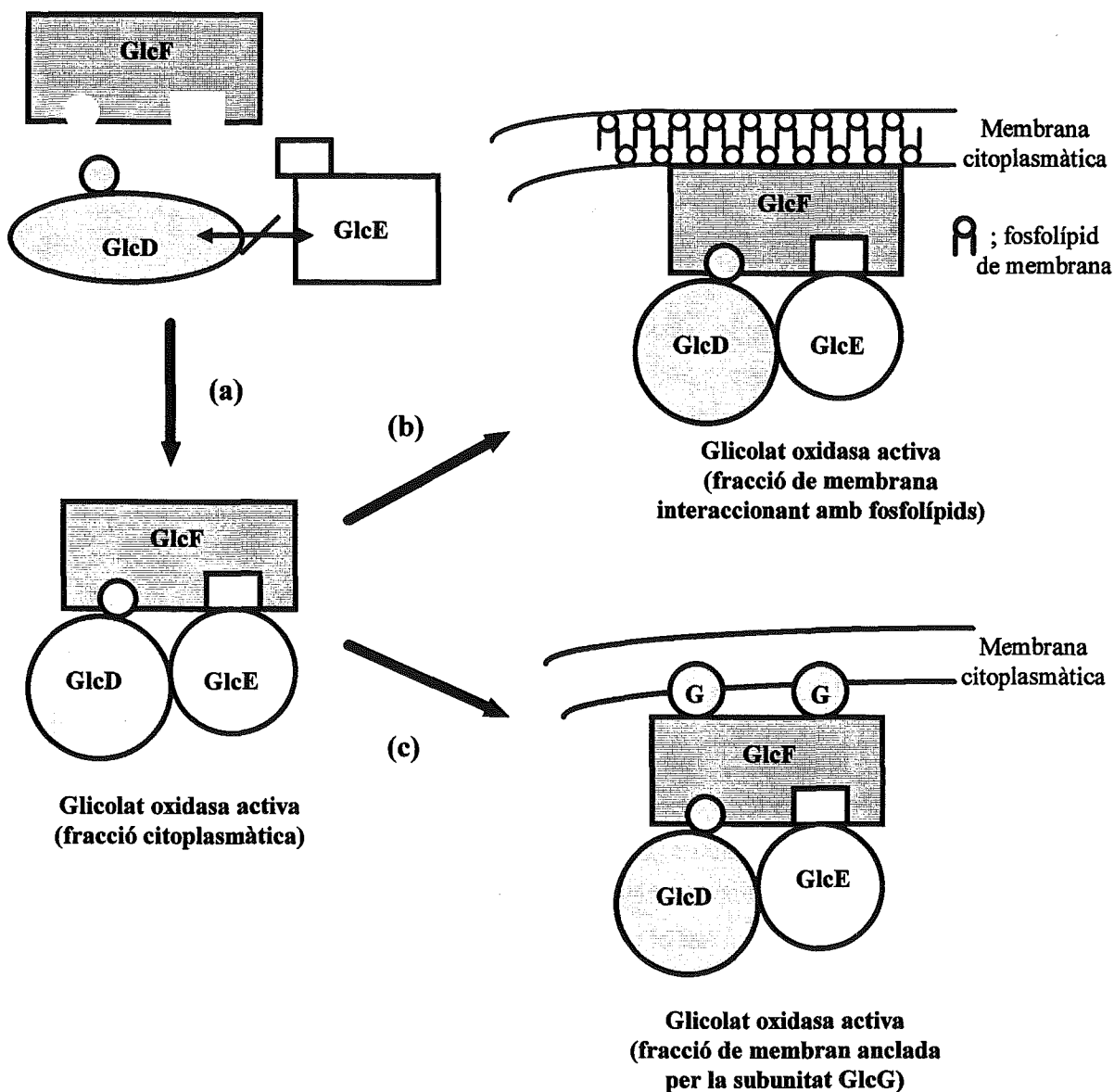


Figura 1. Possibles models estructurals de l'enzim heteroligomèric glicolat oxidasa. Es postula que les subunitats GlcD i GlcE no són estructuralment associables; només quan estan unides a la subunitat GlcF tenen una estructura terciària activa (a). El trímer GlcDEF està associat a la membrana citoplasmàtica per interaccions amb els fosfolípids (b) o bé per l'anclatge a través de les subunitats GlcG (G) (c).

És possible que aquest petit pèptid es posi en contacte a l'interior de la membrana amb el component de la cadena de transport electrònic al qual transfereixi els electrons. Tot i que es desconeix quin és aquest component, podria tractar-se de la ubiquinona, component al qual la

D-lactat deshidrogenasa i altres flavoproteïnes traspassen els seus electrons. És probable que la participació d'una subunitat d'aquest tipus permeti un major rendiment de l'enzim, ja que els electrons despresos en l'oxidació de glicolat a glioxilat passarien directament a la membrana, on es troben localitzats els components de la cadena respiratòria, i així s'evitaria la pèrdua d'energia provocada pels successius salts electrònics. Per comprovar aquesta hipòtesi es podrien realitzar experiments destinats a comprovar si els nivells d'activitat glicolat oxidasa per mol de glicolat consumit varien en mutants de GlcG.

El gen *glcT*, situat en direcció 3' al gen de la malat sintasa G, podria codificar la proteïna transportadora de glicolat. GlcT té un perfil hidropàtic típic de proteïnes de membrana. Quan es compara la seva seqüència aminoacídica amb bancs de dades de seqüències proteïques s'observa que és semblant a diferents proteïnes transportadores d'organismes diversos; la similitud més important la presenta amb l'L-lactat permeasa (LctP) d'*Escherichia coli*. L'apilament de les seqüències dels dos polipèptids mostra que entre ells mantenen una identitat mitja del 65.5 % i una conservació del 80 % dels residus aminoacídics. Existeix però una zona situada a la regió central (des de la posició 271 a 302 de GlcT i de la 271 a la 294 a LctP) que no té similitud. Aquesta regió podria ser la que donés l'especificitat pel substrat transportat.

La implicació de GlcT en el transport de glicolat sembla estar parcialment demostrada ja que mutants de *glcT* tenen un creixement en glicolat enlentit respecte la soca salvatge. El fet que el fenotip no sigui completament negatiu indica que possiblement el glicolat pot entrar dins el bacteri usant altres transportadors alternatius, que estiguin especialment dissenyats per transportar molècules amb estructures similars. Aquest seria el cas del transportador de l'L-lactat, del D-lactat o de l'acetat. Tant el glicolat com el D-lactat, l'L-lactat i l'acetat són àcids monocarboxílics de tres o dos àtoms de carboni. Seria possible que, davant la pressió ambiental que un medi mínim on només existeix glicolat exerceix sobre el bacteri, aquest aconseguís créixer utilitzant algun altre transportador del qual ell pugui ser substrat. En aquest sentit, s'ha descrit l'existència d'una família de transportadors d'àcids monocarboxílics (família MCT) en mamífers implicada en el transport d'acetat, piruvat o lactat. En aquest cas, l'existència de diferents membres d'aquesta família sembla respondre a requeriments específics de teixit (Price *et al*, 1998).

L'assignació definitiva a GlcT del rol de transportador de glicolat requerirà assajos de transport amb glicolat marcat radioactivament per poder veure la incorporació del mateix a l'interior cel·lular. Per comprovar si aquesta proteïna és capaç de transportar altres compostos estructuralment relacionats, s'hauria de comprovar la competició de diferents substrats pel mateix transportador.