



**Universitat Autònoma de Barcelona**

**Departament de Genètica i de Microbiologia**

**VARIABILITAT DE LA XARXA LexA  
EN ELS BACTERIS**

Jordi Cuñé Castellana

2006

**Universitat Autònoma de Barcelona**

Departament de Genètica i de Microbiologia

# VARIABILITAT DE LA XARXA LEXA EN ELS BACTERIS

Memòria redactada per Jordi  
Cuñé i Castellana per optar al  
Grau de Doctor en  
Microbiologia per la  
Universitat Autònoma de  
Barcelona.

Vist-i-plau  
El Director de tesi

Dr. Jordi Barbé i García

**BELLATERRA 2006**

*“La frase més excitant que es pot sentir en ciència,  
la premonitòria de nous descobriments, no és  
“¡Eureka!”, sinó “És estrany...”*

Isaac Asimov

*“La vida és el que et va passant mentre fas plans  
per una altra cosa”*

Sam Shepad

“El Gran Somni del Paradís”

A totes les persones importants en la  
meva vida.

## RESUM

El sistema SOS és una xarxa multigènica sota el control negatiu del repressor LexA, amb importància vital per a la supervivència cel·lular ja que la seva desrepressió es dona quan la cèl·lula presenta danys en el material genètic que poden ser transmesos a la descendència. Aquest fet, junt amb la seva ampla distribució en el domini *Bacteria* ens presenten el coneixement del seu funcionament en els diferents grups com una eina filogenètica idònia.

Per a tal fi, i com a fil conductor d'aquest treball, es va procedir a l'estudi del reguló LexA a *Dehalococcoides ethenogenes*, un bacteri verd no sulfurós; *Magnetococcus sp.* soca MC-1, un bacteri magnetotàctic; i *Leptospira interrogans* serovar Lai, la primera espiroqueta amb un *lexA* identificat. Per a tal es va procedir a la clonació i purificació del gen *lexA* de cada un d'ells i a la purificació del seu producte.

En el primer d'ells es va definir la caixa d'unió de LexA (AGAACN<sub>4</sub>GTTCT), comprovant a partir d'ella els seus vincles amb els bacteris grampositius, així com la no regulació del gen *recA*, considerat com a canònic, pel mateix.

A *Magnetococcus sp.* soca MC-1, la caixa SOS descrita fou GTTCN<sub>7</sub>GTTTC. Fins el moment, aquest microorganisme es trobava ubicat dins la subclasse Alfa del Proteobacteris, però tant el coneixement del motiu d'unió del seu LexA com els gens que integraven el reguló, ens el mostren com una branca estretament relacionada amb aquesta classe, però clarament independent.

A *Leptospira interrogans* serovar Lai, el resultat fou sorprenent, ja que la caixa SOS d'aquest microorganisme, la varem haver d'identificar a partir de la regió promotora de *recA*, ja que no era present en la de *lexA*, esdevenint així el primer exemple en el que LexA no autoregula la seva expressió. Aquesta dada així com el fet de no poder detectar cap altre gen sota el seu control, ens mostren *Leptospira*, com un pas entremig en la tendència evolutiva seguida en les espiroquetes de perdre el seu gen *lexA*.

Els següents articles es presenten a la secció d'annexes, i en ells es descriuen i discuteixen els resultats sobre els quals està basada aquesta tesis.:

- I. **Fernandez de Henestrosa, A.R., J. Cuñé, I. Erill, J.K. Magnuson, i J. Barbé.** 2002. A green nonsulfur bacterium, *Dehalococcoides ethenogenes*, with the LexA binding sequence found in gram-positive organisms. *J. Bacteriol.* **184**(21): 6073 - 6080.
  
- II. **Fernandez de Henestrosa, A.R., J. Cuñé, G. Mazón, B.L. Dubbels, D.A. Bazylnski, i J. Barbé.** 2003. Characterization of a new LexA binding motif in the marine magnetotactic bacterium strain MC-1. *J. Bacteriol.* **185**: 4471 - 4482.
  
- III. **Cuñé, J., P. Cullen, G. Mazon, S. Campoy, B. Adler, i J. Barbé.** 2005. The *Leptospira interrogans* *lexA* gene is not autoregulated. *J. Bacteriol.* **187**: 5841 - 5845.

## **1. INTRODUCCIÓ**

# 1. INTRODUCCIÓ

Les poblacions, i cada un dels individus que les integren, poden ser considerats com a vectors de transmissió del fons genètic comú de cada una de les espècies a les que pertanyen. Aquest fet, no implica que la projecció en el temps de les espècies hagi estat lineal i invariable, sinó com podem observar en l'actual diversitat present en la biosfera, aquestes han divergit i s'han diferenciat de manera extraordinària. La maquinària que ha originat aquesta variabilitat, amb força motriu la diversitat ambiental i els errors imprescindibles en la transmissió del material hereditari, és el que coneixem com a evolució.

Òbviament, tot i l'elevat grau de divergència i conseqüentment el gran nombre de diferències, totes i cada una de les formes de vida actuals, tenen per llunyà que sigui, un ancestre comú. Aquest fet, permet l'intent d'un rastreig a través de diferents indicadors, de quins han estat els camins evolutius i les relacions de filiació entre individus de diferents espècies, essent la utopia d'aquesta branca de coneixement, l'elucidació de l'actualment anomenat Luca: l'ancestre universal comú últim (sigles de la seva escriptura en anglès *Latest universal common ancestor*). Els indicadors escollits, cal que siguin universals en el nostre grup d'estudi i que desenvolupin un paper cel·lular transcendent, minimitzant així possibles artefactes deguts a la deriva gènica. Aquestes premisses ens condueixen a estimar com a eines més idònies pel nostre estudi, elements moleculars implicats en funcions cel·lulars bàsiques, com per exemple, i parlant del món dels bacteris, el sistema SOS.

La gènesis de tot el que entenem com a vida es troba en la informació codificada en el DNA, per tant, la perpetuació de l'espècie, dependrà en gran mesura de la transmissió a la descendència del material genètic el més inalterat possible. També és cert, que la presència de certa variabilitat, repercutirà positivament en el futur de l'espècie.

Aquestes dues afirmacions, tot i que antagòniques, cal considerar-les com a complementàries: la variabilitat gènica proporciona recursos a l'espècie davant determinades exigències ecològiques, però si els errors que causen aquesta variabilitat, es concentren en el temps o són de gran importància, tenen un efecte deleteri, tot i que de manera individual i amb intervals suficients no ho siguin.

Per a tal fi, tant les cèl·lules eucariotes com procariotes disposen de diferents mecanismes de reparació dels danys que per diferents motius s'originen en el DNA. Aquests danys, són deguts a l'exposició del material genètic a diferents agents nocius, que poden tenir el seu origen en el mateix metabolisme, o bé, ser d'origen exogen. Segons la naturalesa i la gravetat dels danys, la cèl·lula usarà un o altre dels diferents mecanismes dels que disposa per a remeiar-los, de manera total, parcial, definitiva o transitòria.



Podem classificar els diferents mecanismes presents en les cèl·lules procariotes segons el tipus d'actuació. Així podem parlar de reparació per reversió directa, escissió i recombinació. Dins de la reversió directa, podem parlar de la fotoreactivació i de la transferència d'alquils. Aquests sistemes, tenen la capacitat de detectar la presència de la lesió i al mateix temps reparar-la, a través d'enzims que catalitzen la reacció, com són les fotoliasas en el primer cas i les alquil transferases en el segon.

En la reparació per escissió, es produeix una eliminació de la lesió, bé de forma molt precisa (escissió de base) o de forma més àmplia (escissió de nucleòtid). En ambdós casos, hi intervenen una sèrie d'enzims amb diferents activitats a fi d'acomplir la missió.

La reparació per recombinació es dona quan les lesions introduïdes presenten un major grau de dificultat de reparació: enllaços entre les hebres, trencaments uni- o bicatenaris o alteracions en la cadena motlle durant la replicació. Les vies d'actuació d'aquest sistema són diverses, però en totes elles intervé la proteïna RecA en el pas d'aparellament de cadenes o sinapsi.

El microorganisme més estudiat, i per tant també en el que més bé es coneixen aquests mecanismes de reparació és *Escherichia coli*. En estudis recents, s'ha detectat l'expressió de més de mil gens davant d'exposició d'un agent lesionant com la mitomicina C (Khil i Camerini-Otero, 2000). La majoria d'aquests 1.000 gens es troben formant part de xarxes multigèniques capaces d'activar-se selectivament segons el tipus de dany i l'abast del mateix. Les més estudiades a *E. coli* fins el moment, són:

- i) **Resposta adaptativa a agents alquilants.** Aquest mecanisme s'activa quan la cèl·lula es sotmesa a un estrès degut a l'exposició a elevades concentracions d'espècies químiques alquilants del DNA tals com la N-metil-N'-N-nitrosoguanidina (MNNG), el metil metanosulfonat (MMS) o l'etil metanosulfonat (EMS) entre d'altres. El reguló està format pels gens *ada*, *alkA*, *alkB* i *aidB*, essent el producte del primer, la proteïna Ada, el regulador positiu al mateix temps que implicat en les tasques de reparació (Landini i Volkert, 2000).
- ii) **Resposta a l'estrès oxidatiu.** S'entén com a estrès oxidatiu la presència de compostos químics coneguts com a ERO: espècies reactives de l'oxigen. Aquestes poden ser d'origen endogen, o provenir del medi. Entre elles podem citar el peròxid d'hidrogen ( $H_2O_2$ ), radicals superòxid ( $O_2^-$ ) o el nitril (NO). Els regulons implicats descrits fins el moment són l'OxyR i el SoxR, que agrupen més de 20 gens, i que són els responsables d'actuar davant les lesions que es deriven d'aquests agents. L'OxyR, codifica catalases i reductases que actuaran en presència de peròxid d'hidrogen. El SoxR per la seva banda, regula l'expressió de diferents enzims,

destacant entre ells el de la superòxid dismutasa (Demple, 1997a; 1997b; Michán *et al.*, 1999; Zheng *et al.*, 1999).

- iii) **La resposta SOS o sistema de reparació d'emergència.** En ser aquest sistema l'eix central de la present tesi, a continuació es passarà a exposar el seu funcionament de manera més extensa.

### **1.1. El sistema SOS a *Escherichia coli***

Al 1975, Miroslav Radman (Radman, 1975) publicava “*Una hipòtesi va ser formulada fa uns anys respecte si E. coli posseeix un sistema induïble de reparació del DNA (reparació SOS) que és el responsable també de la mutagènesis induïda. Algunes característiques del sistema SOS són (1) és induït o activat de manera posterior a l'aparició de danys en el DNA, (2) requereix la síntesi de proteïnes de novo, (3) consta de diferents gens dels quals els més estudiats fins el moment són recA i lexA, i (4) els requeriments fisiològics i gènics per l'actuació de la reparació SOS són substancialment similars als necessaris per la inducció de profags...*”.

Avui en dia, coneixem més i més profundament el sistema SOS, però les bases canòniques del seu funcionament són les anteriorment citades.

Quan Radman feia referència als requeriments gènics necessaris i més estudiats, ja s'apuntava clarament la dualitat de regulació present en el sistema SOS, tot i algunes excepcions: el producte del gen *lexA* n'és el repressor i el del *recA*, el promotor de la inducció.

Actualment, també es coneixen la resta de requeriments per a un correcte funcionament del sistema: una seqüència específica d'unió per la proteïna LexA al DNA en els zones pròximes dels gens o operons sota el seu control, que com a mínim a *E. coli* són 40 (Fernandez de Henestrosa *et al.*, 2000). Així mateix, es coneix que aquests varien entre els diferents grups filogenètics, degut a l'ampli ventall de les necessitats de reparació del DNA i del control del cicle cel·lular, derivades de la varietat de nínxols ecològics possibles (Mazón *et al.*, 2004).

#### **1.1.1. Inducció del sistema SOS.**

El senyal que inicia el seguit de fets que conduiran a la inducció del sistema SOS, és la presència de danys en el DNA, específicament, l'aparició de regions de cadena senzilla (ssDNA), que s'originen quan el complex de replicació del DNA, i en concret la DNA

polimerasa III, intenta avançar en la seva tasca en un punt del genoma on s'ha produït una lesió (Sassanfar i Roberts, 1990). La presència de zones amb ssDNA, es veu incrementada degut a l'activitat del complex RecBCD (holoenzim exonucleasa V), i més exactament, a la seva activitat helicasa sobre el dsDNA (Churcill *et al.*, 1999).

El receptor de la senyal inductora (la presència de regions d'ssDNA) és la proteïna RecA, la qual pren una conformació activa formant nucleofilaments al voltat de les monohebres de DNA, amb consum d'energia en forma d'ATP (Yu i Egelman, 1993). Estudis estèrics, han provat que la seva unió es produeix per agrupació de 6 monòmers per volta de DNA (Egelman, 1998) de manera cooperativa (Sattin *et al.*, 2004). En el procés de nucleació, existeixen altres elements que intervenen en la formació dels nucleofilaments, com ara les proteïnes Ssb (que faciliten l'accés a les zones d'ssDNA a la proteïna RecA) i el sistema RecFOR (que fa que el procés en global sigui molt més eficient) (Kuzminov *et al.*, 1999; Rangarajan *et al.*, 2002).

Mentre el DNA no presenta danys, trobem el sistema SOS reprimint mitjançant la unió de la proteïna LexA (en forma de dímer) a la seqüència específica d'unió, o caixa SOS, present en cada un dels gens del reguló, *lexA* inclòs, impeding una taxa de transcripció alta, ja que sempre trobarem una expressió basal en tots i cada un dels gens.

La interacció dels polímers nucleoproteïcs del RecA amb el LexA, facilita l'autohidròlisi d'aquest últim (figura 1.1), que es produeix entre els residus Ala<sub>84</sub> i Gly<sub>85</sub>. Els fragments resultants d'aquesta autohidròlisi són ràpidament destruïts, en quedar exposades regions reconegudes per la proteasa ClpXP, que fins llavors quedaven amagades (Neher *et al.*, 2003). D'aquesta manera, els gens del reguló, els promotors dels quals estaven reprimits, augmenten la seva taxa de transcripció, i així serà fins que les lesions en forma d'ssDNA desapareguin, amb el que la proteïna RecA tornarà a la seva conformació inactiva, deixant llavors de facilitar l'autohidròlisi de la proteïna LexA sintetitzada *de novo*, permetent acumular-ne de nou un quantitat suficient per a reprimir un altre cop l'expressió dels gens del reguló, retornant tots ells, al nivell d'expressió basal (figura 1.1).

Estudis recents han assenyalat la presència de moduladors de l'activitat del RecA, pertanyents al reguló *lexA*: el producte dels gens *dinI* i *recX*. En el primer d'ells, la proteïna DinI, la seva participació fou estudiada per mitjà de l'estudi de la seva capacitat d'hidrolitzar ATP (Lusetti *et al.*, 2004a). En el principi de la resposta SOS, quan es troba en baixes concentracions, contribuiria a estabilitzar la nucleació del RecA, facilitant així la resposta. Posteriorment, al final de la resposta SOS, i quan es troba en elevades concentracions, juga exactament el paper contrari. El RecX per la seva part, tindria la funció de dificultar la capacitat coproteàsica. Per tant, ambdós es poden considerar com a moduladors de l'activitat del RecA (Drees *et al.*, 2004; Lusetti *et al.*, 2004b).

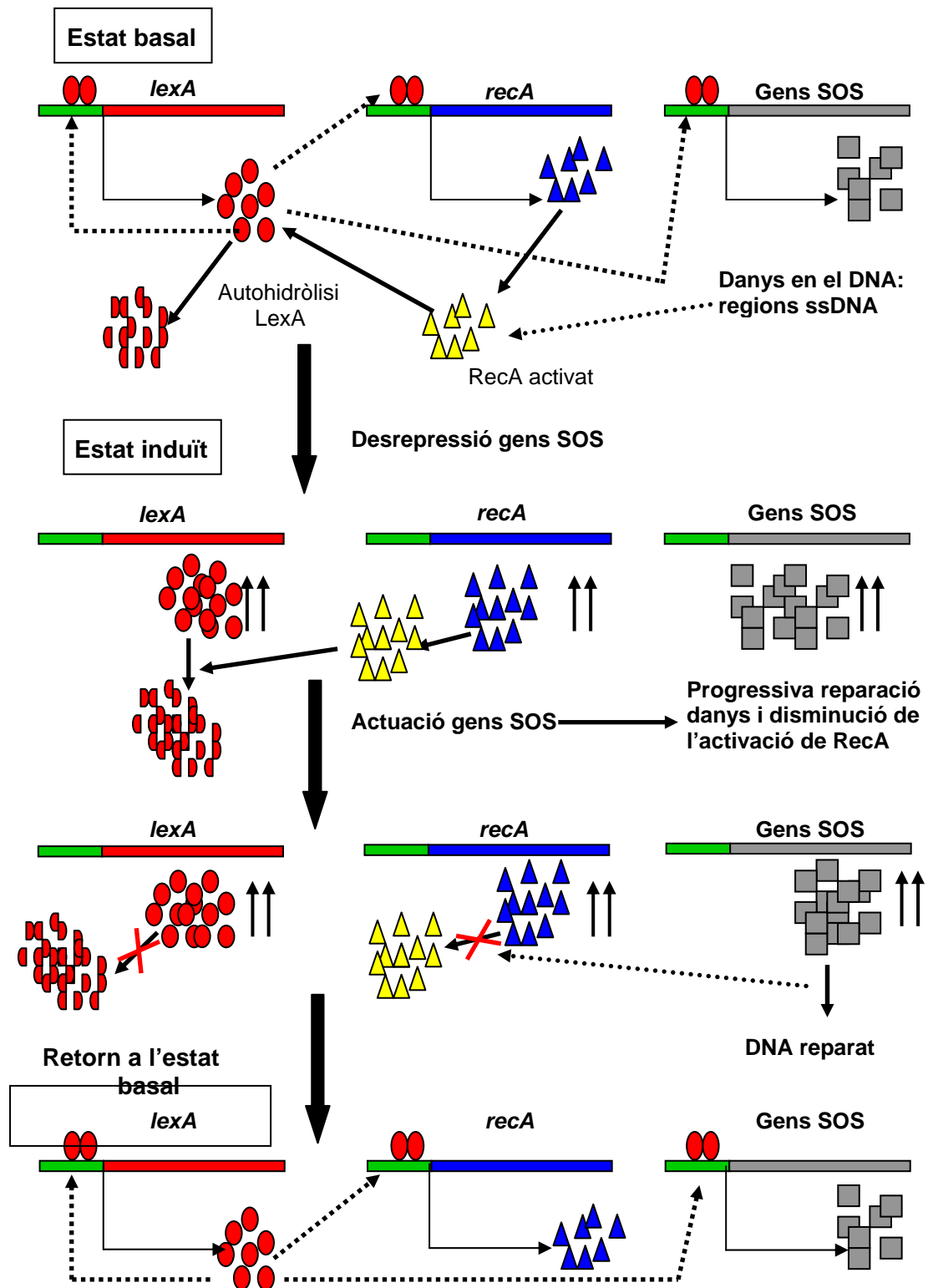


Figura 1.1. Representació esquemàtica del funcionament del sistema SOS.

### 1.1.1.1. La proteïna LexA

La proteïna LexA d'*E. coli*, està formada per 202 aminoàcids i té un pes molecular de 22,7 kDa (Horii *et al.*, 1981). La seva regió codificant es troba formant unitat transcripcional amb el gen *dinF* a *E. coli* (Heide *et al.*, 1993).

La seva funció en el sistema SOS, és la repressió dels gens del reguló, inclosos el *recA* i el mateix *lexA* (Brent i Ptashne, 1981; Little *et al.*, 1981), mitjançant la unió a una seqüència nucleotídica específica o caixa SOS: 5'-CTGT(AT)<sub>4</sub>ACAG-3' (Walker, 1984; Wertman i Mount, 1985). Per mitjà de la unió, la proteïna impedeix el correcte accés de la RNA polimerasa al promotor dels gens, impedit-ne una excessiva transcripció.

En el mecanisme d'inhibició de la transcripció, juga també un paper important la capacitat de dimerització que presenta LexA, reconeixent cadascun dels monòmers un dels motius inversament simètrics de la caixa SOS, gràcies a la seva estructura palindròmica.

Pel seu funcionament com a regulador, junt amb la capacitat d'unió al DNA i dimerització que el permeten actuar com a repressor, també és important la capacitat d'autohidròlisi. Aquest procés es dona *in vivo* degut a un pH bàsic o a la presència de danys en el DNA, que indueixen la proteïna RecA, a actuar com a catalitzadora del procés.

#### i) Estructura

Pel que fa a l'estructura i funció podem diferenciar tres regions en la proteïna LexA:

a) Domini N-terminal: podem confinar-lo entre els aminoàcids 1 i 72 de la proteïna. Aquest domini és l'implicat en la unió de la proteïna LexA al DNA, per mitjà d'una estructura molt semblant al motiu HTH (*helix-turn-helix*), comú en moltes proteïnes que requereixen de la unió al DNA, com són els repressors transcripcionals (Harrison i Aggarwal, 1990; Schnarr *et al.*, 1991).

El coneixement tridimensional de la proteïna LexA, va arribar gràcies a estudis de ressonància magnètica nuclear (RMN). Aquests mostraren la presència de 3 estructures hèlix  $\alpha$  entre els nucleòtids 6-21, 28-35 i 40-52, batejades com a 1, 2 i 3. Aquestes es troben connectades per regions *turn* i també trobem 2 regions  $\beta$  antiparal·leles que comprenen els residus 50-58 i 66-68 (Fogh *et al.*, 1994) (figura 1.2).

De les tres hèlix  $\alpha$ , són les dues últimes les més importants en el reconeixement de les seqüències d'unió. De fet, sembla que és l'hèlix  $\alpha_3$  la que presenta una major importància en l'especificitat, i per tant, els seus residus aminoacídics variaran segons el

microorganisme, en funció de la seqüència nucleotídica de la seva caixa SOS (Thliveris i Mount, 1992; Groban *et al.*, 2005).

b) Regió connectora (*hinge region*): té la funció d'enllaçar la regió amino i carboxi terminal. Comprèn els aminoàcids situats entre els residus 73 i 94, i per tant, engloba l'enllaç Ala<sub>84</sub>-Gly<sub>85</sub>, per on es produeix el trencament en l'autohidròlisi. Presenta certa torsió, necessària per a facilitar el trencament de LexA, facilitant l'entrada en contacte amb el domini C-terminal (Little *et al.*, 1984) (figura 1.2).

c) Domini carboxi terminal: s'extén del residu 94 al 202. La seva estructura es bàsicament en forma de fulles  $\beta$  (Hurstel *et al.*, 1986). És la part més conservada de la proteïna LexA, amb una gran similitud amb la proteïna UmuD i amb alguns repressors de bacteriòfags (Battista *et al.*, 1990a), tot i que respecte aquests últims, divergeix en el mecanisme de trencament (Mustard i Little, 2000). Per tant, vèiem que en el procés de regulació és important la interacció amb la proteïna RecA activada i l'autohidròlisi (figura 1.2).

## **ii) Unió a les caixes SOS**

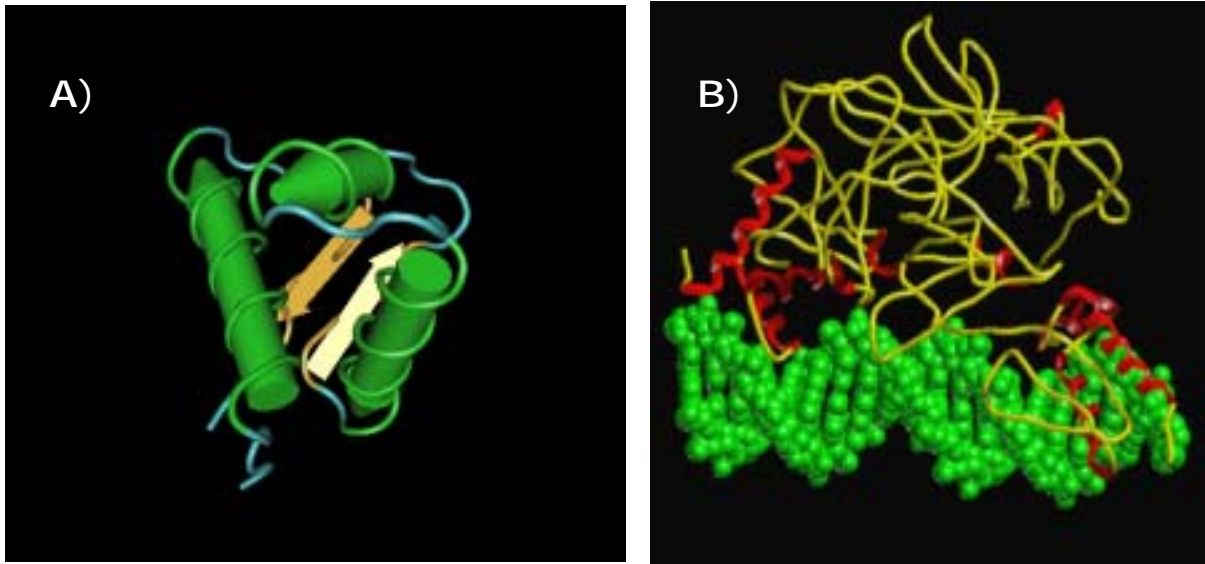
La unió de la proteïna al DNA, es produeix gràcies al seu domini N-terminal.

Com qualssevol reacció química reversible, és un procés dinàmic. Experimentant, s'ha pogut conèixer com transcórrer el procés de dimerització (veure dimerització, pàg 9):

Com s'ha citat anteriorment, les dues últimes hèlix  $\alpha$  són les que tenen un paper més rellevant en la interacció amb el DNA. Estudis basats en mutagènesis de proteïnes, han provat quins són els residus aminoacídics més importants. Dins de la tercera hèlix (40-52), destaquen els residus situats entre el 42 i el 47 com a primordials en la unió a la regió operadora (Thliveris *et al.*, 1992), ja que mutacions d'aquests, inhibeixen la capacitat d'unió però no la dimerització amb LexA salvatges. Variants d'aquest estudi provaren la capacitat de modificar la seqüència reconeguda per part de la proteïna, reforçant la hipòtesi de que és en aquesta reduïda zona on resideix l'especificitat d'unió (Groban *et al.*, 2005).

Estudis de ressonància magnètica nuclear (RMN), han aportat llum respecte quines són les interaccions que s'estableixen entre la seqüència 5'-CTGT(AT)<sub>4</sub>ACAG-3' i la proteïna LexA (Oertel Buchheit *et al.*, 1990; Oertel Buchheit *et al.*, 1992). La naturalesa d'aquestes relacions bioquímiques es variada, destacant els ponts d'hidrogen entre els residus N<sub>41</sub>, E<sub>44</sub> i E<sub>45</sub> amb les bases CTGT i les interaccions hidrofòbiques entre la primera timina de la caixa

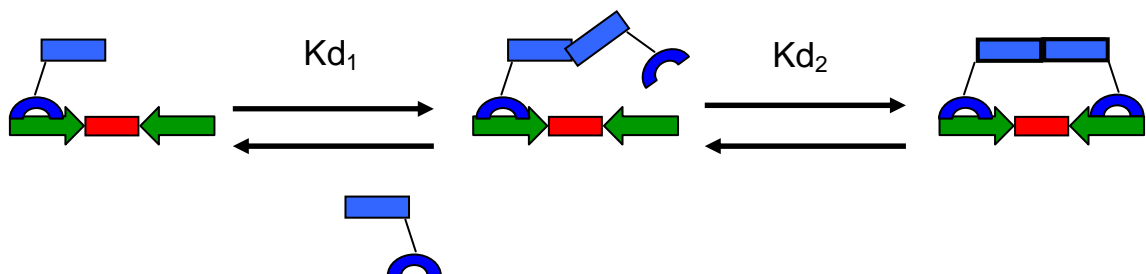
amb els residus S<sub>39</sub>, A<sub>42</sub> i N<sub>41</sub>. A més d'aquestes, cal esmentar les relacions que existeixen entre diferents residus aminoacídics i l'esquelet fosfat, junt amb una gamma d'atraccions i repulsions hidrofòbiques i hidrofíliques que garanteixen l'estabilitat de la unió (figura 1.2).



**Figura 1.2.** a) Modelatge tridimensional de la proteïna LexA obtingut amb el programa Cn3D 4.1 disponible en el National Center of Biotechnology Information (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Domains>). b) Model de la interacció de la proteïna LexA amb el DNA.

### iii) Dimerització

Per mitjà de diferents estudis s'ha pogut provar que és un procés seqüencial, on el primer pas és la unió d'un monòmer de LexA a una meitat del motiu (figura 1.3), seguit de la unió d'un altre monòmer a la segona part del operador. Aquesta unió s'estabilitza per mitjà de la interacció dels dominis carboxi terminal (Kim i Little, 1992; Mohana-Borges *et al.*, 2000).



**Figura 1.3.** Model de la dimerització de la proteïna LexA en la seva unió a la caixa SOS. El domini N-terminal està representat per un semicercle i el C-terminal per un rectangle. La zona connectora, o *hinge region*, es simbolitza per mitjà d'una línia. Modificat de Kim i Little (1992).

#### iv) Autohidròlisi

El fet que el fenomen sigui denominat com a autohidròlisi, tot i que el paper del RecA en la dualitat de regulació sigui el d'induir aquest trencament, es degut a que tant el punt de trencament com el centre catalític del procés es troben en la proteïna LexA. El primer el trobaríem en l'enllaç Ala<sub>84</sub>-Gly<sub>85</sub> i el segon en els residus Ser<sub>119</sub> i Lys<sub>156</sub> (figura 1.4). Aquesta reacció també s'ha observat que es pot donar de manera espontània a pH bàsic (Little, 1984; Little *et al.*, 1991).

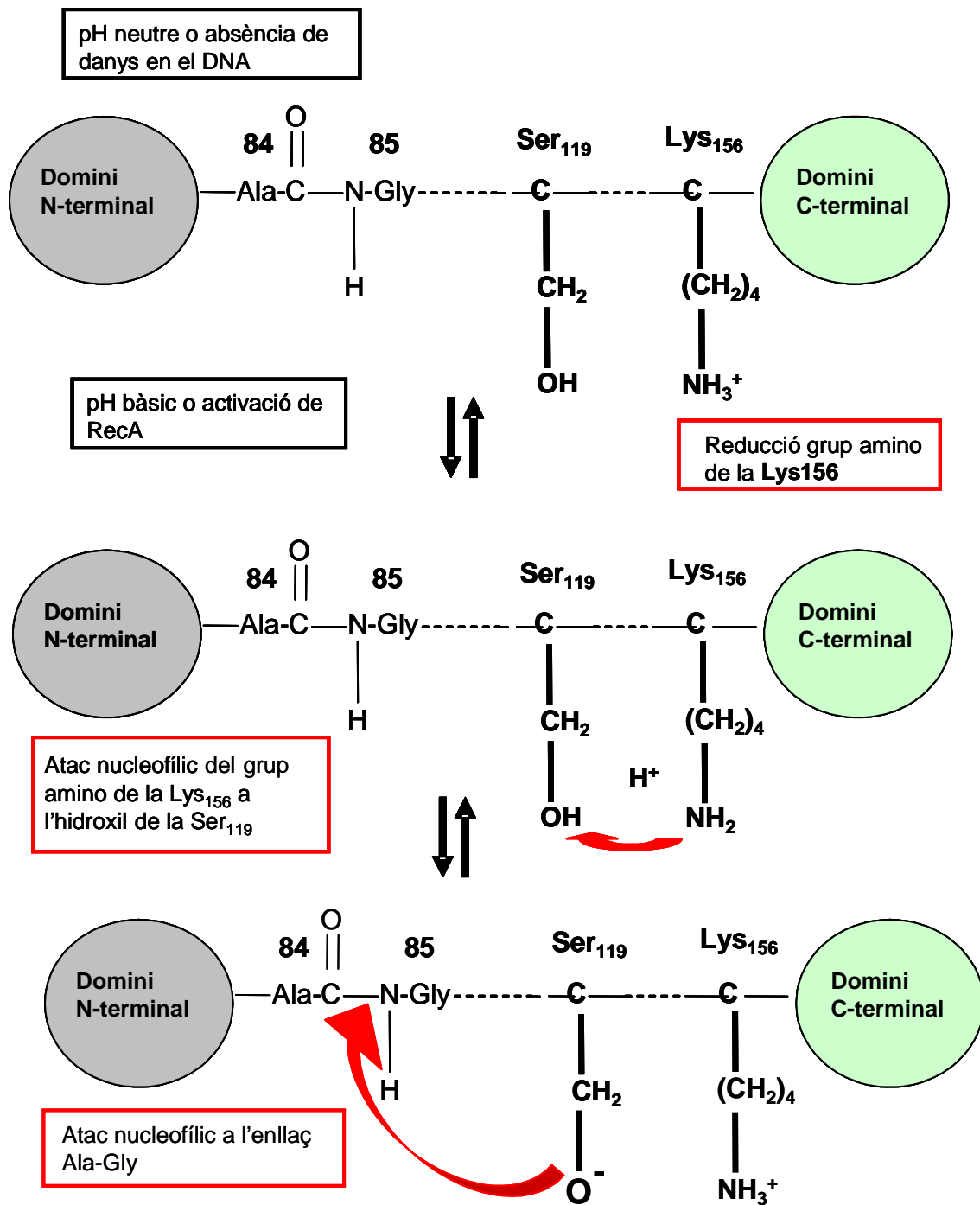
La naturalesa bioquímica del trencament, consisteix en una desprotonització inicial del grup amino de la Lys<sub>156</sub>, que condueix a un atac nucleofílic del mateix grup químic al radical hidroxil de la Ser<sub>119</sub>, que al seu temps, neutralitza la seva reducció amb un nou atac, de la mateixa afinitat que l'anterior, sobre l'enllaç Ala<sub>84</sub>-Gly<sub>85</sub>, produint el trencament de la proteïna per la part connectora (*hinge region*).

La ubicació d'aquests residus en les proteïnes LexA de diferents espècies bacterianes, es manté constant al llarg de l'evolució, amb una variació en la posició d'un o dos residus, fet que ha estat emprat per a la localització d'aquesta proteïna en els diferents genomes bacterians dels que es disposa fins el moment. Aquest fet no ha d'estranyar, ja que com s'ha esmentat anteriorment, és la regió N-terminal la que presenta una major variabilitat interespecífica, mentre que la C-terminal (incloent-hi la zona corresponent a la *hinge region*) és la menys variable.

Mecanismes molt semblants a l'anterior s'han descrit en serin-proteases, amb les quals mostra elevada homologia la regió C-terminal de LexA (Roland *et al.*, 1992). Tot i així, dades obtingudes en treballs posteriors, apunten a un mecanisme més semblant al desenvolupat per les  $\beta$ -lactamases (Little, 1993).

Sembla doncs, que la proteïna LexA, podria presentar dues conformacions: una seria compatible amb els fets esmentats anteriorment, mentre que l'altra no. El paper que tindria la proteïna RecA activada, seria el d'estabilitzar la conformació que facilita el trencament (Oertel-Buchheit *et al.*, 1998; Luo *et al.*, 2001).





**Figura 1.4.** Representació esquemàtica de l'autohidrolisis de la proteïna LexA. S'han omittit les característiques estructurals globals per a representar tant sols els residus implicats en el mecanisme de trencament.

### 1.1.1.2. La proteïna RecA

La proteïna RecA d'*E. coli* està formada per 352 aminoàcids, té un pes molecular de 38 kDa, un pI de 5,6 (Cox, 2003) i presenta polifuncionalitat en diferents processos cel·lulars. És una proteïna molt conservada al llarg de l'evolució, motiu pel qual ha estat utilitzada com a marcador filogenètic molecular (Eisen *et al.*, 1995).

#### i) Estructura

L'estructura tridimensional de la proteïna ha pogut ser determinada gràcies a eines de resolució atòmica, permetent detectar diferents dominis, tant a nivell estèric com funcional. Del primer tipus, destacar el domini central (corresponent als residus aminoacídics que ocupen la posició 36 fins la  $266 \pm 4$ ) (figura 1.5), que conté la regió d'unió a nucleòtids i dos regions de baixa densitat electrònica: el *bucle 1* i *bucle 2* implicats en la unió al DNA. Acompanyant aquest domini, en trobem dos més de grandària menor. Un a l'extrem N-terminal ( $1-36 \pm 8$ ) i l'altre a l'extrem C-terminal (266-352) (Story *et al.*, 1992; Rehrauer i Kowalczykowski, 1996) (figura 1.5).

A nivell funcional, podem citar els que es mostren a continuació, correlacionats amb les tasques que duen a terme cada un d'ells:

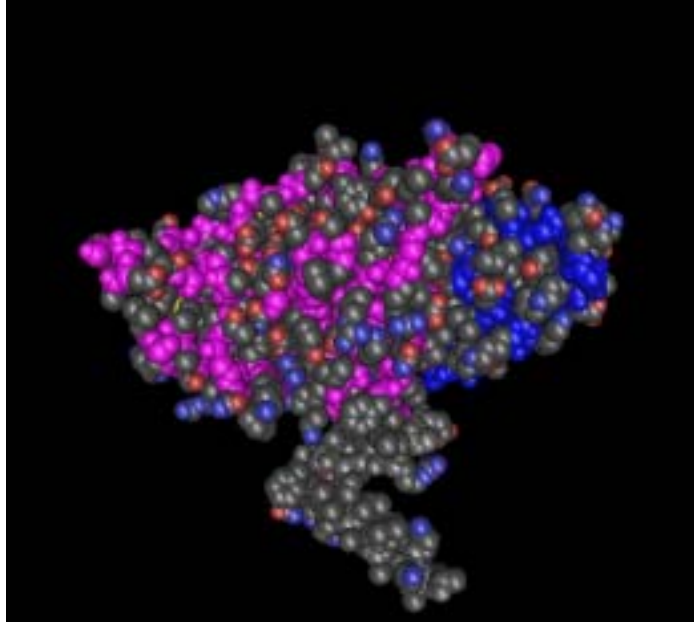
**a) Dominis d'unió a nucleòtids i hidròlisi.** Són dos dominis altament conservats que es troben en la regió d'unió A, compresa entre els residus 66 i 73 (GPESGKT), i la regió B, delimitada pels residus 140-144. La regió A s'uneix a ribonucleòsids trifosfat (figura. 1.5).

**b) Dominis d'unió al DNA.** Els *bucles 1* i *2* anteriorment esmentats, sembla que serien els implicats en aquesta funció, degut a la seva ubicació central en la proteïna i a la formació de plegaments (figura 1.5). Tot i així, sembla que podrien existir altres residus implicats en aquestes funcions d'unió al DNA. Entre aquests trobaríem els que ocupen la posició 61 fins la 72, de la 178 a la 183 i de la 233 a la 243 (Rehrauer i Kowalczykowski, 1996).

**c) Dominis d'interacció entre unitats monomèriques i entre filaments.**

Com ja s'ha exposat, la unitat funcional de la proteïna RecA és un oligòmer, fruit de la interacció entre 6 monòmers per volta de l'hèlix de DNA, gràcies a interaccions hidrofòbiques al llarg de tota la proteïna, del que en deriva la dificultat d'especificar la posició de tots i cada un dels punts d'unió (figura 1.5). Pel que fa a la interacció entre els filaments, les zones de major importància es troben situades en els extrems de cada monòmer, és a dir, en les seves regions N i C-terminal (37-38 i 298-301, respectivament).

**d) Dominis d'unió a LexA i altres proteïnes.** Els residus implicats en aquestes unions, són els mateixos si la proteïna és un LexA, un UmuD o es tracta de repressors fàgics. La zona implicada va dels residus 229 al 243 (Karlin i Brocchieri, 1996), junt amb el *bucle 1* (Yu i Egelman, 1993).



---

**Figura 1.5.** Imatge tridimensional obtinguda amb el programa Cn3D 4.1 disponible en (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Domains>). En color rosa, el domini central on trobem les funcions d'unió a nucleòtids, unió al DNA, unió entre els filaments del RecA i la interacció amb altres proteïnes amb funcions diverses, com per exemple, facilitant la hidròlisi del LexA. Acompanyant aquest domini principal trobem, de color gris, el corresponent al domini N-terminal, i en color blau el C-terminal. En aquest últim també tenen lloc funcions relacionades amb la interacció entre els filaments.

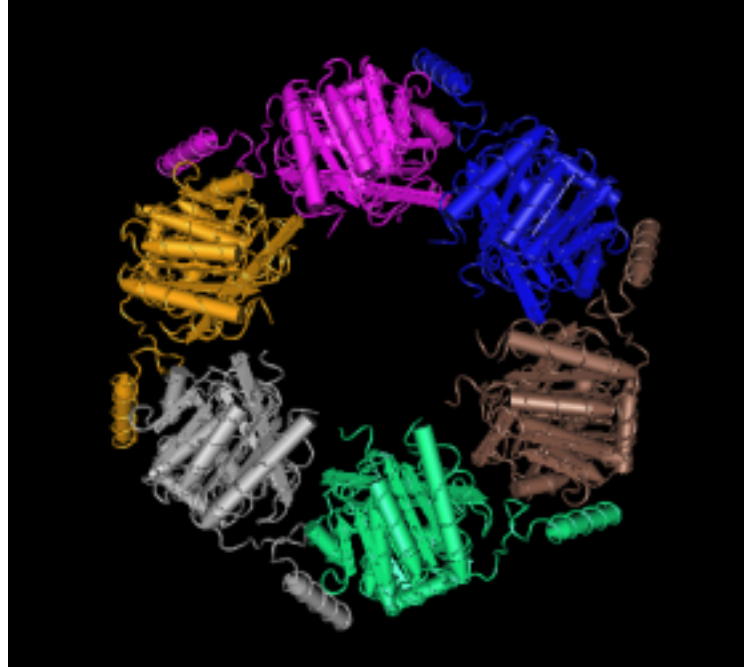
---

## ii) Nucleació

La unitat funcional de la proteïna RecA, és un polímer helicoïdal regular (figura 1.6) format per la unió de sis monòmers del RecA per volta de DNA que ho fan de manera ràpida i cooperativa en direcció 5' (Kojima *et al.*, 1990; Konforti i Davis, 1992; Kim *et al.*, 1993), en estructura molt similar al complex RuvB (George *et al.*, 2000). Aquesta estructura es troba normalment envoltant regions de cadena senzilla del DNA, tot i que en alguns casos també ho fa en el dsDNA. Quan es troba en aquesta associació, podem parlar de complex nucleoproteïc activat.

Estèricament, podem parlar de la presència d'un solc longitudinal llis i un rugós (degut a les prominències dels monòmers). Aquest últim és on es troba el punt d'interacció amb el LexA i també del DNA de doble cadena. D'aquest fet, podem

deduir que existeix competència entre la recombinació (funció desenvolupada per RecA) i l'activitat coproteàsica de la proteïna RecA (Harmon *et al.*, 1996; Rehrauer i Kowalczykowski, 1996)



---

**Figura 1.6.** Representació tridimensional de la nucleació de la proteïna RecA obtinguda amb el programa Cn3D 4.1 disponible en (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Domains>). Cada un dels 6 monòmers està representat per un color, i cal apreciar la interacció que existeix per mitjà d'una hèlix  $\alpha$  entre cada un d'ells.

---

### iii) Funcions

La proteïna RecA, com ja s'ha dit, intervé en nombroses activitats molt importants per la cèl·lula, com són la recombinació homòloga, la regulació dels gens SOS i la mutagènesi SOS (un tipus de reparació del DNA que implica una taxa de mutagènesi molt alta).

Entre les reaccions i activitats que realitza intervenint en totes les funcions anteriorment esmentades, trobem:

- a) Intercanvi i aparellament de bases.
- b) Activitat coproteàsica dependent d'ATP i del DNA.
- c) Interacció amb factors proteics mutagènics que faciliten la reparació tendent a error davant lesions al DNA (Tang *et al.*, 1998).

Derivat de la seva participació en aquests importants processos cel·lulars, el seu nivell basal és alt (entre 1.000 i 10.000 monòmers per cèl·lula), quantitat que, en moments puntuals d'inducció del sistema SOS, pot augmentar entre 20 i 50 vegades.

### 1.1.2. Els gens SOS

El conjunt de gens que formen l'operó *lexA* no és el mateix en tots els microorganismes, variant en relació les necessitats adaptatives de cada un d'ells segons el nínxol ecològic que ocupen. Tot i així, existeixen una sèrie de gens que podríem catalogar com a canònics del sistema, com per exemple *lexA* i *recA*, per la seva importància de regulació, tot i que s'han descrit algunes excepcions a aquesta màxima, que posteriorment es tractaran amb més detall (Narumi *et al.*, 2001; Campoy *et al.*, 2005). A *E. coli*, els gens que es troben sota la regulació negativa de LexA, són els que es mostren en la taula 1.1. En la mateixa taula, també es detallen les funcions atribuïdes a cada un dels productes proteics dels gens dels que se'n coneix la funció.

L'eficiència de la repressió no és invariable, si no modulable segons la distància de la caixa respecte l'inici de transcripció i les variacions en la seqüència respecte la consens, entenent com a consens, la seqüència unitària més pròxima al mínim comú múltiple de la totalitat de les detectades en els gens SOS d'*E. coli*.

Aquest fet encara confereix una major control del grau d'expressió de cada gen i per tant de combinatòria en el funcionament del reguló, poden filar més prim en l'ajustament de les necessitats de cada microorganisme en qüestió, en relació amb les funcions i per tant beneficis que li pot comportar l'expressió justa i adequada d'un gen determinat. Així existeixen treballs on s'ha descrit numèricament la repressió dels diferents gens en tant la major o menor afinitat de LexA per la seva caixa (Lewis *et al.*, 1994). En aquests estudis, es va definir l'índex heteròleg (HI), entès com a indicador de l'afinitat d'unió específica de LexA a cada una de les caixes. Empíricament es va calcular, que sota el patró HI (Lewis *et al.*, 1994), la unió es donava si el valor d'aquest era menor de 15 (taula 1.2). Per tant, a major valor de l'índex heteròleg, major serà la facilitat amb que s'expressarà un gen. És possible també la presència de més d'una d'aquestes seqüències operadores per gen, com passa per exemple en *lexA* i *recN* (taula 2) (Rostas *et al.*, 1987).

**Taula 1.1.** Gens del reguló *lexA* a *E. coli*.

<b>Gens cromosòmics</b>	<b>Funció</b>
<i>dinA (polB)</i>	DNA polimerasa III
<i>dinB (dinP)</i>	DNA polimerasa IV
<i>dinD (pscA)</i>	Desconeguda
<i>dinF</i>	Desconeguda. Està en el mateix operó que <i>lexA</i>
<i>dinG</i>	Desconeguda
<i>dinI</i>	Inhibició de l' <i>umuD</i> . Estabilitzador dels filaments del RecA
<i>dinL (sosC, yjiW)</i>	Desconeguda
<i>dinM (sosD, ydjK)</i>	Possible homòleg de l' <i>uvrC</i>
<i>dinO (sosF, molR)</i>	Desconeguda
<i>dinQ</i>	Desconeguda
<i>dinS</i>	Possible transposasa
<i>fsk (dinH)</i>	Divisió cel·lular i segregació cromosòmica
<i>hokE (ydbY)</i>	Possible proteïna assassina
<i>lexA</i>	Repressor del sistema SOS
<i>recA</i>	Recombinació i coproteasa. Regulador positiu del sistema
<i>recN</i>	Recombinació per la via RecF; reparació de trencament de cadena doble i senzilla
<i>recX</i>	Possible modulador del sistema SOS: inhibició funcions RecA
<i>ruvAB</i>	Recombinació per la via RecF. Resolvasa
<i>sbmC</i>	Resistència a la microcina B17
<i>ssb</i>	Unió al DNA de cadena senzilla
<i>sulA (sfiA)</i>	Inhibició de la divisió cel·lular
<i>umuDC</i>	Mutagènesi SOS. DNA polimerasa V
<i>uvrA (dnaE)</i>	Reconeixement dels dímers de pirimidina i reparació dels mateixos
<i>uvrB</i>	Reconeixement dels dímers de pirimidina i reparació dels mateixos
<i>uvrD</i>	Helicasa II
<i>ybfE</i>	Funció desconeguda
<i>ydjM</i>	Funció desconeguda
<i>yebG</i>	Funció desconeguda
<i>ysdAB</i>	Funció desconeguda
<b>Gens extracromosòmics</b>	
<i>colA (caa)</i>	Producció de colicines (CoIA)
<i>colB (cea)</i>	Producció de colicines (CoIB)
<i>impAB</i>	Paper semblant al de l' <i>umuDC</i> . Implicat en mutagènesi SOS
<i>mucAB</i>	Paper semblant al de l' <i>umuDC</i> . Implicat en mutagènesi SOS
<i>samAB</i>	Paper semblant al de l' <i>umuDC</i> . Implicat en mutagènesi SOS
Repressors fàgics*	Inducció de profags

**Taula 1.2.** Valor de l'índex heteròleg de les caixes SOS del reguló *lexA* a *E. coli*

Gens	Caixa SOS*	HI
<i>dinA</i> ( <i>polB</i> )	GACTGTATAAAACCACAGCC	12.09
<i>dinB</i> ( <i>dinP</i> )	CACTGTATACTTTACCAGTG	12.84
<i>dinD</i> ( <i>pscA</i> )	AACTGTATATAAATACAGTT	3.34
<i>dinG</i>	TAT <b>TGG</b> CTGTTTATACAGTA	8.98
<i>dinI</i>	ACCTGTATAAATAACCAGTA	6.24
<i>dinL</i> ( <i>sosC</i> , <i>yjiW</i> )	TACTG <b>A</b> TGATATATACAGGT	9.40
<i>dinM</i> ( <i>sosD</i> , <i>ydjQ</i> )	CACTG <b>G</b> ATAGATAAACAGCA	9.54
<i>dinO</i> ( <i>sosF</i> , <i>molR</i> )	AACTG <b>G</b> ATAAAATTACAGGG	10.68
<i>dinQ</i>	TACTGTATGATTAT <b>CC</b> AGTT	4.84
<i>dinS</i>	AGCTGTATTTGTCT <b>CC</b> AGCA	9.60
<i>ftsK</i> ( <i>dinH</i> )	TCCTGTTAATCCATACAGCA	8.61
<i>hokE</i> ( <i>ybdY</i> )	CACTGTATAAATAAACAGCT	6.72
<i>lexA/dinf</i> (1)	TGCTGTATATACTC <b>AC</b> AGCA	6.34
<i>lexA/dinf</i> (2)	TGCTGTATATACAC <b>CC</b> AGGG	8.42
<i>recA</i>	TACTGTATGCTC <b>AT</b> ACAGTA	4.31
<i>recN</i> (1)	TACTGTATATAAAA <b>CC</b> AGTT	5.16
<i>recN</i> (2)	TACTGTACACAATAACAGTA	9.38
<i>recN</i> (3)	TAA <b>ATGG</b> TTTTTCATACAGGA	11.47
<i>ruvAB</i>	CGCTG <b>G</b> ATGTCTAT <b>CC</b> AGCA	9.19
<i>sbmC</i>	TACTGTATATAAAAA <b>AC</b> AGTA	2.77
<i>Ssb</i>	ACCTG <b>A</b> ATGAATATACAGTA	6.98
<i>sulA</i> ( <i>sfiA</i> )	TACTGTACATCC <b>AT</b> ACAGTA	4.65
<i>umuDC</i>	TACTGTATATAAAAA <b>AC</b> AGTA	2.77
<i>uvrA</i> ( <i>dinE</i> )	TACTGTATATTCAT <b>CC</b> AGGT	6.98
<i>uvrB</i>	AACTGTTTTTTTTAT <b>CC</b> AGTA	6.11
<i>uvrD</i>	ATCTGTATATATAC <b>CC</b> AGCT	8.80
<i>ybfE</i>	AACTG <b>A</b> TTAAAA <b>ACC</b> AGCG	14.00
<i>ydjM</i> (1)	TACTGTACGTATCG <b>AC</b> AGTT	13.6
<i>ydjM</i> (2)	CACTGTATAAAAT <b>CC</b> TATA	9.07
<i>yebG</i>	TACTGTATAAAATC <b>AC</b> AGTT	7.03
<i>ysdAB</i>	CACTGTTTTATTTATACAGTA	2.35
<b>Consens</b>	TACTGTATATATATACAGTA	

\*Les bases ombrejades assenyalen els motius més conservats. Les bases en negreta assenyalen desviacions dels motius conservats CTGT i ACAG.

### 1.1.3. Funcions del sistema SOS

El sistema SOS centra la major part de les seves funcions en la reparació de DNA, amb la finalitat de garantir la supervivència cel·lular. Podríem classificar les funcions en 3 grups: reparació, mutagènesi SOS i funcions no directament relacionades amb la reparació. A continuació es detallaran les funcions integrades en cada un dels grups.

#### 1.1.3.1 Reparació per escissió de nucleòtid

Mitjançant aquest mecanisme, es repara la lesió eliminant les bases danyades i substituint-les per d'altres, usant com a motlle, la cadena inalterada. Aquest per tant, en condicions normals, és un mecanisme sense tendència a l'error. Entre els agents que poden causar lesions reparades per aquest sistema, podem citar la radiació UV, junt amb un ampli ventall d'agents físics i químics (Friedberg *et al.*, 1995).

En aquest mecanisme hi intervenen els productes dels gens *uvr*: *uvrA*, *uvrB*, *uvrC* i *uvrD*. Tot i així, el gen *uvrC* no forma part del reguló *lexA*. En aquest procés però, hi intervenen altres proteïnes com són la DNA polimerasa I i la DNA lligassa. Una descripció més detallada del procés seria:

- i) **Identificació de la lesió i unió del DNA.** Un dímer de la proteïna UvrA (UvrA<sub>2</sub>) es desplaça al llarg de la cadena de DNA, detectant la presència d'una lesió fruit de l'alteració que d'ella es deriva en la topologia del DNA. Quan el complex UvrA<sub>2</sub> s'atura, la verificació de la presència del dany es realitza per part de la proteïna UvrB (UvrA<sub>2</sub>B) que amb activitat helicasa i consum d'energia en forma d'ATP, s'uneix de manera estable en el punt de la lesió (Moolenaar *et al.*, 1994), formant el que es coneix com a complex de pre-incisió. En la funcionalitat de la proteïna UvrB té especial importància la estructura  $\beta$ -harpin que presenta (Moolenaar *et al.*, 2001). En aquest punt del procés, es produeix l'alliberament de la proteïna UvrA.
- ii) **Escissió de nucleòtid.** En aquest moment es quan intervé la proteïna UvrC, que en interacció amb el complex de pre-incisió, exerceix com a nucleasa amb la finalitat d'eliminar la cadena lesionada, practicant el tall a nivell dels enllaços fosfodiéster. El tall a l'extrem 5' de la lesió, es realitza per mitjà de l'acció d'UvrC a 8 nucleòtids de la lesió, en canvi, l'eliminació del fragment es realitzarà per l'acció d'UvrB a 4 nucleòtids de la lesió en direcció 3'. Quan el fragment es alliberat, la proteïna UvrC es desprèn del complex proteïc, permetent l'entrada de la proteïna



UvrD amb activitat helicasa (Helicasa II) (Van Houten i Snowden, 1993; Zou *et al.*, 1997).

- iii) **Síntesi *de novo* de la cadena.** Com s'ha dit anteriorment, la proteïna UvrD presenta activitat helicasa (Helicasa II) que junt amb la DNA polimerasa I, intervindran en la fase final de la reparació, usant com a motlle la cadena complementària inalterada (Moolenaar *et al.*, 2000). La síntesi per part de la DNA polimerasa I, a la qual cedeix el seu lloc en el complex la proteïna UvrB, es possible gràcies a la presència de dNTPs i a l'acció de les lligasses, l'activitat de les quals culminarà la reparació.

El mecanisme exposat, és el que es podria considerar com a canònic, en la reparació per escissió de nucleòtid, però podríem trobar variants en el seu desenvolupament. Així, últimament ha estat estudiat i descrit el gen *ydjQ*. El seu producte proteic Cho, presenta una funció anàloga a UvrC, però a diferència d'aquest últim, sí que forma part del reguló *lexA*. La proteïna Cho, junt amb la proteïna UvrB, formaria un complex actiu que podria prendre part en la reparació de determinades lesions, conjuntament amb la proteïna UvrC (Moolenaar *et al.*, 2002).

També cal dir que el sistema descrit és el que es coneix com a *short patch repair*, però en la cèl·lula també trobem el conegut com a *long patch repair*, que actua quan el primer no ha donat resultats satisfactoris. La diferència entre els dos, rau en la grandària del fragment eliminat: en el primer la longitud no és superior als 12 nucleòtids, mentre que el segon, pot arribar als 1.500.

### 1.1.3.2. Reparació per recombinació

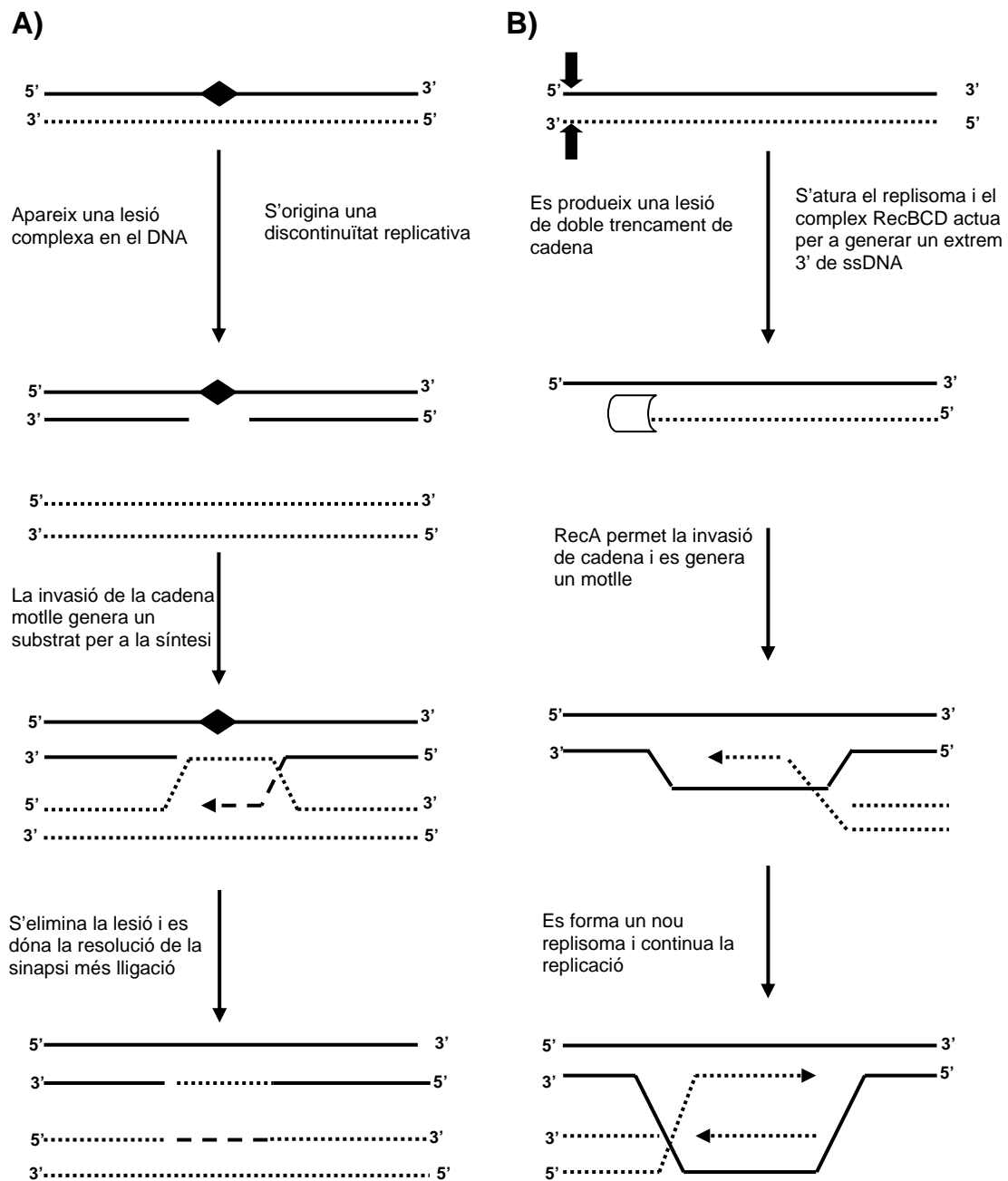
Aquest tipus de reparació el trobem quan les lesions no poden ser resoltes pels mecanismes de reparació per escissió, degut a l'existència d'enllaços entre les dues hebres, trencaments de cadena, discontinuïtats o danys que han implicat alteracions en la cadena motlle en períodes de síntesi nucleotídica. Un cas quantitativament important d'actuació, és en les discontinuïtats generades quan la replicació es veu aturada per l'existència d'una de les lesions abans esmentades: la DNA polimerasa literalment fa un salt, deixant zones sense replicar que poden ser de milers de nucleòtids (figura 1.7).

El fet que parlem de reparació per recombinació, no implica que aquest sistema actui directament sobre elles, si no que en molts casos crea un substrat susceptible de ser reparat per altres sistemes. Per mitjà d'ell, s'aconsegueix que un dany present en ambdues cadenes, passi tant sols a estar present en una d'elles, essent més fàcil la seva reparació, per exemple, per algun

dels mecanismes anteriorment esmentats (Kuzminov *et al.*, 1999; Dronkert *et al.*, 2001; Dudás *et al.*, 2004).

Existeixen dues vies principals que duen a terme la reparació per recombinació: la via *recBCD* i la via *recF*. És aquesta segona la que està sota el control del sistema SOS, tot i que intervé quan la *recBCD* no és funcional (Ivancic-Bace *et al.*, 2002), i estaria implicada en els casos de bloqueig de la replicació. De tots els gens que formen part de la via *recF* (*recF*, *recJ*, *recO*, *recR*, *recQ* i *recN*) l'únic que estaria sota el control de LexA, seria el *recN* (Kuzminov *et al.*, 1999). La via *recBCD*, que és la principal, estaria implicada en la reparació en els casos de trencament de doble cadena, per mitjà de la formació d'un holoenzim amb activitat mixta exonucleasa-helicasa (Anderson i Kwalczykowski, 1998). Pel funcionament d'aquestes vies, seria imprescindible la intervenció de les proteïnes RecA i Ssb (Roca i Cox, 1990; Cox *et al.*, 2000).

En les fases del procés de reparació per recombinació corresponents a la migració de cadena i la fase postsinàptica, intervenen també altres proteïnes, com per exemple la proteïna RecG amb activitat helicasa i els productes dels gens *ruvABC*. Tanmateix, de tots ells, tant sols els gens *ruvA* i *ruvB*, estan directament regulats per LexA (Kogoma *et al.*, 1997; Kuzminov *et al.*, 1999). També cal destacar, la presència d'altres vies menys importants i que tant sols intervenen en casos especials, com seria la via RecE. Aquesta actua en relació a un fag críptic i mutacions en el gen *sbcA*, en general, junt al producte del gen *recT* : una proteïna anellant important en aquesta via (Kolodner *et al.*, 1994; Kusano *et al.*, 1994; Silberstein *et al.*, 1995).



**Figura 1.7.** Esquema del procés de recombinació davant d'una discontinuïtat replicativa derivada d'una lesió complexa (A) i davant d'un trencament de cadena (B).

### 1.1.3.3. Mutagènesi SOS

Aquest és un procés que es troba sota una estreta regulació i que tant sols s'activa quan el sistema SOS es troba molt induït. L'essència de la seva funció, seria permetre la continuació de

la replicació tot i l'elevada presència de danys en el DNA, esperant resoldre'ls de manera posterior per mitjà de diferents mecanismes. Per això, gràcies a la polimerasa V, una base provisional es ubicada davant de les diferents lesions, no necessàriament la correcta, capacitat que no presenta la polimerasa III. Per aquest motiu, la mutagènesi SOS ha estat anomenada també com a síntesi de translesió o reparació tendent a error. El fet de que estigui sotmesa a una estreta regulació, es degut a que el seu mecanisme d'actuació engendra certs riscos per la supervivència cel·lular.

Així, tot i que l'actuació de *bypass* per mitjà de la inserció indiscriminada de bases, és un mecanisme que permet concloure la replicació, implica una taxa elevada de mutagènesi, amb els perills que això comporta. No obstant, aquest fet és una arma de doble fil, ja que a nivell poblacional estem generant una variabilitat que pot ser útil com alternativa davant la pressió selectiva, que es pot estar donant ja en el moment de la intervenció, i que d'ella en deriven els nombrosos danys en el DNA.

Els dos elements centrals d'aquest sistema són els gens *umuDC* (que a *E. coli* es troben en la mateixa unitat transcripcional), el producte proteic dels quals, correctament complexats, donaran lloc a la polimerasa V. A *E. coli*, els gens *umuDC* es troben en la mateixa unitat transcripcional. La proteïna UmuD, presenta un domini carboxi terminal molt semblant al del LexA, amb un punt de trencament entre l'enllaç Cys<sub>24</sub>-Gly<sub>25</sub> (Sutton *et al.*, 2000; Sutton *et al.*, 2001). Degut a l'activació del sistema SOS, augmenta l'expressió dels gens *umuDC*, i la presència del complex nucleoproteïc RecA\* (conformació activa), és el que permet el pas d'UmuD a UmuD' (Nohmi *et al.*, 1988). Seguidament, es formarà el complex UmuD'<sub>2</sub>UmuC, conegut com a polimerasa V (Tang *et al.*, 1999) que junt amb l'ajut de la proteïna RecA\* (RecA activada) (Ekaterina *et al.*, 1993), és el responsable de dur a terme la translesió (González i Woodgate *et al.*, 2002).

De nou, en aquesta via, la proteïna Ssb jugarà un paper important en facilitar la intervenció de la proteïna RecA. De fet, s'ha postulat que tant RecA com Ssb, són vitals per una correcta realització del *bypass* (Reuven *et al.*, 2001).

Com s'ha esmentat, aquest és un sistema que pot comportar perills per al *fitness* cel·lular. Això explica les múltiples mesures fixades per l'evolució, a fi de gaudir d'un estricte control de l'expressió i activitat d'aquests elements. Per exemple, la caixa SOS present en la seva regió reguladora, és per la que presenta major afinitat la proteïna LexA (Lewis *et al.*, 1994); la proteïna RecA actua més eficientment com a coproteasa en l'autohidròlisi de LexA que en la d'UmuD; la formació d'homodímers UmuD'-UmuD' es veu estequiomètricament desfavorida davant la formació del complex UmuD'-UmuD (Battista *et al.*, 1990b). Per últim, la proteïna DinI, també regularia el procés de translesió, ja que la proteïna RecA presenta més afinitat d'unió per ella que per la proteïna UmuD, limitant així la possibilitat d'activació d'aquesta última en un cas de gran inducció del sistema SOS (Yasuda *et al.*, 2001).

#### 1.1.3.4. Altres funcions del sistema SOS

De manera col·lateral a la inducció del sistema SOS degut a la l'activació de la proteïna RecA, també es donarà la inducció de profags (Sauer *et al.*, 1982), com a conseqüència de l'analogia estructural i funcional dels repressors del cicle lític dels bacteriòfags i la proteïna LexA (Mustard i Little, 2000).

També cal destacar com a funció no reparativa del sistema SOS, la inhibició de la divisió cel·lular, per l'acció del producte del gen *sulA* (Huisman *et al.*, 1981). El mode com actua, es impeding la septació cel·lular per actuació sobre la proteïna FtsZ fins que totes les lesions són reparades, resultant-ne un fenotip conegut com a filamentós. Tot i això, la proteïna SulA és molt làbil i per tant, tant sols es donarà la filamentació en casos d'una elevada inducció del sistema SOS (Walker, 1996).

## 1.2. El sistema SOS en altres microorganismes

### 1.2.1. El sistema SOS en els bacteris grampositius

El sistema SOS en els microorganismes grampositius també ha estat àmpliament estudiat, inicialment a *Bacillus subtilis* (Winterling *et al.*, 1997), i més recentment en *Mycobacterium tuberculosis* (Davis *et al.*, 2002).

Les primeres evidències de la presència d'un sistema homòleg al descrit a *E. coli*, daten de principis dels 90 (Wojciechowski *et al.*, 1991), en demostrar-se la capacitat d'autohidròlisi de la proteïna LexA d'*E. coli* en cèl·lules de *B. subtilis* tractades a fi de provocar la inducció del sistema SOS.

El repressor del sistema SOS a *B. subtilis*, rep el nom de *dinR* (*damage induction*) i la seva autohidròlisi es produeix entre els residus aminoacídics Ala<sub>91</sub>-Gly<sub>92</sub> (Winterling *et al.*, 1997). Tot i així, actualment s'accepta la terminologia de *lexA* per aquest gen.

El punt de major homologia amb altres proteïnes LexA, com s'ha dit ja, es troba en el domini carboxi terminal, amb els residus Ser<sub>119</sub> i Lys<sub>156</sub> (implicats en l'autohidròlisi), i en les zones de reconeixement entre els monòmers per la dimerització, que bàsicament es realitza com a *E. coli* (Winterling *et al.*, 1998). Evidentment, en el domini amino terminal, es troben més divergències, en ser el punt de reconeixement amb el DNA (Knegtel *et al.*, 1995).

Per altra part, la proteïna RecA presenta major homologia que el LexA amb els seus homòlegs en altres microorganismes, per ser com s'ha dit, extremadament important, i per tant molt conservada al llarg de l'evolució.

El motiu d'unió al DNA de la proteïna LexA a *B. subtilis*, és conegut com caixa "Cheo" (Cheo *et al.*, 1991; Yasbin *et al.*, 1992), i en una primera instància es descrigué com el motiu: GAAC-N4-GTTC. En estudis i revisions posteriors, es fixar la seqüència consens mitjançant tècniques de mutagènesis puntual, essent el resultat CGAACRNRYGTTYC (Winterling *et al.*, 1998). A posteriori i mercè a disposar de major quantitat de caixes SOS dels gens del reguló (taula 1.3), s'ha pogut fixar com a seqüència consens: CGAACATATGTTCG (Au *et al.*, 2005).

En tant la composició del reguló a *B. subtilis*, estudis molt recents han detectat la presència de 54 gens organitzats en 34 operons, que presenten una caixa SOS on DinR s'hi uneix específicament (Au *et al.*, 2005). D'aquests 54 gens, tant sols 33 (dels quals 25 es troben organitzats en 10 operons) mostren inducció posterior a tractaments amb UV i MC (taula 1.3), encaixant així dins el funcionament canònic del reguló (Au *et al.*, 2005). Tanmateix, també trobem gens sense cap motiu hipotètic d'unió, que mostren inducció per RecA, com d'altres amb caixa "Cheo" que no en mostren (Au *et al.*, 2005).

Per estudis de similitud, s'ha pogut determinar les funcions d'alguns d'aquests gens, així es pot parlar de funcions tals com la reparació per escissió, recombinació, filamentació (inhibició de la divisió cel·lular), mutagènesis SOS i altres funcions, com intervenció en reparació d'alquilacions (Au *et al.*, 2005) (taula 1.3).

De manera diferent a *E. coli*, cal dir que a *B. subtilis* el sistema SOS també s'indueix quan les cèl·lules presenten competència natural (Love *et al.*, 1985), però es fa de manera independent del RecA, per mitjà de la proteïna CTF, codificada pel gen *comK* (Haijema *et al.*, 1996; Hamoen *et al.*, 2001).

Una composició i funcionament del sistema SOS molt semblant al descrit (taula 1.4), el trobem en altres microorganismes grampositiu com ara: *Clostridium perfringens*, *Corynebacterium glutamicum*, *Mycobacterium leprae*, *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, *Staphylococcus aureus*, *Streptomyces lividans* i *Streptomyces coelicor*.

Com a contrapunt a aquesta notòria homogeneïtat, trobem que a *M. tuberculosis*, s'ha descrit una caixa SOS lleugerament diferent a la Cheo de *B. subtilis*: TCGAAC-N<sub>4</sub>-GTTCGA (Davis *et al.*, 2002). Amb aquesta informació es va realitzar un anàlisi dels gens que pertanyen al seu reguló *lexA*, mostrant que alguns gens induïbles per danys en el DNA, ho fan de manera independent del LexA. En treballs més recents, s'ha provat que aquest són majoria (Rand *et al.*, 2003).

Destacar que a *M. tuberculosis*, no trobem els gens homòlegs a l'*umuDC* d'*E. coli*, fet que suggereix la presència d'un sistema més acurat que el d'*error-prone*, per a garantir la seva supervivència, potser adaptada a les necessitats que impliquen medis més hostils (Movahedzadeh *et al.*, 1997a; 1997b; Brooks *et al.*, 2001).

**Taula 1.3.** Gens SOS de *B.subtilis*. Modificat de Au *et al* (2005)

<b>Gen<sup>a</sup></b>	<b>Caixa SOS</b>	<b>Factor d'inducció<sup>b</sup></b>
<b><i>ybaK</i><sup>c</sup></b>	aGAACATtTGTTc <sup>d</sup>	2.3
<i>cwlD</i>		2.1
<b><i>dinB</i></b>	aGAACtcATGTTcG	55
<b><i>ydiO</i></b>	aGAACATtcGTTcT	3.5
<i>ydiP</i>		3.6
<b><i>pcrA</i></b>	aGAACgTATGTTt	2.0
<i>ligA</i>		2.0
<b><i>yhaZ</i></b>	aGAACgTAcATTcC	11
<b><i>yhaO</i></b>	aGAACgTgcaTTcG	3.5
<i>yhaN</i>		1.6
<i>yhaM</i>		2.0
<b><i>yhjD</i></b>	aGAACAaAcGTTcC	16
<i>yhjC</i>		2.9
<i>yhjB</i>		3.8
<b><i>xkdA</i></b>	aGAACAcAcGTTcG	3.8
<b><i>recA</i></b>	CGAAATgCgTTcG	9.4
<b><i>aprX</i></b>	CGAACAAaAcGTTcT	2.7
<b><i>lexA</i></b>	gGAAtgTtTGTTcG	3.4
<b><i>yneA</i></b>	CaAACATAgGTTcG	37
<i>yneB</i>		44
<i>ynzC</i>		8.1
<b><i>parE</i></b>	CaAACATAcGTTcT	2.9
<i>parC</i>		3.0
<b><i>yqjW</i></b>	CGAACATAaGTTcT	3.5
<i>yqjX</i>		5.7
<i>yqjY</i>		1.8
<i>yqjZ</i>		1.9
<b><i>ruvA</i></b>	CGAACATATGTTaa	2.2
<i>ruvB</i>		2.2
<b><i>uvrC</i></b>	aaAACAAaAcGTTcG	1.6
<b><i>uvrB</i></b>	CGAACtTtaGTTcG	11
<i>uvrA</i>		12
<b><i>dinC</i></b>	aGAACAagTGTTcT	175

<sup>a</sup>Els gens escollits han estat els que mostraren dependència del RecA i una caixa SOS funcional.

<sup>b</sup>Factor d'inducció posterior al tractament amb Mitomicina C.

<sup>c</sup>En negreta els gens que es troben en l'extrem 5' de l'operó; en estil normal els subsegüents.

<sup>d</sup>En majúscula les bases que coincideixen amb la seqüència consens, en minúscula, els que no.

**Taula 1.4.** Possibles caixes SOS presents en diferents bacteris grampositius.

Microorganisme	Gen	Seqüència
<i>Bacillus subtilis</i>	<i>recA</i>	CGAATATGCCTTCG
	<i>dinA</i>	CGAACTTTAGTTCG
	<i>dinB</i>	AGAACTCATGTTTCG
	<i>dinC</i> (1)	CGAACGTATGTTTG
	<i>dinC</i> (2)	AGAACAAAGTGTTCG
	<i>dinR</i> (1)	CGAACCTCAGTTTG
	<i>dinR</i> (2)	CGAACAAACGTTTC
	<i>dinR</i> (3)	GGAATGTTTGTTTCG
	<i>uvrB</i>	AAAACAAACGTTTCG
	<i>dnaX</i>	CGAACCAAGGTTCA
<i>Bifidobacterium breve</i>	<i>recA</i> (1)	CGTTCACCCGCATC
	<i>recA</i> (2)	CGAACAAATGTTTCG
<i>Clostridium perfringens</i>	<i>recA</i>	AGAACTTATGTTTCG
<i>Corynebacterium glutamicum</i>	<i>recA</i>	CGTAGGAATTTTCG
<i>Corynebacterium pseudotuberculosis</i>	<i>recA</i>	AGAAATGGTCGTTAG
<i>Lactococcus lactis</i>	<i>recA</i>	TGATAGAAAGTTCC
<i>Mycobacterium leprae</i>	<i>recA</i> (1)	CGAACAGATGTTTCG
	<i>recA</i> (2)	CGTACTGCGATTTCG
	<i>lexA</i> (1)	CGAACACATGTTTG
	<i>lexA</i> (2)	CGAACATTTCGATCG
<i>Mycobacterium smegmatis</i>	<i>recA</i> (1)	CGAACAGGTGTTTCG
	<i>recA</i> (2)	GGAACACCGGTCAG
<i>Mycobacterium tuberculosis</i>	<i>recA</i>	CGAACAGGTGTTTCG
	<i>lexA</i>	CGAACACATGTTTG
<i>Spiroplasma melliferum</i>	<i>recA</i>	XGATCACGAGAACG
<i>Spiroplasma citri</i>	<i>recA</i>	TGATCACGAGAACA
<i>Staphylococcus aureus</i>	<i>recA</i>	CGAACAAATATTTCG
<i>Streptococcus mutans</i>	<i>recA</i>	CGAACATGCCCTTG
<i>Streptococcus pneumoniae</i>	<i>recA</i>	GGATCATTAGAATG
	<i>dinF</i>	TGAACTTGAAATCG
<i>Streptococcus pyogenes</i>	<i>recA</i>	CGATTAGGAGAACG
<i>Streptomyces lividans</i>	<i>recA</i>	CGAACATCCATTCT
<i>Streptomyces coelicolor</i>	<i>recA</i>	AGAATGGATGTTTCG
	<i>lexA</i>	CAAACACACGTTTCG
<i>Streptomyces clavuligenes</i>	<i>lexA</i>	CGTTTCGAGTGAAAA
<i>Streptomyces rimosus</i>	<i>recA</i>	CGAACGTCTATTCA
<i>Streptomyces ambofaciens</i>	<i>recA</i>	CGAACATCCATTCA
Consens		CGAACRNRYGTTTCG

Les bases ombrejades assenyalen els motius més conservats. Les bases en negreta assenyalen desviacions dels motius conservats CGAAC i GTTCG. R = G o A; Y = C o T.



## 1.2.2. El sistema SOS en diferents Proteobacteris

### 1.2.2.1. Grup dels Proteobacteris Alfa

Taxonòmicament parlant, dins aquest grup hi pertanyen entre altres els ordres *Caulobacterales*, *Rhizobiales* i *Rhodobacterales*. La majoria d'ells són organismes de vida lliure amb habitats molt variats igual que les seves rutes metabòliques. Estudis del sistema SOS realitzats en els microorganismes *Rhodobacter sphaeroides*, *Rhodobacter capsultus*, *Sinorhizobium melitoti* i *Paracoccus denitrificans* (Tapias *et al.*, 1997; 1999; Fernandez de Henestrosa *et al.*, 1998; Labazi *et al.*, 1999; del Rey *et al.*, 1999), provaren la presència d'un motiu d'unió del LexA, diferent al que es coneixia a *E.coli*, tant en l'estructura com en la seqüència: GTTC-N<sub>7</sub>-GTTC.

Com a organisme representant de l'ordre *Rhodobacterales* i per extensió dels  $\alpha$ -Proteobacteris, el més estudiat en tant el funcionament del sistema SOS ha estat

*R. sphaeroides*. En els esmentats estudis, es presenten una sèrie de divergències respecte el funcionament del sistema SOS a *E. coli*. Així, podem esmentar la viabilitat cel·lular de les soques *lexA*(Def) obtingudes pel nostre grup d'investigació (Tapias *et al.*, 2000).

Estudis posteriors a *R. sphaeroides*, varen permetre establir una relació no descrita fins llavors, entre la transcripció del gen *recA* i la proteïna LexA (Tapias *et al.*, 2002). El mecanisme, en el fons esdevindria un potenciador del control ja descrit a *E. coli*: LexA a baixes concentracions facilitaria un increment de la transcripció del gen *recA*, mentre que a altes concentracions, potenciaria el seu propi efecte inhibidor intrínsec.

Anàlisis informàtiques practicades en els genomes de diferents representants del grup, disponibles en les bases de dades, varen permetre obtenir la seqüència consens de la caixa SOS per aquest grup, així com determinar quin seria el nucli de gens canònics del reguló pels Proteobacteris Alfa, apareixent els següents: *lexA*, *recA*, *uvrA*, *ssb*, *sulA* i *dinP*. Aquesta selecció es realitzar en tant puntuacions atorgades a les diferents caixes SOS detectades en tant afinitat d'unió de la proteïna LexA (Erill *et al.*, 2004) (taula 1.5).

En aquest mateix treball (Erill *et al.*, 2004), es va poder determinar la importància relativa de cada una de les bases en les dues seqüències consens interespecífiques que es varen obtindre a partir de les dades experimentals de diferents  $\alpha$ -Proteobacteris (figura 1.8).

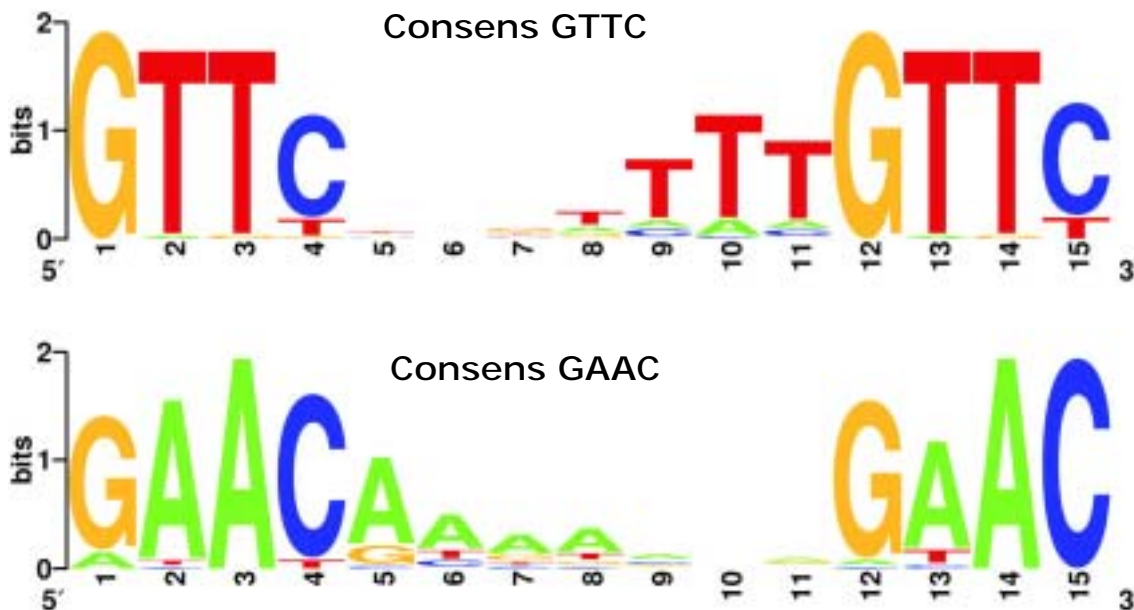
Com a excepcions als trets característics esmentats en els  $\alpha$ -Proteobacteris, citar els paràsits intracel·lulars *Rickettsia conorii* i *Rickettsia prowazekii*, que no presenten gen *lexA*, segurament degut a motius de *fitness* cel·lular que han conduït a una dràstica disminució de la grandària del seu genoma (Erill *et al.*, 2004), i amb ella la pèrdua d'alguns gens, incidint de nou en la varietat de regulacions, o fins i tot de no regulacions, idònies per cada nínxol.

**Taula 1.5.** Caixes SOS del grup dels  $\alpha$ -Proteobacteris.

Microorganisme	Gen	Seqüència <sup>a</sup>
<i>Rhodobacter sphaeroides</i>	<i>recA</i> (1)	GTT <b>C</b> GCCTTATG <b>A</b> TC
	<i>recA</i> (2)*	GTT <b>C</b> GCCTCAAGTTC
	<i>lexA</i> (1)*	GTT <b>C</b> TGCCCGCGTTC
	<i>lexA</i> (2)*	GTT <b>C</b> ACCGCCTGTTC
	<i>uvrA</i>	GTT <b>C</b> ATACTATGTTC
	<i>uvrB</i> (1)	G <b>C</b> TCCGCCCTTGTTC
	<i>uvrB</i> (2)*	G <b>A</b> TCCGTTTTTGTTC
<i>Rhodobacter capsulatus</i>	<i>recA</i> (1)	GTT <b>C</b> CGAAATTGTTC
	<i>recA</i> (2)*	GTT <b>C</b> TGCTTTCGTTC
	<i>uvrA</i>	GTT <b>C</b> CTGTTCCGTTC
<i>Rhodopseudomonas viridis</i>	<i>recA</i> (1)	GTT <b>C</b> TCTTCTTGTTC
	<i>recA</i> (2)*	G <b>T</b> A <b>C</b> ACGATTTGTTC
<i>Rhodopseudomonas palustris</i>	<i>recA</i> (1)*	GTT <b>C</b> GCAATATGTTC
	<i>recA</i> (2)*	G <b>T</b> A <b>C</b> CCCTATTTGTTC
<i>Agrobacterium tumefaciens</i>	<i>recA</i> (1)*	GTT <b>C</b> TGCTTTCGTTC
	<i>recA</i> (2)*	G <b>T</b> A <b>C</b> TCTATTTGTTC
<i>Rhizobium etli</i>	<i>recA</i> (1)*	GTT <b>C</b> TATATTTGTTC
	<i>recA</i> (2)*	G <b>T</b> A <b>C</b> CCCTATTTGTTC
<i>Sinorhizobium meliloti</i>	<i>recA</i> (1)*	GTT <b>C</b> GATTCTTGTTC
	<i>recA</i> (2)*	G <b>T</b> A <b>C</b> ATGTTTTGTTC
	<i>uvrA</i>	GTT <b>C</b> TTTTTTTGTTC
<i>Mesorhizobium loti</i>	<i>recA</i> (1)*	GTT <b>C</b> TTTTTTTCG <b>T</b> A <b>C</b>
	<i>recA</i> (2)*	G <b>T</b> A <b>C</b> CTTTTTTGTTC
<i>Brucella abortus</i>	<i>recA</i> (1)	GTT <b>C</b> GTGGATAGTTC
	<i>recA</i> (2)*	GTT <b>C</b> CATTCTTGTTC
	<i>uvrA</i>	GTT <b>C</b> GATATTTGTTC
	<i>Ssb</i> *	GTT <b>C</b> CTGTTTTGTTC
<i>Paracoccus denitrificans</i>	<i>recA</i> *	GTT <b>C</b> ACGGGTTGTTC
	<i>uvrA</i>	GTT <b>C</b> CTGTGATGTTC
<i>Aquaspirillum magnetotacticum</i>	<i>recA</i>	GTT <b>C</b> TCCTCTCGTTC
<i>Acidiphilium facilis</i>	<i>recA</i> *	GTT <b>T</b> TGTCAACGTTC
<i>Zynomonas mobilis</i>	<i>recA</i>	GTT <b>C</b> ACCTTATGTTC
	<i>uvrA</i> (1)	<b>A</b> TTCCCCCTTTGTTC
	<i>uvrA</i> (2)	<b>A</b> TTCTGCTACCGTTC
<i>Sphingomonas aromaticivorans</i>	<i>recA</i> (1)	GTT <b>C</b> CCCCCTTGTTC
	<i>recA</i> (2)*	G <b>T</b> A <b>C</b> TCGTTGTGTTC
Consens		GTT <b>C</b> YYYTTTTGTTC

<sup>a</sup>Les bases ombrejades assenyalen els motius més conservats; les que apareixen en negreta assenyalen desviacions dels motius conservats GTTC.

(\*) Seqüències reverses complementàries de la caixa original per facilitar la comparació



**Figura 1.8.** Caixes SOS consens interespecífiques del grup dels  $\alpha$ -Proteobacteris, obtingudes per mitjà de dades experimentals i anàlisi preliminar.

### 1.2.2.2. Grup dels Proteobacteris Gamma

Com a màxim representant d'aquest grup trobem *E. coli*, l'organisme del que més es coneix el seu sistema SOS i exemple d'un dels dos grups diferenciabls en aquesta subclasse: un de patògens de vertebrats sovint entèrics, i d'altres que el seu nínxol ecològic es el sòl. Són organismes amb una relació filogenètica molt estreta, amb el que el seu sistema SOS, tant en la seqüència del motiu d'unió com en la composició del reguló o la seqüència de les proteïnes clau, per exemple LexA i RecA, es preveu també força conservat. Aquest fet s'ha pogut corroborar en diferents espècies: *Pseudomonas sp.*, *Salmonella sp.*, *Erwinia sp.*, *Providentia rettgeri* i *Aeromonas sp.* (Calero *et al.*, 1991; Garriga *et al.*, 1992; Riera i Barbé, 1993; Riera i Barbé, 1995). Com a contrapunt a l'apuntada homogeneïtat en aquest grup, citar el patògen vegetal *Xylella fastidiosa*, en el que es va descriure un motiu d'unió diferent al conegut a *E. coli*: TTAGN<sub>6</sub>TACTA (Campoy *et al.*, 2002).

En tant la composició del reguló, s'ha detectat certa divergència (Erill *et al.*, 2003), tot i que es podria definir un nucli de gens que són comuns a totes les espècies analitzades fins ara: *lexA*, *recA*, *ssb* i *uvrA* (Erill *et al.* 2004).

### 1.2.2.3. Grup dels Proteobacteris Delta

Els organismes estudiats en tant el seu sistema SOS pertanyents a aquest grup taxonòmic, presenten elevada heterogeneïtat en els seus habitats i cicles vitals, del que se'n deriva la disparitat dels diferents sistemes SOS, tant en la seqüència d'unió de LexA com en el funcionament del reguló en general.

Com a exemple dels mixococs trobem *Myxococcus xanthus*. Aquest és un bacteri que presenta certes peculiaritats, com per exemple, que en el seu cicle vital apareix un estadi reproductiu en forma de cossos fructífers. El motiu o caixa SOS de *M. xanthus* és CTRHAMRYBYGTTTCAGS (Campoy *et al.*, 2003). Amb la variabilitat que permet la degeneració, aquest motiu recorda al descrit a *E. coli* (5'-CTGT(AT)<sub>4</sub>ACAG-3'), tot i que la proteïna LexA d'aquest últim, no s'uneix al motiu del primer. En el seu genoma, presenta una sola còpia del gen *lexA*, però no així del gen *recA*: *recA*(1) i *recA*(2) (Norioka *et al.*, 1995). D'ambdós gens, tant sols el *recA*(2) és induïble per danys en el DNA, en altres paraules, es troba sota el control de la proteïna LexA, ja que s'expressa constitutivament en una soca de laboratori *lexA*(Def) (Campoy *et al.*, 2003). Aquest mateix mutant, mostra viabilitat i les seves característiques no divergien significativament de les detectades en la soca salvatge. Un altra peculiaritat, és que a diferència del que succeeix a *E. coli*, els gens *recN* i *ssb* en la soca ja esmentada *lexA*(Def), eren induïts en presència de danys en el DNA, fet que suggereix un mode de regulació diferent al descrit fins ara, reafirmant la validesa del sistema SOS com a marcador evolutiu i filogenètic, en plasmar-se en ell les diferències degudes a la pertinença a grups taxonòmics diferents.

Un altre membre de la subdivisió Delta dels Proteobacteris del que es té coneixement del funcionament del seu reguló *lexA*, és *Geobacter sulfurreducens*. Aquest microorganisme destaca per la seva capacitat d'oxidar a CO<sub>2</sub> diferents compostos orgànics, usant com a acceptor final d'electrons fumarat, Fe(III), sofre elemental o malat (Lovley, 2000). Primerament es va ubicar com a anaeròbic estricte, però recentment s'ha observat que pot tolerar la presència d'oxigen.

Estudis *in silico* realitzats sobre el seu genoma revelaren la presència de dos gens *lexA* anomenats *lexA1* i *lexA2* (Jara *et al.*, 2002). En ambdós casos, i comprovat per mitjà d'RT-PCR, es troben formant unitat transcripcional amb un homòleg al gen *dinB*

d'*E. coli*. Els assaigs EMSA ens mostraren que ambdós proteïnes s'unien al palíndrom imperfecte GGTTN<sub>2</sub>CN<sub>4</sub>GN<sub>3</sub>ACC, que trobem corrent amunt de l'hipotètic inici de traducció dels dos gens *lexA*. Aquest motiu també s'ha detectat davant del gen *lexA* de *Geobacter metallireducens*, indicant que aquesta és la caixa SOS del gènere.

Un altre organisme estudiat en el nostre grup en tant la composició i estructura del seu reguló SOS, és *Bdellovibrio bacteriovorus*. Aquest és un bacteri gramnegatiu, amb forma de vibri, que s'alimenta d'altres microorganismes i que presenta un cicle vital amb alternança de fases: en

una d'elles presenta una proliferació independent mentre que en l'altra es veu obligat a envair l'espai periplasmàtic d'algun altre bacteri gramnegatiu (Starr *et al.*, 1971). Estudis en aquest organisme, revelaren que el motiu d'unió de LexA era TTACN<sub>3</sub>GTAA (Campoy *et al.*, 2004). Sense tenir en compte el gen *lexA*, tant sols es va detectar aquest motiu d'unió davant de l'operó Bd0386-Bd0385-Bd0384-Bd0383, que junt amb altres evidències, conduïren a pensar que es podia tractar de l'equivalent a l'operó *sulA-dinP-dnaE*. El motiu d'unió descrit, no es reconegut pel LexA d'altres integrants del grup com *G. sulfureducens* i *M. xanthus*, però en canvi si que ho fa el LexA d'*Anabaena sp.* PCC7120. S'ha provat que la caixa SOS d'aquest últim, està íntimament relacionada amb la dels bacteris grampositius (Mazón *et al.*, 2004), fet que hipotèticament ens ubica *B. bacteriovorus* com a possible origen de la resta d'espècies pertanyents als grups més recents de Proteobacteris (Campoy *et al.*, 2004).

#### 1.2.2.4. El sistema SOS als cianobacteris

Els cianobacteris són organismes procariotes fotoautotròfics amb capacitat per créixer en habitats molt variats tals com ambients aquàtics continentals, marins o en el sòl.

Malgrat es tenien evidències de l'existència del sistema SOS en aquest grup taxonòmic i dades del gen *recA* a *Anabaena* PCC7120, *Synechococcus* PCC7002 i *Synechocystis* PCC6308 (Geoghegan *et al.*, 1987; Murphy *et al.*, 1987; Owtrim i Coleman, 1987), fins recentment no s'ha disposat de més informació del reguló *lexA* en aquest filum.

L'organisme més estudiat en aquesta divisió es *Anabaena* PCC7120 (Mazón *et al.*, 2004). Mitjançant la informació obtinguda en els estudis del reguló SOS d'aquest cianobacteri, es pogué elucidar la seqüència consens per aquest grup filogenètic: RGTACNNNDGTWCB.

La cerca d'aquest motiu en els genomes d'altres cianobacteris (Mazón *et al.*, 2004), es el que va permetre afirmar que el motiu RGTACNNNDGTWCB era àmpliament present en tot el grup. Així fou detectada en *Nostoc punctiforme*, *Synechocystis* PCC6803, *Synechococcus sp.* WH8102, *Prochlorococcus marinus* MIT9313 i *Prochlorococcus pastoris* CCMP1378. Com a nucli de gens catalogables de canònics en el reguló *lexA* dels cianobacteris, es podria acceptar el format pels gens *lexA*, *recA*, *uvrA* i *ssb*, que coincidirien parcialment amb els descrits a *E. coli* (Fernández de Henetrosa *et al.*, 2000).

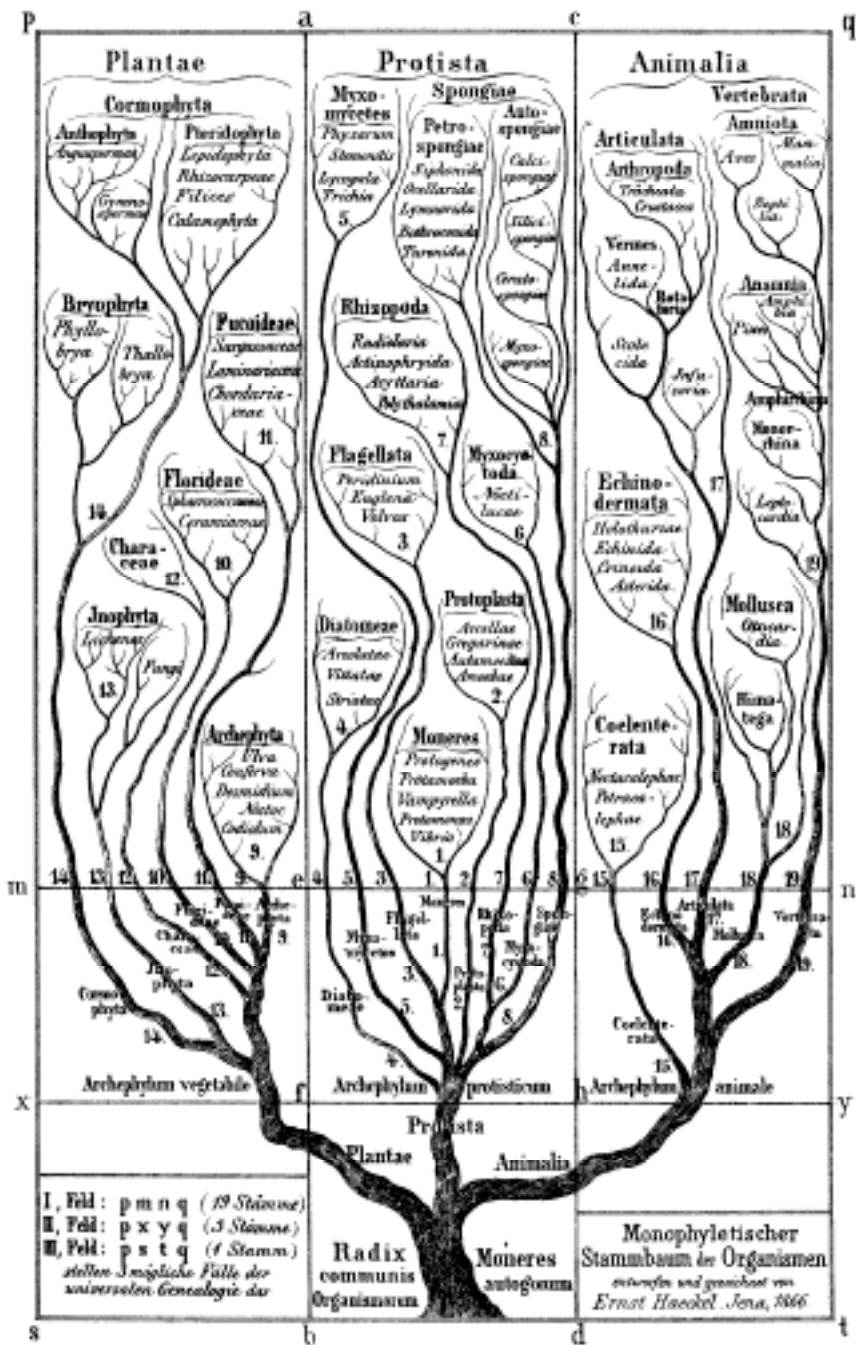
L'estudi detallat de la caixa SOS dels cianobacteris (RGTACNNNDGTWCB), revela una estreta relació amb els bacteris grampositius (CGAACRNRYGTTYC), provada mitjançant diferents experiments (Mazón *et al.*, 2004) i que estava en comunió amb la hipòtesi derivada dels resultats obtinguts a través de l'estudi de les delecions i insercions de diferents residus aminoacídics en proteïnes *house keeping* (Gupta i Griffiths, 2002). D'acord amb la hipòtesi formulada, els bacteris grampositius haurien provingut d'un ancestre comú anterior als bacteris

verds no sulfurosos, que al seu temps ho hauria fet d'un anterior als cianobacteris (Gupta i Griffiths, 2002).

### **1.3. Estudis filogenètics**

Ençà que el botànic suec Carl von Linne, a finals de segle XVIII, va proposar la classificació dels diferents organismes en categories jeràrquiques (espècie, gènere, família, ordre, classe, filum i regne), la classificació i seguiment dels lligams de filiació entre els diferents organismes, ha estat un objectiu científic destacat. Posteriorment a la publicació de la teoria de l'evolució per part de Charles Robert Darwin, en la seva obra *L'origen de les espècies* que data del 1859, el biòleg i filòsof alemany Ernst Haeckel, gràcies a l'ajuda d'un microscopi, introduí el món microbià en la classificació de la vida (figura 1.9). Tot i que avui en dia sabem que l'organització que proposava es incorrecta, cal valorar el fet de fixar l'origen de totes les espècies a partir d'un ancestre comú, com es pot observar en la representació gràfica original del seu discurs (figura 1.9), esdevenint així una aproximació al concepte anteriorment esmentat del Luca.

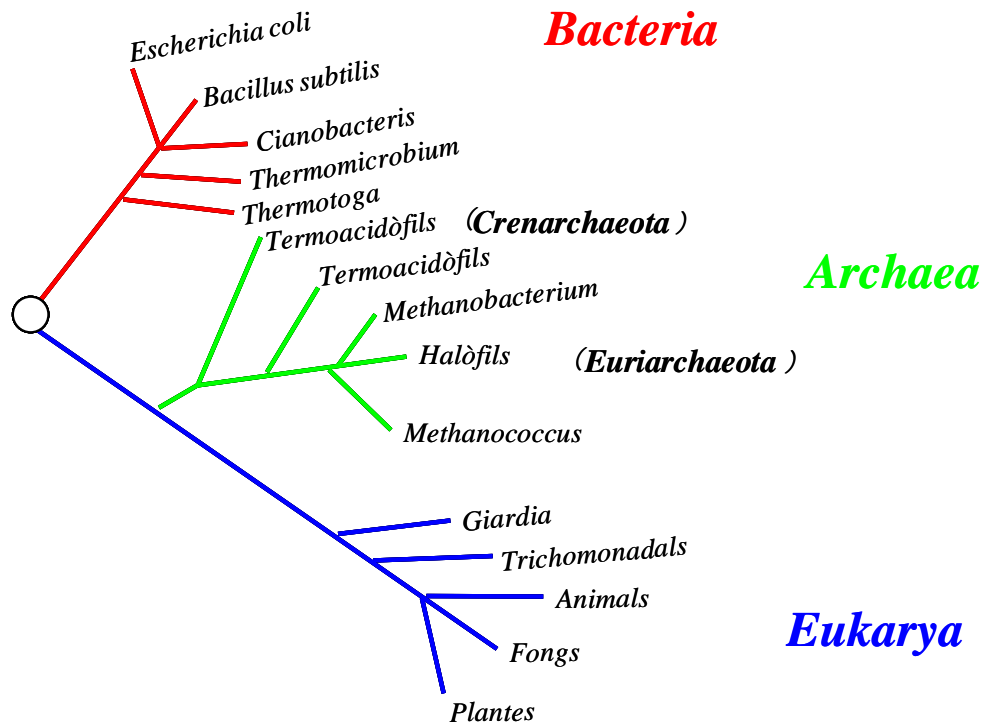
En aquests inicis de la ciència filogenètica, es varen usar patrons que actualment es sap que no són vàlids per a elucidar els vincles reals, com era la semblança morfològica. Avui en dia, en disposar d'una comprensió més elevada de la biologia, les eines escollides són a nivell molecular, i especialment, si estan implicades en funcions cel·lulars bàsiques, tant per a garantir la minimització de la deriva gènica (com s'ha dit anteriorment) com per a permetre'n l'aplicabilitat a tots, o la major quantitat possible d'espècies o grups.



**Figura 1.9.** Representació gràfica de l'organització filogenètica proposada per Haeckel. La seva classificació es basava en la ubicació de tots els organismes en tres Regnes: *Plantae*, *Protista* i *Animalia*. Es necessari destacar la seva estructura entroncada donant idea de l'origen comú de totes les espècies presents en la biosfera.

Tot i que posteriorment es continuaren realitzant estudis i aproximacions, una fita important fou la publicació per part del microbiòleg Carl Woese i el seu equip de treball, d'un nou arbre de la vida, realitzat per mitjà de les dades obtingudes de seqüenciar l'RNA ribosomal d'un flavobacteri, un cianobacteri, un bacteri vermell, un bacteri fotosintètic, un bacteri grampositiu, bacteris metanococals, termococals i metanobacteris, un ciliat, un fong, un animal, una planta verda, un flagel·lat i un microsporidi. Un cop es disposar d'aquesta informació, s'establiren les

relacions de filiació entre els grups que representaven cada un dels organismes escollits, a partir de la seqüència de l'rRNA. En referència als procariotes i profunditzat en estudis posteriors (Woese, 1987), la classificació filogenètica amb el 16 sRNA, diferencia dos grans grups: els arquees i els bacteris, originats per divergència evolutiva a partir d'un ancestre comú (Woese, 1987).

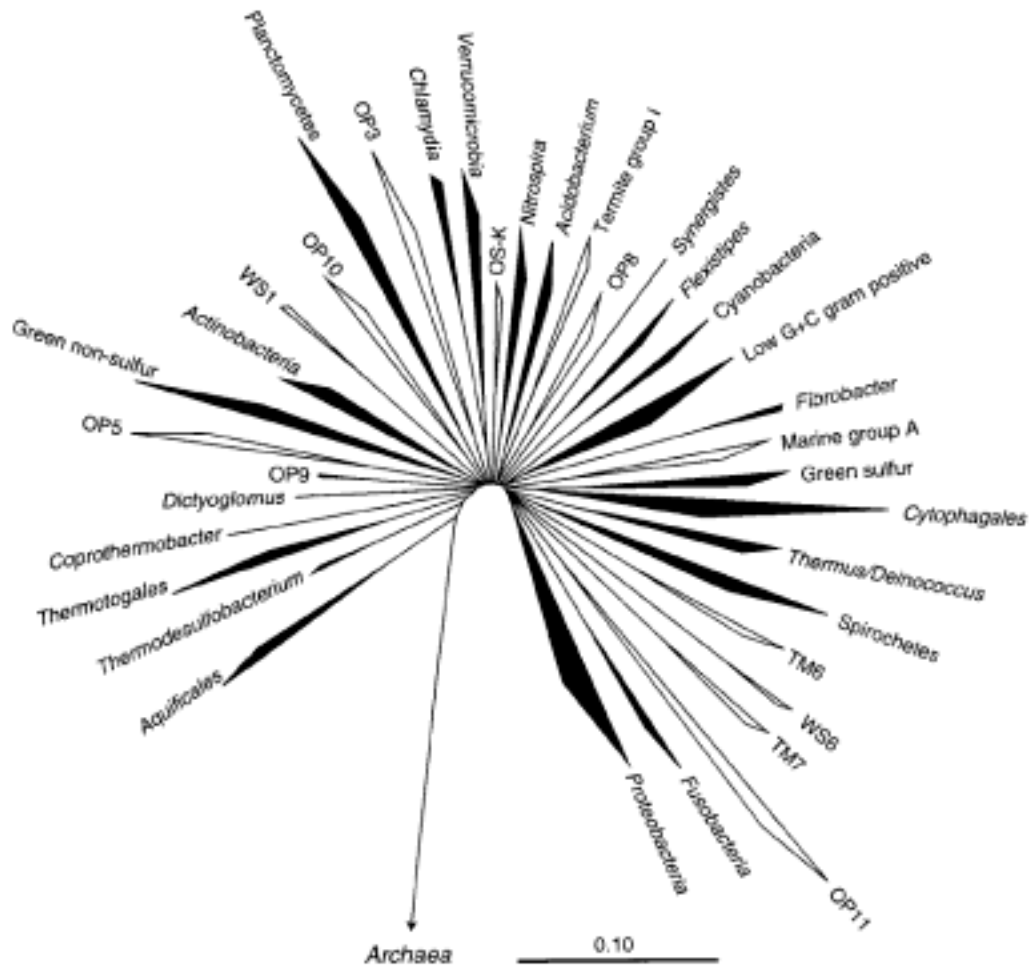


**Figura 1.10.** Model filogenètic de Woese pels eucariotes i els procariotes (1987).

Dins del domini *Bacteria*, segons aquests estudis, trobem dotze divisions o *phyla*: Actinobacteris, Firmicutes, Cianobacteris, Clamídies, Proteobacteries, Espiroquetes, Grup CFB, *Deinococcus/Thermus*, Bacteris Verds dels Sofre i Thermotogals (figura 1.10).

En les comparacions amb el 16S rRNA, es considera que parlem de dos grups diferents, quan aquests presenten unes diferències superiors al 20 o 25% en les seves seqüències, utilitzant diferents fragments o tècniques d'anàlisi dels mateixos (Gupta *et al.*, 1998; Gupta *et al.*, 2000). Per contra, si aquestes no superen aquests percentatges, s'estima que es troben dins el mateix grup. Tot i així, aquesta classificació en dotze divisions del domini *Bacteria*, actualment està essent qüestionada, estimant alguns treballs en més de 40 el nombre de divisions (Hugenholtz *et al.*, 1998) (figura 1.11).





**Figura 1.11.** Il·lustració de l'ampliació de divisions en el domini *Bacteria* proposada per Hugenholtz *et al.* (1998). En negre es mostren les branques corresponents a microorganismes que han pogut ser cultivats en el laboratori, mentre que en blanc es mostren les *divisions candidates*, basades en la seqüenciació de l'rRNA directament de mostres ambientals. Alguns termes apareixen en anglès. Modificat de Hugenholtz *et al.* (1998).

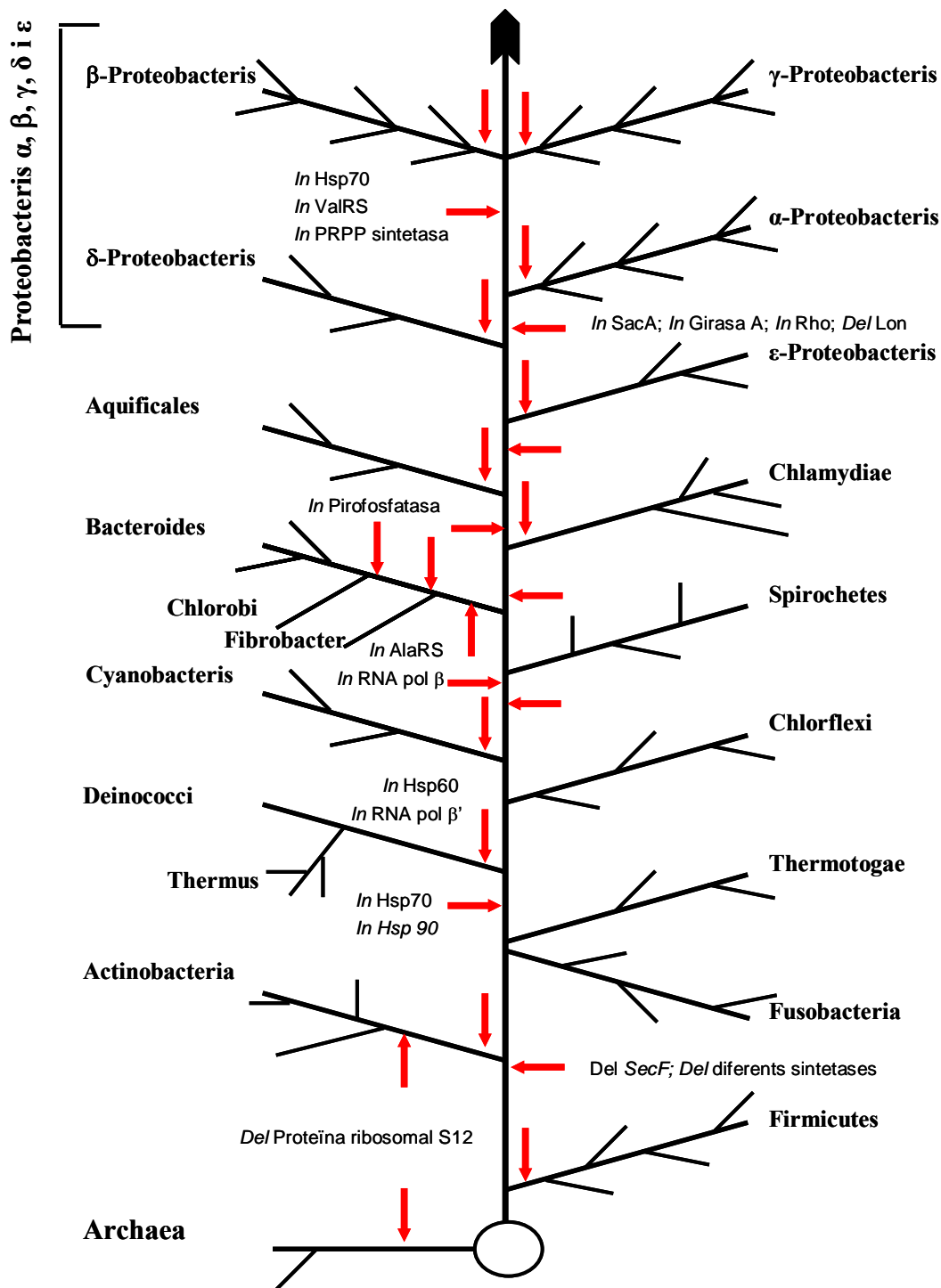
El motiu d'aquest espectacular increment, es deu a l'augment del nombre de genomes disponibles i en disposar de tècniques moleculars més avançades, que han possibilitat una població d'estudi més important. Així, alguns dels grups proposats coneguts com a *divisions candidates*, tant sols se'n coneix la seqüència de l'rRNA, ja que es tracta de mostres ambientals. Tot i així, encara no es pot assegurar el nombre de divisions necessàries, tant per la immensa variabilitat que estem descartant per l'actual incapacitat d'aïllament i estudi de totes les espècies bacterianes, com per no disposar encara d'uns patrons estandarditzats i universals amb els que construir els diferents arbres filogenètics, fet que porta a disparitats de resultats, ja que petits canvis en les matrius usades en els càlculs informàtics poden dur a grans diferències en els resultats obtinguts.

De llavors ençà, s'han realitzat altres estudis filogenètics usant altres elements com eines de rastreig. Així, pròxim a l'exemple exposat, trobem treballs filogenètics basats en tRNA i aminoacil-tRNA-sintetasa (Wolf *et al.*, 1999) o d'altres realitzats per mitjà de la comparació del

sistema de la fosfotransferasa (Barabote i Saier, 2005). La seqüència aminoacídica de les proteïnes LexA i RecA, elements propis del reguló *lexA*, també han estat emprades amb aquest objectiu, en ser elements àmpliament presents i altament importants en la funció cel·lular bacteriana (Eisen *et al.*, 1995; van Waasbergen *et al.*, 2000; Krawczyk *et al.*, 2002). Es necessari remarcar, que en el cas de la seqüència aminoacídica del RecA (Eisen, 1995; Karlin *et al.*, 1995; Karlin i Brocchieri, 1996), els resultats que s'obtenen són pràcticament idèntics als obtinguts amb l'rRNA.

Darrerament, s'està concedint un paper rellevant a la transferència o intercanvi horitzontal d'informació (HTG) (Woese, 2000; Gupta i Griffiths, 2002; Collyn *et al.*, 2005; Novichkov *et al.*, 2004). Aquest corrent, postularia que les actuals espècies bacterianes, representen un calaix de sastre de tots els *inputs* genètics adquirits per transferència vertical o horitzontal, essent aquesta última especialment destacada i important. En relació amb l'HTG, per a poder filar més prim entre els diferents vincles, diversos grups d'investigació han estudiat i seguit el rastre de proteïnes altament conservades, proteïnes *house keeping*. En aquest cas, es fa un seguiment del que s'anomenen seqüències signatura, entenent com a tal, regions de la seqüència primària proteica mantingudes en un o varis grups, i que es diferencien i divergeixen en alineaments amb les seves homòlogues en altres grups filogenètics. Especialment interessants són les possibles insercions i delecions presents en les seves seqüències aminoacídiques, conegudes com a *indels* (*insertion-deletion*) (Gupta *et al.*, 1998; Gupta, 2000) (figura 1.12).

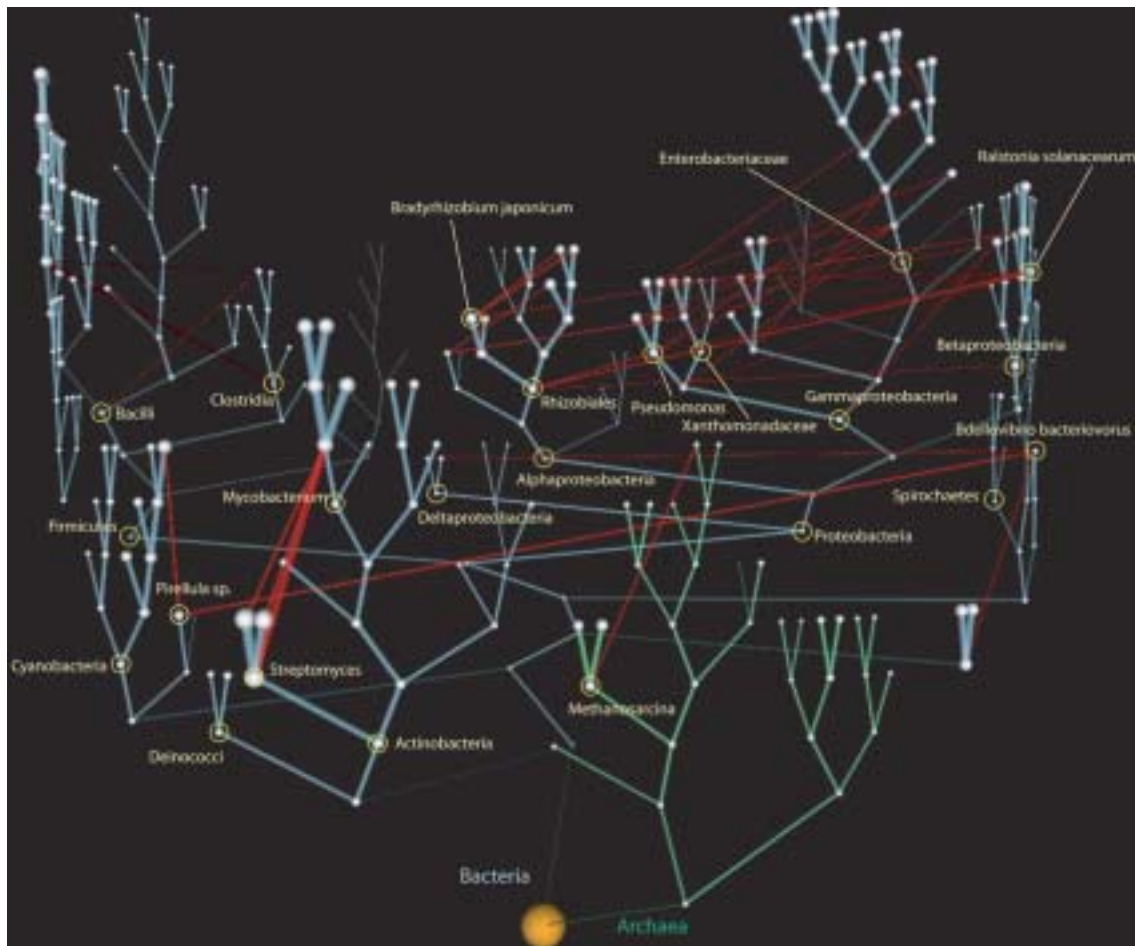
Aquests, són especialment atractius com a eina filogenètica, ja que no es troben relacionats amb mutacions puntuals independents i poden ser considerats com un segell identificatiu dels organismes que les presenten: quan un *indel* es troba de manera conservada i definida en diferents grups, l'explicació més lògica i ajustada, és que es produí la seva introducció de manera discreta en un ancestre comú, essent aquest l'origen dels diferents grups que el presenten actualment (Gupta *et al.*, 1998; Gupta, 2000). Estudis dels *indels* de proteïnes tals com Hsp70, GroEL (Hsp60), GSI (glutamat-1-semialdeid 2,1 aminomutasa), ribosomals i d'altres amb funció anabòlica, en els dominis *Archaea* i *Bacteria*, proven l'estreta relació entre els arquees i els bacteris grampositius (figura 1.12). Aquest fet, junt amb les característiques dels seus embolcalls cel·lulars, conduïren a Gupta i al seu equip de treball a proposar una nova classificació: monoderms (Arqueobacteris i Grampositius) i diderms (Gramnegatius). On els primers són ancestres dels segons i els eucariotes provindrien d'una fusió o quimera d'ambdós grups (Gupta *et al.*, 1998; Rivera i Lake, 2004).



**Figura 1.12.** Model proposat per Gupta *et al.* (1998), respecte l'evolució del món procariota. Modificat de Gupta *et al.* (1998). *In* és l'abreviació de inserció, mentre que *Del*, és la de deleció.

Amb tot això, sembla clar, que la representació gràfica dels camins evolutius que han dut fins l'actual diversitat present en la biosfera, més que un arbre, seria una xarxa (Kunin *et al.*, 2005), on les interaccions filogenètiques s'establirien no tant sols amb el llinatge directe, sinó amb altres grups allunyats. Amb aquest propòsit, Kunin i els seus col·laboradors (Kunin *et al.*, 2005), recentment han realitzat un primera aproximació al que podria ser aquesta xarxa, prestant especial atenció a la transferència genètica horitzontal. Per a tal, s'usaren un grup de gens ortòlegs, en els que s'aplicaren diferents tractaments informàtics, intentant eliminar artefactes

per mitjà de la realització del procés sota diferents criteris tals com el contingut en GC o el *codon usage* (Kunin *et al.*, 2005). Amb aquests estudis, s'aconseguien mostrar els diferents lligams entrecreuats, que indiquen les possibles transferències horitzontals i verticals. Tot i aportar llum respecte aquests tipus d'intercanvis establerts entre diferents grups, una limitació que en la majoria de casos no pot ser resolta, és qui és l'acceptor i qui és el donor (figura 1.13).



**Figura 1.13.** Representació tridimensional de la xarxa filogenètica proposada per Kunin i col·laboradors (2005). L'esquelet de l'arbre ha estat generat usant el percentatge de similitud entre els gens. L'arrel comuna es representa en color groc. Els diferents grups del domini Bacteria estan indicats amb esferes en les branques de color blau cel; els Arquea com nodes en les branques verdes. Les línies vermelles mostren les interrelacions degudes a l'HGT.

## **2. OBJECTIUS**

## 1.4. Objectius

D'acord amb tot l'exposat anteriorment, existeixen ja un nombre elevat d'organismes pertanyents a diferents grups taxonòmics, en els que s'ha descrit el seu reguló *lexA*.

Amb el coneixement de les caixes SOS i quins són els gens que les presenten, podem arribar a hipotetitzar quina ha estat la història evolutiva d'aquests, atenent als possibles canvis i reorganitzacions, i així inferir el fil conductor que ens ha dut a l'actual diversitat bacteriana. Al mateix temps, aquests coneixements ens poden aportar llum de quina és la relació present entre l'hàbitat (junt amb el cicle vital dels microorganismes) i el seu sistema SOS, que esdevé vital per una correcta adaptació i supervivència.

Amb aquest propòsit varem decidir estudiar els microorganismes *Dehalococcoides ethenogenes*, *Magnetococcus sp.* soca MC-1 i *Leptospira interrogans* serovar Lai, ja que per les seves peculiaritats, podien resultar interessants pel nostre propòsit investigador.

Així primerament, s'estudià *D. ethenogenes*, un bacteri anaeròbic capaç de declorar productes tals com el tetracloroetè, que és un dels contaminants aquàtics més comuns. La seva paret cel·lular presenta una composició que no coincideix ni amb la dels grampositius ni amb la dels gramnegatius, sinó que sembla estar més relacionada amb la paret cel·lular dels arqueobacteris. Pel que fa als estudis amb el 16S rRNA,

*D. ethenogenes* ha estat assignat al grup dels bacteris verds no sulfurosos, però totes les seves peculiaritats en tant propietats estructurals, vies metabòliques i la inusual classificació taxonòmica, el convertien en un bacteri interessant des del nostre punt de vista.

Respecte *Magnetococcus sp.* soca MC-1, es tracta d'un microorganisme gramnegatiu microaerofilic, que representava el primer coc magnetotàctic aïllat en cultiu pur. Aquest fet ja el feia mereixedor d'estudi, en tant la informació que ens podia facilitar d'aquest tipus de microorganismes. Filogenèticament, aquest microorganisme es trobava filiat amb la subclasse Alfa dels Proteobacteris, però estudis *in silico* en el seu genoma incomplet, però amb determinats contigs disponibles, no ens va permetre identificar ni la caixa SOS descrita en aquest grup, GTTCN<sub>7</sub>GTTC, ni cap de les descrites fins el moment, fet que junt amb la presència del gen *lexA*, ens feia preveure la presència d'un nou motiu d'unió. Totes aquestes característiques, revelaven interessant el coneixement tant del funcionament del sistema com del conjunt de gens que englobava el reguló.

Per últim, l'interès pel coneixement del reguló *lexA* de *Leptospira interrogans* serovar Lai, atengué a diversos motius. En primer lloc, era el primer microorganisme pertanyent a l'ordre de les espiroquetes, en el que es podia identificar una proteïna hipotètica amb unes característiques que feien plausible la possibilitat de que es tractés del repressor LexA. També és destacable, que en aquest ordre, trobem espècies patògenes d'especial rellevància, com per exemple espècies del

mateix gènere *Leptospira* o del gènere *Treponema*. Així mateix, també és rellevant el paper que se'ls hi atribueix en la teoria endosimbiòtica, que hipotetitzava respecte l'origen de les cèl·lules eucariotes (Margulis, 1996).

*L. interrogans* és un organisme de morfologia marcadament helicoidal, amb mobilitat i aeròbic. Una altra característica a destacar, és que el seu genoma es troba substancialment reduït, fet que el presentava atractiu des de la nostra òptica d'estudi, en relació a les repercussions que tenia aquest fet respecte el funcionament del reguló *lexA*.

En tots tres treballs, el primer objectiu ha estat la identificació i clonació del gen *lexA* en un vector d'expressió, a fi d'obtenir la proteïna pura i amb ella, passar a la detecció per mitjà dels assajos EMSA (electrophoretic mobility shift assay) de la caixa SOS amb una major fiabilitat i comoditat. En cas de detectar-s'hi l'existència d'un motiu d'unió al DNA, aquest ens servirà per a buscar, gràcies a diferents eines informàtiques els diferents gens que es troben sota el control de LexA, arribant així al coneixement, si més no aproximat, de la composició del reguló en aquests microorganismes.

Un cop recapitulades totes les informacions extretes, s'intentarà elucidar quina és la ubicació taxonòmica dels diferents grups dels que han estat escollits com a representants, així com aportar llum a la història evolutiva.

### **3. RESULTATS I DISCUSSIÓ**



## 2. RESULTATS I DISCUSSIÓ

### 2.1. Estudi del reguló *lexA* a *Dehalococcoides ethenogenes*

#### 2.1.1. Identificació del gen *lexA* i purificació del seu producte

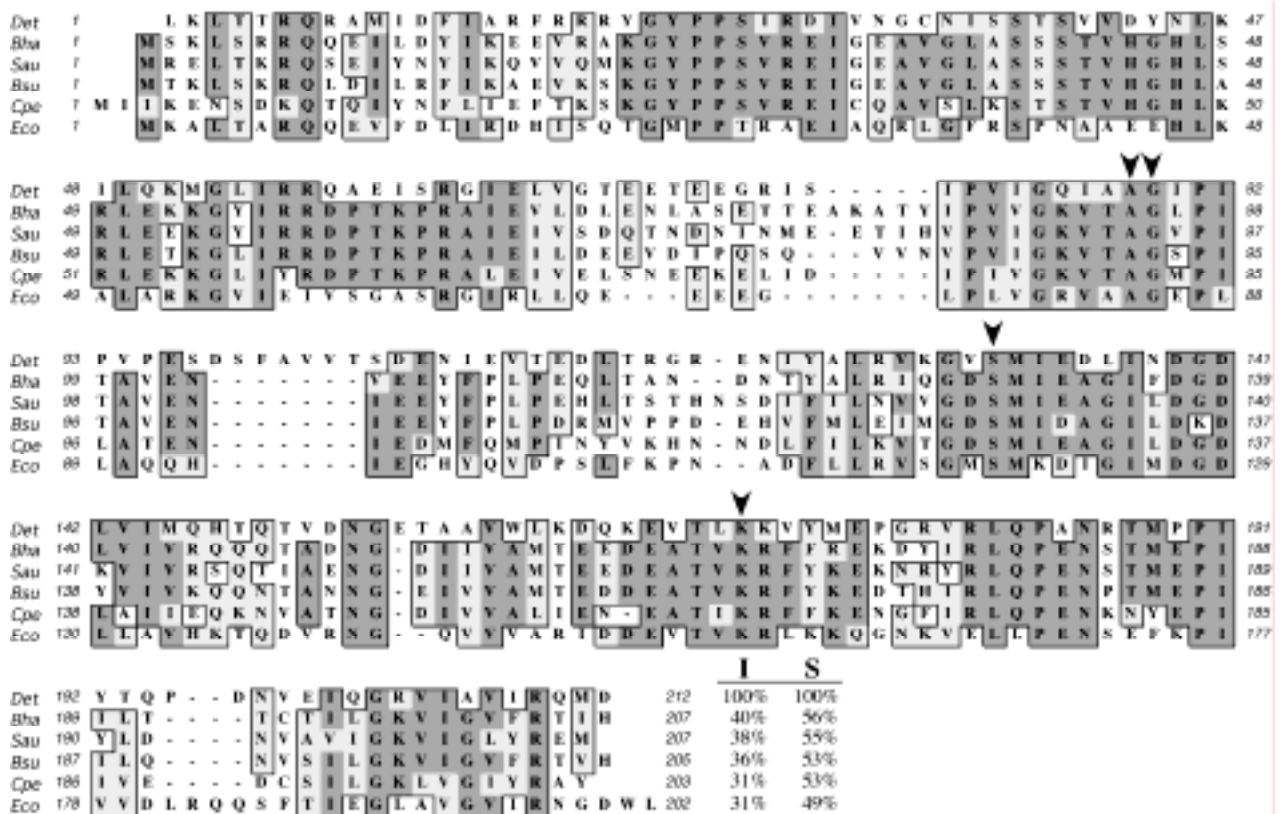
La seqüència aminoacídica de la proteïna LexA d'*E. coli*, fou usada com a subjecte en la cerca per mitjà del programa TBLASTN aplicat en el genoma de *D. ethenogenes* sota 195, que en el moment del treball es trobava sense un assemblatge complet, però que malgrat això, si que hi trobàvem disponibles contigs que ens van permetre realitzar anàlisis del tipus BLAST, en el recurs electrònic del *Institute for Genomic Research and the National Center of Biotechnology Information* ([http://www.ncbi.nlm.nih.gov/Microb\\_blast/unfinishedgenomes.html](http://www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenomes.html)). El programa TBLASTN, ens revelar una proteïna homòloga de 212 aminoàcids.

Tot i que *D. ethenogenes* no és un bacteri grampositiu, amb qui presenta major homologia la seva proteïna LexA és amb la de *Bacillus halodurans*, exactament un 40%, mentre que amb *E. coli* és un 31%. Altre bacteris d'aquest filum amb qui presenta homologia són per exemple *Staphylococcus aureus* i *Clostridium perfringens*.

Dels alineaments de diferents proteïnes LexA per mitjà del programa CLUSTAL W, varem poder deduir que conservava tots els residus importants implicats en el seu autotrencament: Ala<sub>88</sub>, Gly<sub>89</sub>, Ser<sub>131</sub> i Lys<sub>170</sub> (figura 2.1).

L'anàlisi BLAST, ens va remetre 1 kilobase (en endavant kb) de DNA que flanquejava la regió codificant del gen *lexA*, que fou emprada per a dissenyar els primers AR31 i AR34.

De la PCR realitzada amb aquests dos primers, varem obtindre un fragment de 866 parells de bases (en endavant pb) que posteriorment fou clonat en el vector pGEM-T. Un cop assegurat l'èxit de l'amplificació mitjançant seqüenciació, passarem a anomenar aquest plasmidi pUA969. Per a procedir a la purificació de la proteïna, varem dissenyar el primer AR61, que juntament amb el ja esmentat AR34, foren usats per a amplificar de nou el *lexA* de *D. ethenogenes* des del plasmidi pUA969, essent clonat un altre cop en el pGEM-T, generant el pUA970. L'oligonucleòtid AR61, presenta en el seu extrem 5' (que coincideix amb el 5' d'aquest fragment portador del gen *lexA*), la diana de restricció *EcoRI*, fet que permet la fusió en fase amb la proteïna GST, que es troba codificada en la regió immediatament anterior al punt de clonació en el vector d'expressió pGEX-4T1 (Amersham Pharmacia Biotech).



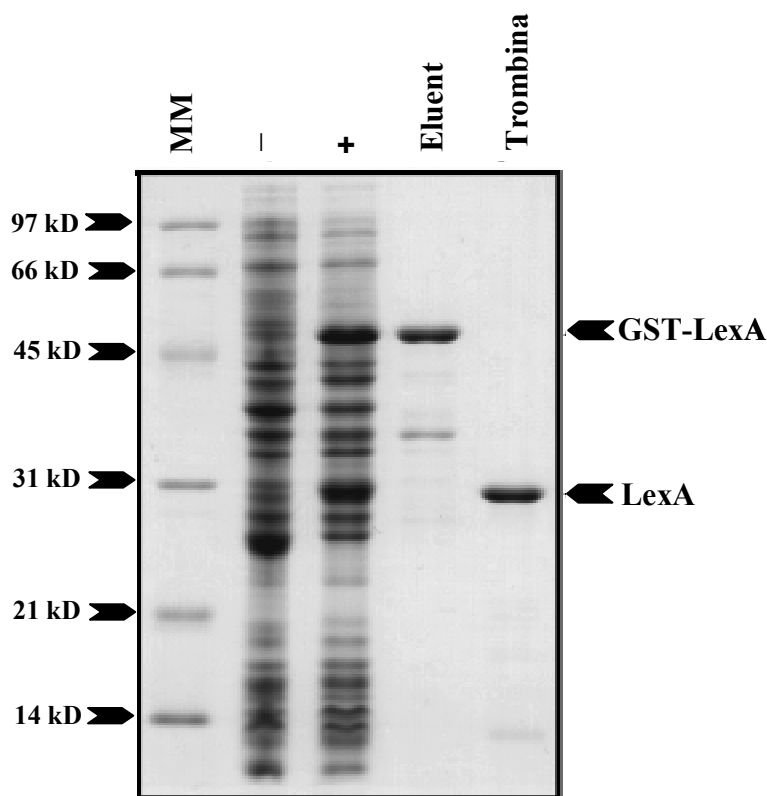
**Figura 2.1.** Alineament CLUSTAL W realitzat amb el Mac Vector 6.5 (Oxford Molecular) usant les proteïnes LexA de *D. ethenogenes* (Det), *B. halodurans* (Bha), *St. aureus* (Sau), *B. subtilis* (Bsu), *C. perfringens* (Cpe) i *E. coli* (Eco). Els residus idèntics presenten un fons gris fosc, mentre que les similituds es mostren en gris clar. Els residus Ala<sub>88</sub> i Gly<sub>89</sub> implicats en el trencament de la proteïna, i els Ser<sub>131</sub> i Lys<sub>170</sub> necessaris per la hidròlisi, estan indicats amb puntes de fletxa en la seqüència del LexA de *D. ethenogenes*. A continuació de l'últim residu de cada seqüència aminoacídica, apareix el percentatge d'identitat i de similitud respecte el LexA de *D. ethenogenes*. Els números d'accés en la base de dades proteica Entrez, són els següents: *B. halodurans*, Q9KAD3; *S. aureus*, BAB57501; *B. subtilis*, P31080; *C. perfringens*, BAB80867 i *E. coli*, P03033.

Del plasmidi pUA970, fou extret un fragment de 0,65 kb, mitjançant la digestió amb els enzims *EcoRI* i *SalI*, que fou clonat en el pGEX-4T1, obtenint el plasmidi pUA971.

El pUA971, fou transformat per xoc tèrmic en cèl·lules competents d'*E. coli* BL21 Codon Plus (Stratagene) preparades amb CaCl<sub>2</sub>, per a que durant el seu creixement post inducció acumulessin grans quantitats de proteïna LexA gràcies a la sobreexpressió facilitada per l'anterior clonació en el vector d'expressió pGEX-4T1.

La proteïna GST, presenta en la seva estructura primària la següent seqüència aminoacídica: Leu-Val-Pro-Arg-Gly-Ser. Aquesta seqüència es troba ubicada en la regió entre la proteïna GST i la proteïna fusionada en qüestió, i es reconeguda com a diana d'actuació per part de la proteasa trombina, que actua sobre l'enllaç Arg-Gly. Com a conseqüència de sotmetre l'eluent obtingut en la sobreexpressió de la proteïna LexA a l'acció de la trombina, varem poder disposar de la

proteïna LexA de *D. ethenogenes* pura en un 95%, tot i que amb els residus Gly-Ser-Pro-Glu-Phe lligats al seu extrem N-terminal (figura 2.2).



**Figura 2.2.** Purificació de la proteïna LexA de *D. ethenogenes*. S'obtingué la proteïna amb un percentatge de puresa superior al 95 %. En la imatge del gel desnaturalitzant de poliacrilamida al 13 %, cada carril mostra els diferents passos en el procés de purificació: marcador molecular, carril MM; extracte cru de BL21 Codon Plus/PUA971 (-); extracte cru de BL21 Codon Plus/pUA971 induït amb 1 mM de IPTG (+); *D. ethenogenes* GST-LexA purificat per afinitat amb el *glutathione-Sepharose 4B beads*, eluent; proteïna LexA de *D. ethenogenes* posterior a la digestió proteàsica de la trombina, trombina.

### 2.1.2. Caracterització del motiu d'unió al DNA de la proteïna LexA

Estudis realitzats en el contig que contenia el *lexA* de *D. ethenogenes*, ens van mostrar que la seva ubicació estava entre els gens *rpsT* i el *fucA* (figura 2.3A). Des del codó de finalització de traducció del gen *rpsT*, fins a l'hipotètic codó d'inici de traducció del *lexA*, trobem 104 pb on presumiblement es trobava el motiu d'unió de la proteïna LexA al DNA.

Mitjançant PCR i usant els primers AR31 i AR33DIG, varem obtenir un fragment marcat amb digoxigenina, que fou usat amb posterioritat per a detectar si la proteïna LexA purificada tenia capacitat d'unir-se a aquesta regió, fet que es pretenia comprovar mitjançant assaigs de mobilitat electroforètica (EMSA).

Els assaigs EMSA realitzats emprant el fragment anterior, en endavant LexA1, corroboraren la hipòtesis inicial, és a dir, que la regió reguladora es trobava en els 104 pb immediatament anteriors al gen *lexA*. Dues proves posteriors, ens confirmaren que els resultats inicials eren fidedignes, aquestes foren 1) l'aparició de la banda retardada d'unió, es veia abolida per la presència de quantitats molt més grans del mateix fragment no marcat amb digoxigenina, és a

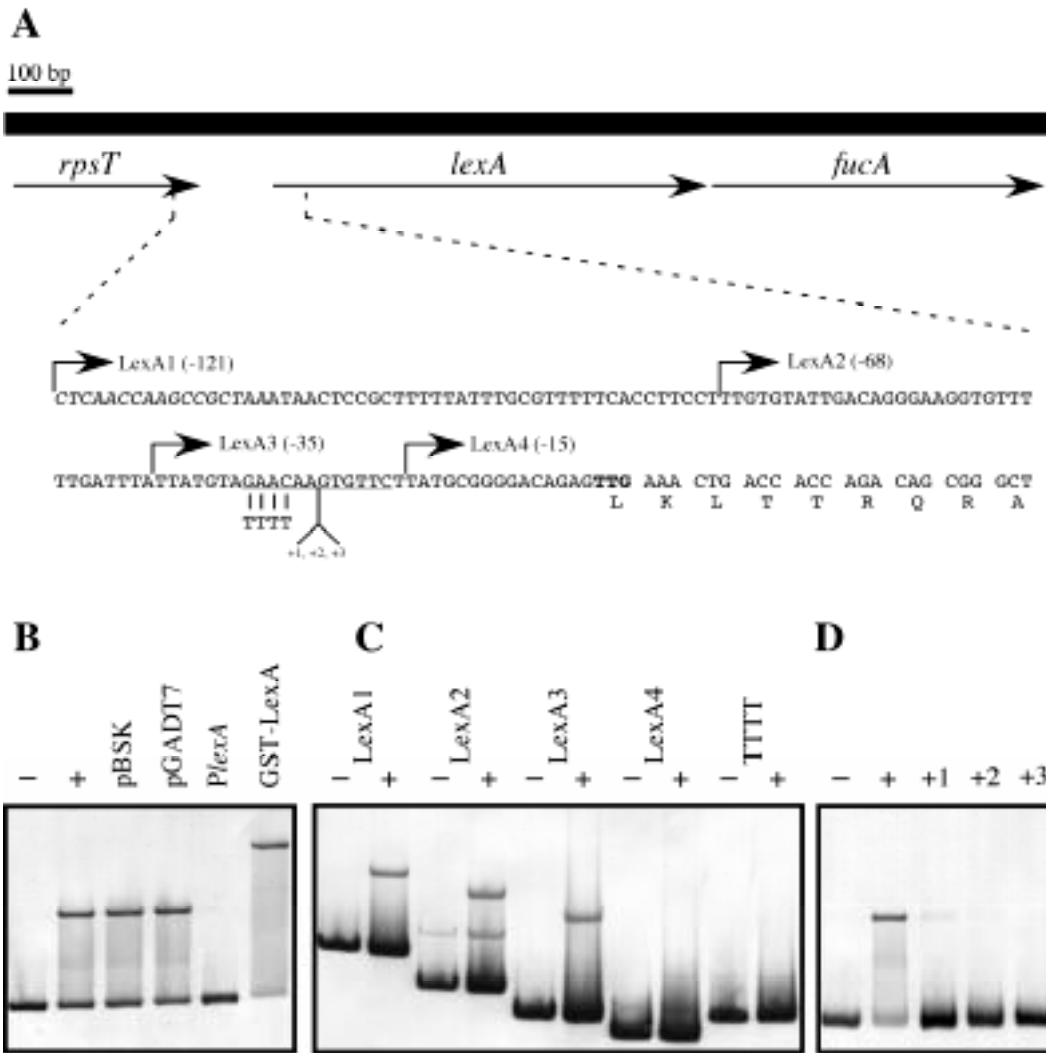
dir, actuant com a competidor específic; 2) l'esmentada unió no es veia alterada per l'addició de DNA no específic, mostrant-se per tant, específica la unió (figura 2.3B).

Amb la finalitat d'acotar més acuradament el punt d'unió del LexA de *D. ethenogenes*, es varen dissenyar primers a fi d'obtenir fragments successivament menors en longitud en el seu extrem 5'. El darrer d'ells que mostrar unió fou el LexA3, ja que el fragment LexA4, tot i que es repetí la prova tres cops, no mostrava banda de retard (figura 2.3C). Aquest fet implicava que la seqüència reconeguda per LexA, es trobava a menys de 35 parells de bases del codó d'inici de traducció del gen *lexA*. Una examinació visual, va desvelar la presència d'una seqüència potencial a -28 pb del codó TTG: GAACAAGTG TTC (figura 2.3A). Aquesta seqüència, coincideix substancialment amb la reconeguda per DinR (*Cheo box*): CGAACN<sub>4</sub>GTTCG (Winterling *et al.*, 1998).

Per mitjà de mutagènesi dirigida mitjanant PCR, es substituï el submotiu GAAC per TTTT en el fragment LexA3: aquest canvi, eliminava la unió del LexA de *D. ethenogenes*, fet que demostrava la pertinença d'aquestes 4 bases a la seva caixa SOS (figura 2.3C). La importància de la orientació rotacional relativa de les repeticions inverses que componen cada un dels submotius, es va analitzar mitjançant la inserció de determinats nucleòtids en la regió connectora de la seqüència GAACN<sub>4</sub>GTTC. Mitjançant diferents primers, s'introduïren d'una a tres adenines en la posició -21 en relació al codó d'inici de traducció (figura 2.3A). En el cas en que tant sols s'incrementava en 1 nucleòtid la longitud de la regió variable, es mantenia certa unió de la proteïna LexA en els assaigs EMSA (figura 2.3D). No obstant, els fragments que presentaven dues i tres adenines afegides no originaren cap unió (figura 2.3D).

Indagant amb la finalitat de conèixer profundament la importància de cada un dels residus nucleotídics en la unió del LexA a la caixa SOS de *D. ethenogenes*, així com el de les bases flanquejants de l'hipotètic motiu, varem passar a obtenir nous fragments derivats del salvatge per mitjà de mutagènesi dirigida. Per a tal fi, es dissenyaren diferents primers (del AR41 al AR57 en combinació amb el AR76) que presentaven els canvis menys permissius possibles, atenent els estudis per a la caracterització de la caixa de DinR duts a terme per Winterling (Winterling *et al.*, 1998). Les mutacions que afectaven als dos submotius palindròmics, disminuïen críticament o eliminaven la unió del LexA (figura 2.4A). Així mateix, també es revelaren importants les bases que acompanyen aquestes repeticions inverses: els residus d'adenina de les posicions -29 i -24, així com els residus de timina presents en les posicions -16 i -21. Tots aquests resultats, van confirmar els obtinguts en treballs previs (Winterling *et al.*, 1998), és a dir, que els segments GAAC i GTTC, són crítics per la unió.

Així doncs, la seqüència específica d'unió pel repressor LexA de *D. ethenogenes*, respòn al motiu: AGAACN<sub>4</sub>GTTC.



**Figura 2.3.** A) Representació esquemàtica de la regió cromosòmica de *D. ethenogenes* que conté el gen *lexA*. Tal com es pot observar, es troba ubicat entre els gens *rpsT* i *fucA*. També es mostren la seqüència de la regió promotora de *lexA* i els 9 primers aminoàcids de la proteïna. L'extrem 3' del gen *rpsT* es troba indicat en lletra cursiva, l'hipotètic codó d'inici de la traducció del LexA es troba en negreta i la seqüència homòloga de la caixa DinR es troba subratllada. Els inicis de cada un dels fragments emprats LexA1, LexA2, LexA3 i LexA4, es troben indicats per fletxes, mentre que les seves distàncies relatives al TTG del LexA es troben en parèntesis. B) Unió específica de la proteïna LexA de *D. ethenogenes*. Els assaigs EMSA es realitzaren utilitzant diferents DNAs com a competidors. El fragment LexA1 fou incubat a 30°C en absència (-) o presència (+) de 25 ng de proteïna LexA pura (50 mM). Per a mostrar l'especificitat de la unió del LexA, el fragment LexA1 fou incubat en presència de la proteïna LexA junt amb els següents DNAs competidors no marcats amb DIG: 2 µg de pBSK, 2 µg de pGADT7 i 1 µg de fragment LexA1 no marcat (*PlexA*). En l'últim carril es va incubar el fragment LexA1 amb la fusió GST-LexA de *D. ethenogenes*. C) Delimitació de la localització del punt d'unió de la proteïna LexA de *D. ethenogenes* a la seva regió promotora. LexA1, LexA2, LexA3 i LexA4 i el fragment derivat del LexA3 on el motiu GAAC fou substituït per TTTT (TTTT) foren incubats en presència (+) i absència (-) de proteïna LexA purificada. D) Imatge dels resultats obtinguts amb els assaigs EMSA realitzats amb el fragment LexA3 on s'havia afegit una (+1), dues (+2) o tres (+3) adenines respectivament en la posició -21 respecte el codó d'inici de traducció proposat. Diferents proves foren realitzades i carregades en una gel nadiu de poliacrilamida Tris-Glicina-5%. Les reaccions amb el fragment LexA3, incubat en absència (-) i presència del LexA (+), varen actuar com a controls negatiu i positiu respectivament.

### 2.1.3. El reguló *lexA* a *D. ethenogenes*

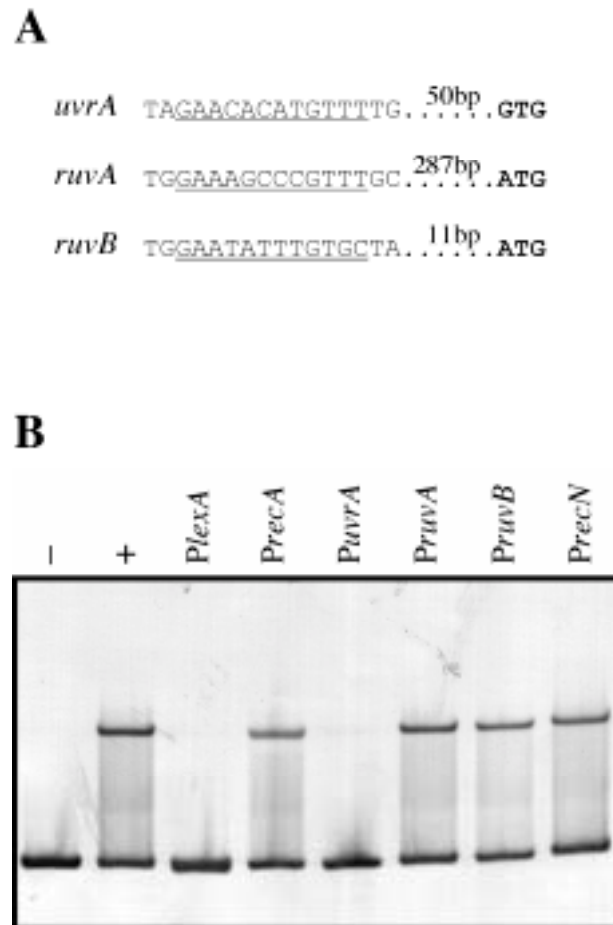
Els assaigs EMSA realitzats amb els fragments mutats mostraren la importància de les bases A i T en les posicions 1 i 14 en relació a la seqüència AGAACN<sub>4</sub>GTTCT (figura 1.4A). Malgrat això, la seqüència de la caixa SOS descrita en *B. subtilis* i altres bacteris grampositius és CGAACN<sub>4</sub>GTTTCG (Winterling *et al.*, 1998). Precisant més, els dos nucleòtids que ocupen les esmentades posicions, tant sols han estat detectats en un nombre baix de grampositius, especialment la timina en posició 14, tant sols observada en una de vint caixes: en l'operador del gen *recA* de *Streptomyces lividans* (Winterling *et al.*, 1998).

Amb la motivació de confirmar que la caixa descrita en *D. ethenogenes* és la mateixa o lleugerament derivada de la dels bacteris grampositius, es va procedir a analitzar la capacitat d'unió de la proteïna DinR a la seqüència en qüestió. Així, el fragment LexA1 marcat amb digoxigenina fou incubat junt amb la proteïna DinR de *B. subtilis* i el LexA de *D. ethenogenes*, en absència i presència de DNA competidor, és a dir, fragments que presentaven o bé la caixa DinR o bé la caixa LexA en qüestió. Els resultats de les proves indicaven que DinR (figura 2.4B) s'unia a la caixa de *D. ethenogenes* i que aquesta unió, era abolida per la presència de qualssevol dels dos DNAs com a competidors. Aquests resultats ens permeteren asseverar que la proteïna LexA de *D. ethenogenes* reconeix específicament el motiu GAACN<sub>4</sub>GTTTC, tal com ho fa DinR.

En *E. coli*, almenys 40 gens estan sota control negatiu directe del repressor LexA (Courcelle *et al.*, 2001; Fernández de Henestrosa *et al.*, 2000). Les seqüències aminoacídiques dels productes de cada un d'aquests gens pertanyents al seu reguló *lexA*, foren usats com a subjecte en les cerques realitzades per mitjà del programa TBLASTN en el genoma incomplet del bacteri objecte d'estudi. Homòlegs del gen *recA*, *uvrA*, *ruvA*, *ruvB*, *uvrD*, *ssb*, *dinB* i *recN* foren identificats, i els seus promotors estudiats, a fi de trobar punts potencials d'unió del LexA a *D. ethenogenes*. En els que no es va trobar homòleg, el motiu pot ser degut a que el seu genoma no es troba completament seqüenciat o potser a algun reajustament genòmic validat en presentar una major adaptació a l'habitat.



A una distància relativa de 64 pb respecte el possible codó d'inici de traducció GTG del gen *uvrA*, varem trobar un motiu molt similar al que apareixia en el gen *lexA* (figura 2.5A). Altrament, també es varen trobar possibles motius d'unió en les regions localitzades davant del gens *ruvA* i *ruvB*: a 301 i 25 pb del codó d'inici de traducció respectivament (figura 2.5A). L'estudi de la resta de gens va ser infructuosa, fet que implicava la no regulació del gen *recA* per part del LexA.



**Figura 2.5.** A) Seqüències anteriors als gens *uvrA*, *ruvA* i *ruvB* de *D. ethenogenes* que hipotèticament presenten caixa d'unió del LexA. Les regions potencials d'unió del LexA es mostren subratllades i els hipotètics inicis de traducció de cada un dels gens es presenten en lletra negreta, junt amb la distància relativa respecte les caixes. B). Resultat d'un assaig EMSA realitzat amb el fragment LexA1 incubat amb proteïna LexA de *D. ethenogenes* en presència de diferents DNAs competidors. El fragment LexA1 fou incubat a 30°C durant 30 minuts en absència (-) i presència (+) de proteïna pura. En les reaccions restants, el fragment LexA1 i la proteïna pura foren incubats junt amb el fragment resultant de l'amplificació de la regió promotora de cada un del gens *lexA*, *recA*, *uvrA*, *ruvA*, *ruvB* i *recN*.

Per a tal de comprovar si aquestes caixes detectades eren funcionals, es va procedir a l'obtenció per PCR dels fragments que contenien les regions operadores dels gens *uvrA*, *ruvA*, *ruvB*, *recN* i



*recA*. Aquests foren usats com a competidors en els assaigs amb el fragment LexA1 marcat amb DIG (figura 2.5B), essent emprats en concentracions 100 cops superiors a les del fragment marcat. En la reacció que contenia el fragment corresponent al gen *uvrA*, no es detectava banda retardada d'unió, no així en les que empràrem els fragments dels gens *ruvA*, *ruvB*, *recN* i *recA* (figura 2.5B). Aquest fet indicava la capacitat de la proteïna LexA per a unir-se a la seqüència present en el fragment *uvrA*: AGAACACATGTTTT. L'apreciació visual de l'absència d'un motiu amb cert grau d'homologia en el promotor del gen *recA*, fou ratificada per la incapacitat que aquest presentava de desplaçar la unió amb el fragment LexA1, confirmant així la teoria de la no regulació negativa, almenys directa, de *recA* per la proteïna LexA. Aquest fet ja havia estat descrit en *Deinococcus radiodurans*, en el que el gen *recA* és induït per radiació gamma de manera independent del LexA. Tanmateix, s'ha suggerit la presència d'un segon gen *lexA* en el genoma d'aquest microorganisme, que potser podria regular directament *recA* (Narumi *et al.*, 2001). Els nostres resultats evidenciaren la no regulació del LexA sobre *recA*, esdevenint així la primera nova, sense possibles alternatives, d'aquest fet.

En base a la informació inferida de les seqüències de diverses proteïnes, recentment s'havia postulat que el grup dels bacteris verds no sulfurosos, al qual pertany *D. ethenogenes*, havia evolucionat d'un ancestre universal posterior als bacteris grampositius, però anterior als cianobacteris (Gupta *et al.*, 2000). Sigui com sigui, sembla clar que existeix una relació filogenètica estreta entre aquest grup de microorganismes i els bacteris grampositius, filiació demostrada per la unió del DinR a la caixa aquí descrita. Aquest fet però, no implica que no existeixin diferències rellevants en el sistema SOS en global. Així per exemple, per mitjà de cerques BLAST *in silico* en el GenBank, es varen poder observar divergències en la composició del reguló: el nombre de gens que presenten amunt del seu codó d'inici de transcripció una seqüència d'unió del LexA és més elevat en organismes tals com *B. subtilis*, *S. aureus*, *Mycobacterium tuberculosis* o *Clostridium acetobutylicum* que en *D. ethenogenes* (figura 2.6). Els motius d'aquest fet poden ser variats, però amb tota seguretat, trobaríem la necessitat d'adaptació i per tant de requeriments, en tant els diferents nínxols ecològics que ocupen cada una de les espècies, junt amb el ja esmentat de la no disponibilitat de la totalitat del genoma.

Per altra banda, derivat dels estudis realitzats amb les seqüències aminoacídiques anteriorment esmentades, semblaria que el grup *Deinococcus-Thermus* (Gupta *et al.*, 2000) hauria aparegut de forma posterior als bacteris verds no sulfurosos, tot i que aquest hipotètic ordre evolutiu no és pot corroborar des del nostre prisma, en no disposar d'informació del motiu d'unió del LexA dels primers.

Intentant obtenir més informació del sistema SOS en els bacteris verds no sulfurosos, es van realitzar anàlisis TBLASTN, usant el LexA de *D. ethenogenes*, en el U.S Department of Energy (DOE) i en el Joint Genome Institute (JGI), que ens mostraren la presència d'una proteïna homòloga a *Chloroflexus aurianticus*, bacteri fotosintètic que també pertany al grup dels

bacteris verds no sulfurosos. Aquesta hipotètica proteïna, es trobava a l'extrem 5' d'un petit contig (694 pb), amb el que l'extrem N-terminal de la mateixa, i per tant, la potencial caixa SOS, no va poder ser identificada. Tanmateix, varem poder observar que la proteïna LexA de *C. aurianticus*, presenta el motiu hèlix-gir-hèlix implicat en la unió al DNA (Fogh *et al.*, 1994), mostrant un 87% d'homologia amb el mateix motiu de la proteïna LexA de *D. ethenogenes*, fet que permetria hipotetitzar respecte a la possible capacitat d'unió del LexA de *C. aurianticus* al motiu GAACN<sub>4</sub>GTTC, amb un comportament per tant, també semblant al dels bacteris grampositius. Un cop es pugui disposar de la totalitat de la seqüència del genoma de *C. aurianticus*, aquest pot esdevenir una eina valuosa per a comprendre la història evolutiva dels bacteris verds no sulfurosos, en tant la homogeneïtat d'aquest filum pel que fa al motiu de reconeixement de la proteïna LexA. Altrament, també podria ser útil aquesta informació per a establir un patró de relació entre les condicions ambientals derivades del nínxol ecològic, i els gens que estan sota el control de la proteïna LexA en els bacteris.

Gen	<i>B. subtilis</i>	<i>S. aureus</i>	<i>M. tuberculosis</i>	<i>C. acetobutylicum</i>	<i>D. ethenogenes</i>	<i>E. coli</i> <sup>a</sup>
<i>recA</i>	+	+	+	+	-	+
<i>lexA</i>	+	+	+	+	+	+
<i>ruvAB</i>	+	-	+	-	- <sup>b</sup>	+
<i>uvrA</i>	-	-	-	-	+	+
<i>uvrB</i>	+	-	-	+	-	+
<i>uvrD</i>	+	-	-	-	-	+
<i>ssb</i>	-	+	-	-	-	+
<i>recN</i>	-	-	-	-	-	+

**Figura 2.6.** Comparació de la presència de caixa SOS davant de diversos gens que hipotèticament formen part del reguló *lexA* de diferents bacteris grampositius, *D. ethenogenes* i *E. coli*<sup>a</sup>

<sup>a</sup> La presència o absència del motiu d'unió del LexA es representa per un signe (+) o (-) respectivament. Les cerques en els genomes s'han realitzat mitjançant el programa RCGScanner, usant un mètode redundat [Rodionov, 2001 #951], que permet la identificació dels nucleòtids escollits així com dels marcs oberts de lectura pròxims [Erill, 2002 #952].

<sup>b</sup> Els gens *ruvA* i *ruvB* no formen un operó a *D. ethenogenes*.

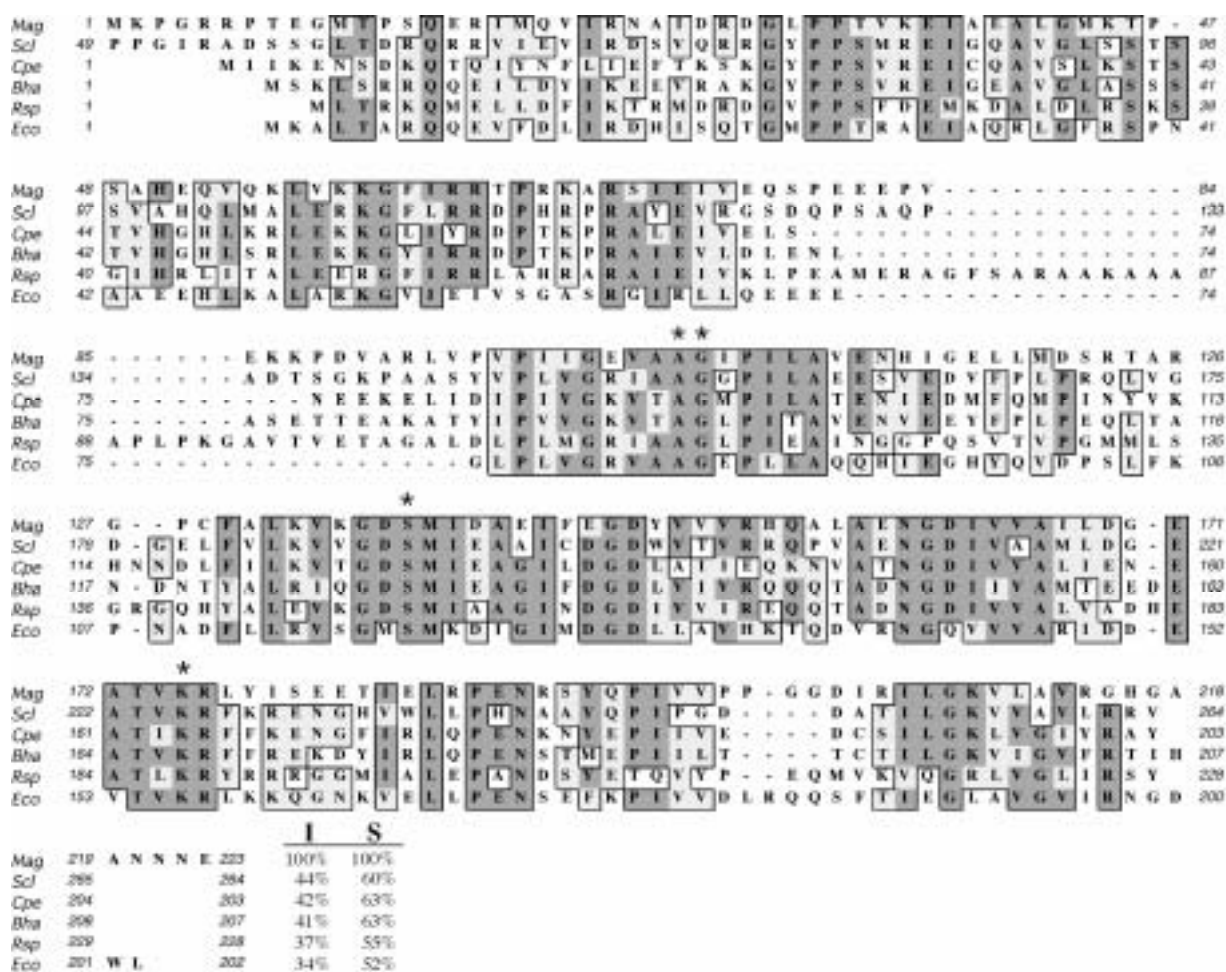
## 2.2. Estudi del sistema SOS a *Magnetococcus sp* soca MC-1

### 2.2.1. Identificació del gen *lexA* i purificació del seu producte

Tot i que el genoma d'MC-1 no es trobava completament seqüenciat, gràcies als contigs disponibles en el DOE i el JGI ([http://www.jgi.doe.gov/JGI\\_microbial/html/magnetococcus/magneto\\_homepage.html](http://www.jgi.doe.gov/JGI_microbial/html/magnetococcus/magneto_homepage.html)) i el National Center of Biotechnology Information (NCBI) ([http://www.ncbi.nlm.nih.gov/Micro\\_blast/unfinishedgenome.html](http://www.ncbi.nlm.nih.gov/Micro_blast/unfinishedgenome.html)), varem poder iniciar aquest treball. Concretament es va obtenir la seqüència del gen *lexA* d'MC-1, realitzant una cerca TBLASTN, essent emprada com a subjecte la proteïna LexA d'*E. coli*. El treball *in silico*, ens reportar una regió que presentava una important homologia amb una proteïna present en el contig 371, suggerint la presència d'un gen *lexA* de 672 pb. El seu producte constava de 223 residus aminoacídics, amb un pes molecular de 24,453 kDa. Aquesta proteïna mostrava les característiques diferencials dels repressors LexA: un domini N-terminal que mantenia el domini hèlix-gir-hèlix (HTH) d'unió al DNA i un de C-terminal amb activitat serin proteasa.

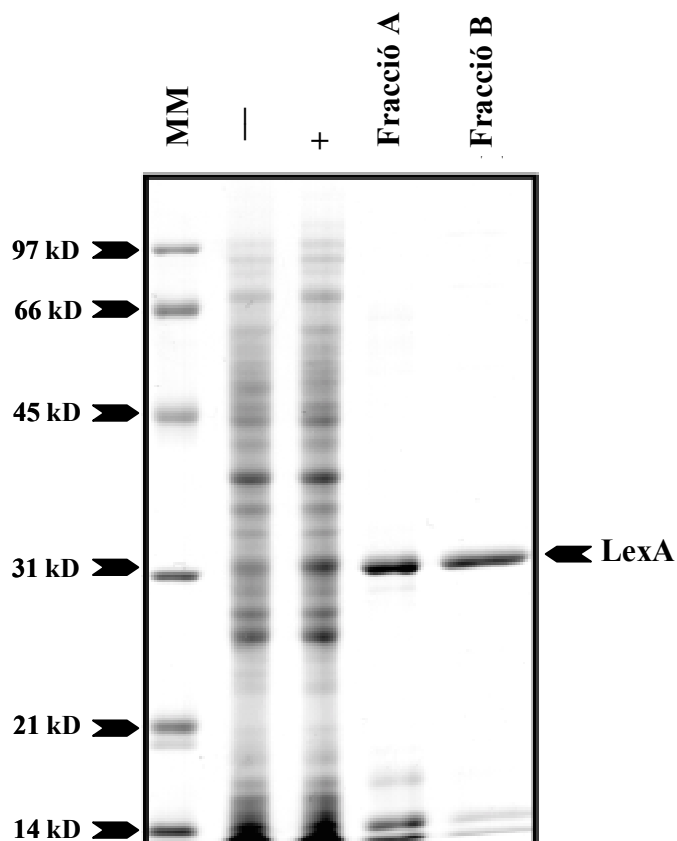
Comparant la seqüència proteica amb la d'altres LexA bacterians disponibles, varem poder observar que amb el que presentava una major homologia era el de *S. clavuligerus* (44%), mentre que amb el de *R. sphaeroides* i *E. coli* mostrava un 37 i 34% d'analogia, respectivament (figura 2.7). Gràcies als resultats obtinguts amb l'eina informàtica d'alineament CLUSTAL W, observarem que el LexA d'MC-1 mantenia els residus importants implicats en el seu autotrencament (Ala<sub>104</sub>, Gly<sub>105</sub>, Ser<sub>138</sub> i Lys<sub>175</sub>) (figura 2.7). Tot i això, importants diferències eren apreciades quan es passava a la realització d'un modelatge proteic en 3D. Per exemple, la segona hèlix del motiu hèlix-gir-hèlix, és de menor grandària que a *E. coli*, i en la regió connectora entre el domini N-terminal d'unió i el C-terminal (*hinge region*), hi trobem 14 nucleòtids addicionals formant una fulla beta no present a *E. coli* (Luo *et al.*, 2001).

Abans de procedir a l'amplificació per PCR del gen *lexA*, mitjançant cerques BLAST varem obtenir les seqüències flanquejants d'aproximadament 1 kb de longitud en total. Amb aquesta informació, es varen dissenyar els primers JC5 i JC6, per a procedir a la realització de la PCR corresponent. Aquests primers, amb vistes a processos posteriors, incorporaven les dianes *NdeI* i *BamHI* respectivament, i un cop realitzada l'amplificació, es procedir al seu clonatge en el vector pGEM-T, generant el plasmidi pUA1011.



**Figura 2.7.** Alineament CLUSTAL W realitzat amb el Mac Vector (versió 6.5; Oxford Molecular) per a la comparació de les proteïnes de la soca MC-1 (Mag), *S. clavuligerus* (Scl) (dels residus 49 al 264), *C. perfringens* (Cpe), *B. halodurans* (Bha), *R. sphaeroides* (Rsp) i *E. coli* (Eco). Els residus idèntics presenten un ombrejat més fosc que els que són similars. I i S, són els percentatges d'identitat i semblança que cada una de les seqüències presenta amb el repressor de la soca MC-1. Per a una millor visualització de la figura, els primers 49 residus del LexA de *S. clavuligerus*, no estan inclosos en l'alineament perquè cap d'ells presenta coincidències en la comparació. Les estrelles indiquen l'Ala, Gly, Ser i Lys implicades en l'autotrencament del LexA. Els codis d'accés dels genomes en la base de dades proteica Entrez del NCBI són els següents: *S. clavuligerus*, CAA12169; *C. perfringens*, BAB80867; *B. halodurans*, Q9KAD3; *R. Sphaeroides*, Q9ZFA4; *E. coli*, P03033.

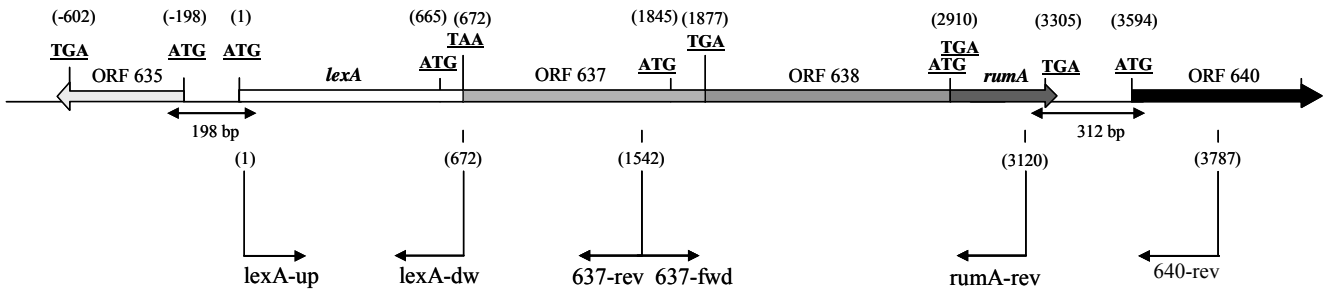
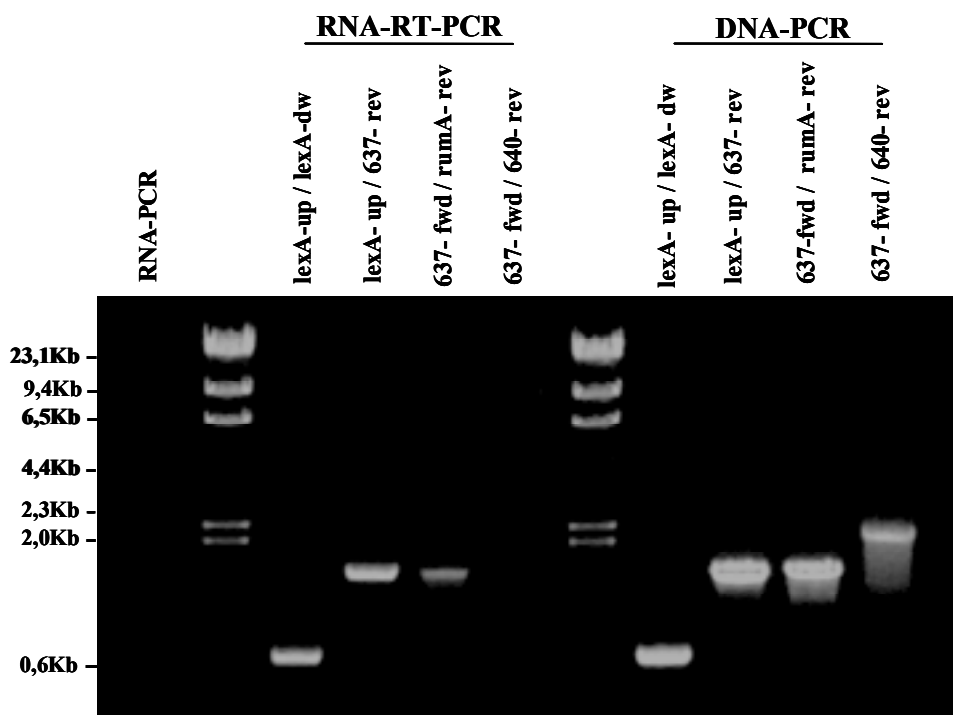
El plasmidi obtingut anteriorment, pUA1011, va ser digerit amb els enzims de restricció *Nde*I i *Bam*HI, i el fragment de 0,7 kb alliberat va ser clonat en el vector d'expressió pET15b (Novagene), generant el pUA1013, a fi de poder obtenir quantitats importants de LexA. El resultat del procés de purificació fou la proteïna LexA d'MC-1 amb un 80 i 95% de puresa respectivament en cada una de les esmentades fraccions, basant-nos en els gels d' SDS-PAGE al 13% (figura 2.8).



**Figura 2.8.** Purificació de la proteïna LexA d'MC-1. La fusió His-LexA s'obtingué en un percentatge de puresa superior al 90%, per mitjà de l'afinitat cromatogràfica d'aquesta pel  $\text{Co}^{2+}$ , usant el kit de purificació TALON<sup>®</sup>. Els carrils del gel desnaturalitzant de poliacrilamida al 13%, representen cada un dels passos realitzats: MW, marcador molecular; -, extracte cru de BL21 Codon Plus/pUA1013; +, extracte cru de BL21 Codon Plus/pUA1013 induït amb 1 mM d'IPTG; fracció A, MC-1 His-LexA eluït amb 150 mM d'imidazol; fracció B, MC-1 His-LexA després de l'elució amb 200 mM d'imidazol.

### 2.2.2. Anàlisi transcripcional del gen *lexA* i caracterització de la caixa LexA

En la direcció transcripcional contrària a la del gen *lexA*, exactament a 198 pb corrent amunt, detectarem un marc obert de lectura que hipotèticament codifava una proteïna de funció desconeguda i formada per 406 residus aminoacídics (figura 2.9A). A continuació del gen *lexA*, però en aquest cas amb la mateixa direcció transcripcional (figura 2.9A), varem apreciar la presència de tres marcs oberts de lectura: dels dos primers no en varem poder determinar la funció, mentre que el tercer d'ells, mostrava un homologia del 58% amb el producte proteic del gen *rumA* d'*E. coli* plasmidi R391, que al seu temps és homòleg al gen *umuD*, que codifica una subunitat de la polimerasa induïble per danys en el DNA o polimerasa V (Tang *et al.*, 1999).

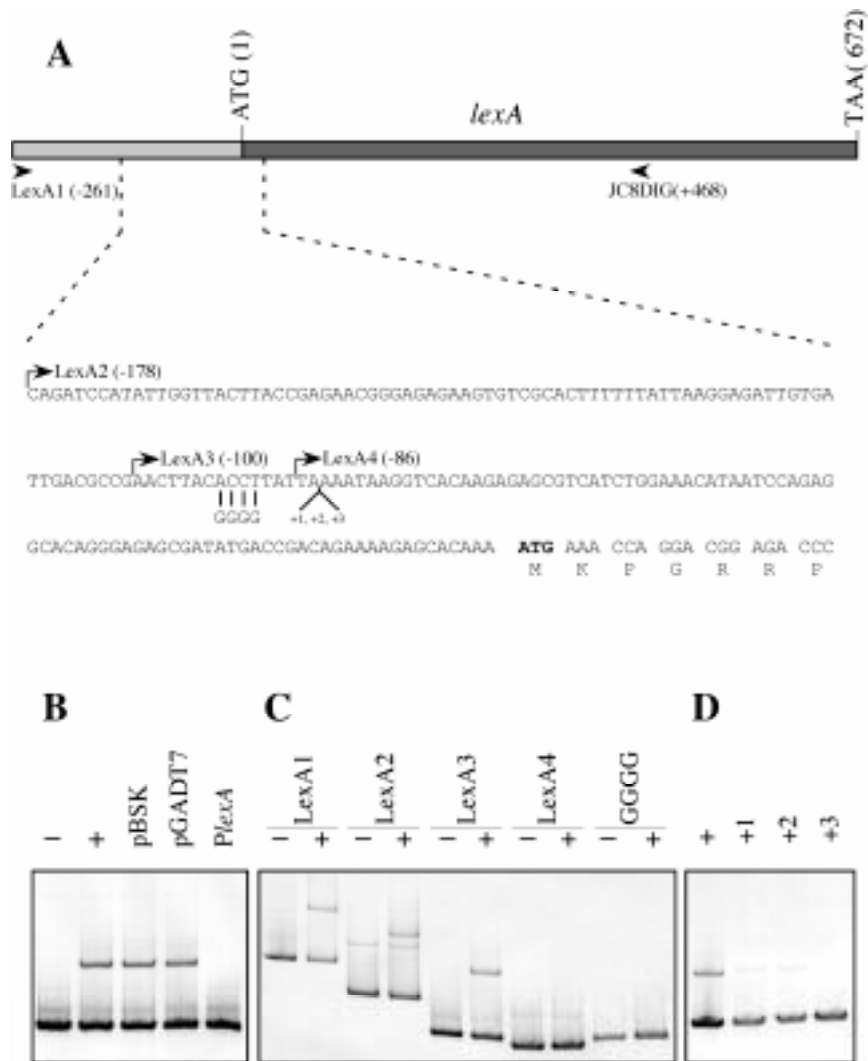
**A****B**

**Figura 2.9.** A) Organització genètica de la regió *lexA* d'MC-1. El codó hipotètic d'inici de traducció de cada un dels gens es mostra en negreta i subratllat. Les fletxes indiquen les posicions dels primers usats per a identificar els diferents transcrits. Les posicions numèriques estan en relació a l'hipotètic inici de traducció del gen *lexA*. B) Resultat dels anàlisis transcripcionals realitzats per mitjà d'assaigs RT-PCR de la regió flanquejant del gen *lexA* usant l'RNA total de cèl·lules MC-1 (RNA-RT-PCR). Com a control, reaccions de PCR foren realitzades amb els mateixos primers però sense retro-transcriptasa i amb RNA (RNA-PCR) o DNA (DNA-PCR) com a motlle. El tamany de les bandes fou estimat mitjançant un marcador molecular (genoma del bacteriòfag  $\lambda$  DNA digerit amb *HindIII*) ubicat en el segon carril de l'esquerra del gel.

El fet que la distància entre *lexA* i aquests tres marcs oberts de lectura sigui bastant curta, suggeriria que tots aquests gens podien ser co-transcrits (figura 2.9A). Per a provar aquest fet, i mitjançant l'ús de diferents primers dissenyats per a tal fi, es va procedir a la realització de les reaccions d'RT-PCR. Els fragments que havíem d'obtenir d'aquestes reaccions, eren d'una longitud de 1542 pb i 1578 pb si el gen *lexA* i els tres ORF que el segueixen, formen un mRNA

policistrònic: la visualització de les proves, ens mostrar uns resultats que corroboraven la nostra hipòtesis (figura 2.9B). És destacable que en el plasmidi R391, un segon gen anomenat *rumB*, localitzat avall de l'esmentat *rumA* (Kulaeva *et al.*, 1995), es co-transcriu amb aquest últim, no essent així a MC-1. La capacitat d'exercir com a reguladors negatiu de determinats gens, impossibilitant o dificultant la seva transcripció, gràcies a la unió a una seqüència nucleotídica específica present en la regió promotora de cada un d'ells, era un element comú a totes les proteïnes LexA descrites fins el moment (Brent *et al.*, 1981; Winterling *et al.*, 1997; Fernández de Henestrosa *et al.*, 1998; Fernández de Henestrosa *et al.*, 2000). Per a la identificació d'aquesta seqüència a MC-1, una regió de 261 pb en direcció 5' en relació l'hipotètic codó d'inici de traducció del gen *lexA*, fou amplificada per mitjà dels primer JC7 i JC8DIG (figura 2.10A), obtenint el fragment LexA1. Aquest fragment fou utilitzat en proves de mobilitat electroforètica, a fi de poder observar el resultat de la seva incubació amb proteïna LexA purificada de la soca MC-1. El revelat cromàtic, s'obtingué gràcies al marcatge amb DIG del cebador JC8DIG, que és reconegut per un anticòs específic de conill que porta conjugada la fosfatasa alcalina que actua sobre una solució cromogènica indicant la posició del DNA en el filtre. En aquest cas aparegué banda de retard en la reacció que contenia el LexA purificat i el fragment LexA1 (figura 2.10B). Per a descartar que es tractés d'un artefacte, i que veritablement la unió era específica, varem estudiar si aquesta es mantenia en presència de DNA inespecífic en la reacció, i per un altre cantó, quin era el resultat d'afegir-hi en gran excés molar LexA1, però sense marcar, és a dir, un competidor específic. En el primer cas, el resultat fou que la banda retardada no es veia afectada, mentre que l'addició de DNA específic, l'eliminava (figura 2.10B). Ambdós, ens confirmaren que el motiu d'unió del LexA es trobava dins aquests 261 parells de bases.

A fi de filar més prim quines eren les coordenades genòmiques del nostre objecte d'estudi, varem procedir a dissenyar oligonucleòtids que ens permetessin obtenir fragments cada cop de menor longitud en detriment dels nucleòtids més distals de l'hipotètic codó d'inici de traducció del *lexA*. Mentre que els fragments LexA1, LexA2 i LexA3 mostraren banda de retard, no era així en el LexA4 (figura 2.10C), del que varem poder deduir que la nostra seqüència diana es trobava entre els posicions -100 i -86 en relació l'ATG de *lexA*. Un cop en aquest punt i com a



**Figura 2.10.** A) Seqüència del gen *lexA* i del seu promotor de la soca MC-1. La seqüència del promotor del *lexA*, els 7 primers residus aminoacídics de la proteïna LexA i els hipotètics codons d'inici i finalització de la traducció també hi són representats. Els punts d'inici de cada un dels fragments del promotor del *lexA* (LexA1, LexA2, LexA3 i LexA4) usats en els assaigs EMSA, estan indicats per mitjà de fletxes i la distància relativa de cada un d'ells respecte el codó ATG, es mostra en parèntesis. B) Estudi de la unió específica de la proteïna LexA d'MC-1 al DNA. Els assaigs EMSA es van realitzar usant diferents DNAs com a competidors. El fragment LexA1 fou incubat en absència (-) i presència (+) de proteïna LexA purificada. A fi de corroborar l'especificitat de la unió de LexA, el fragment LexA1 fou incubat amb proteïna LexA pura d'MC-1 amb diferents DNAs competidors no marcats en la mateixa reacció: 3µg de pBSK, 3µg de pGADT7 i fragment LexA1 (*PlexA*) en una quantitat molar 100 cops superior (1µg). C) Acotament del punt d'unió de la proteïna LexA d'MC-1. LexA1, LexA2, LexA3, LexA4 i el fragment derivat del LexA3 on el tetranucleòtid ACCT, fou canviat per GGGG (en la imatge el carril està indicat com a GGGG), foren incubats en absència (-) i presència (+) de proteïna LexA purificada. D). Assaigs EMSA realitzats amb els fragments derivats del fragment LexA3 on un (+1), dos (+2) o tres (+3) residus d'adenina foren inserits en la posició -84 respecte l'ATG d'inici de traducció. Les diferents reaccions van ser incubades en presència de proteïna LexA i carregades en un gel nadiu de Tris-glicina poliacrilamida. El fragment LexA3 salvatge fou incubat junt amb LexA actuant com a control positiu.



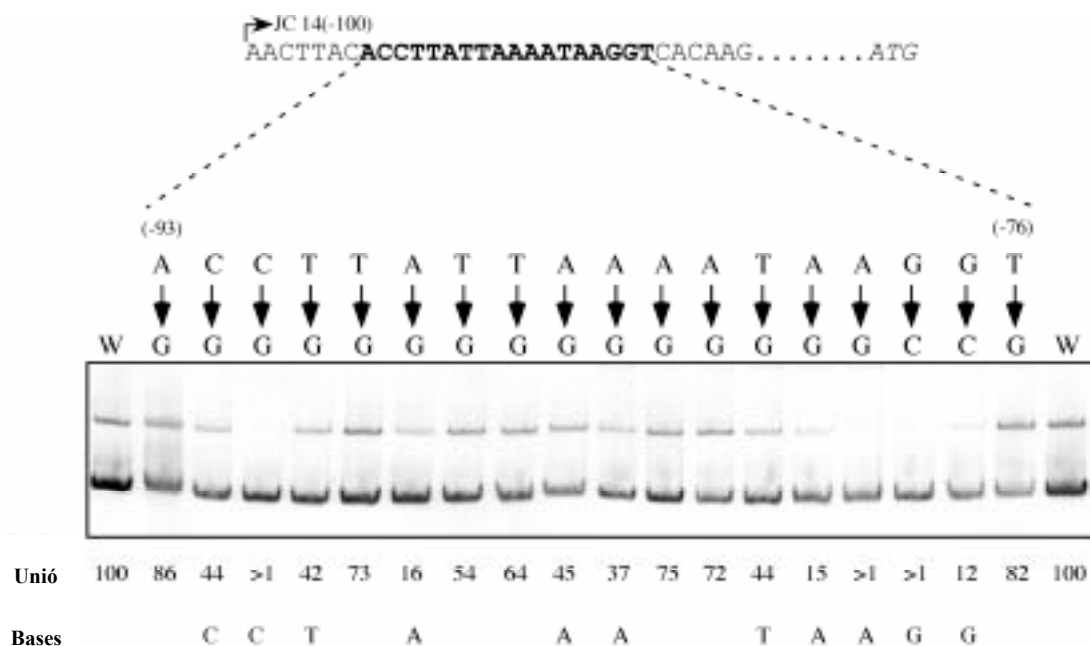
metodologia de treball, varem poder passar gràcies a l'acotament, a la inspecció visual de la regió, detectant una interessant repetició inversa a 75 parells de bases del inici de traducció: ACCTTATTAATAAGGT (ACCTT<sub>8</sub>AAGGT).

Per a poder afirmar amb tota rotunditat que era aquesta la seqüència que estàvem buscant, varem passar a realitzar alteracions en la seva seqüència per mitjà d'oligonucleòtids emprats com a eina per la realització de mutagènesi dirigida. Així en una primera instància, es va substituir el submotiu ACCT per GGGG gràcies al primer JC16 (figura 2.10C), resultant en una abolició de la unió, fet que provava la importància d'aquests 4 residus en la unió específica del LexA (figura 2.10C), i permetia extrapolar que en els correlatius de l'altre submotiu, també era així.

En proves posteriors, es va optar per augmentar la distància entre els submotius ACCT i AGGT, per mitjà de l'addició de nucleòtids addicionals en la zona connectora de la repetició inversa. Així, quan es modificava el fragment LexA3 mitjançant la incorporació d'una, dues o tres adenines a la esmentada regió d'unió (figura 2.10D), apareixien alteracions en la unió del LexA. Per exemple, en els dos primers casos, tant sols detectàvem una unió residual al fragment, i quan augmentàvem en tres la longitud, la unió desapareixia totalment. Aquests, junt amb els anteriors resultats, ens reafirmaren que la caixa SOS a *Magnetococcus sp.* soca MC-1, era: ACCTN<sub>10</sub>AGGT.

En aquest punt, varem intentar determinar quina era la importància relativa de cada una de les bases que integraven la nostra caixa SOS. Per a tal, varem dissenyar i usar els primers JC25-JC43 (figura 2.11), que ens permeteren realitzar una mutagènesi dirigida puntual en el fragment LexA3. Dels assaigs EMSA, se'n derivaren diferents fets 1) les bases CCT i AGG, que pertanyen als motius principals de la repetició inversa, eren vitals per la unió. 2) Precisant més, canvis en les bases -91 (C), -79 (A) i -78 (G) respecte l'hipotètic codó d'inici de traducció, implicaven alteracions dràstiques en l'afinitat del LexA (figura 2.11) i 3) per mitjà del software Macintosh ImageQuant (version 1.2), varem poder determinar, posteriorment a l'obtenció de les bandes cromàtiques en el filtre, que també jugaven un paper important en la unió els següents residus: A(-88, -85, -84), C (-92), T (-90, -81) i G (-77). Com a premissa per a confeccionar aquesta categorització de bases, es fixar un mínim de disminució del 50% de l'activitat d'unió del LexA (figura 2.11).

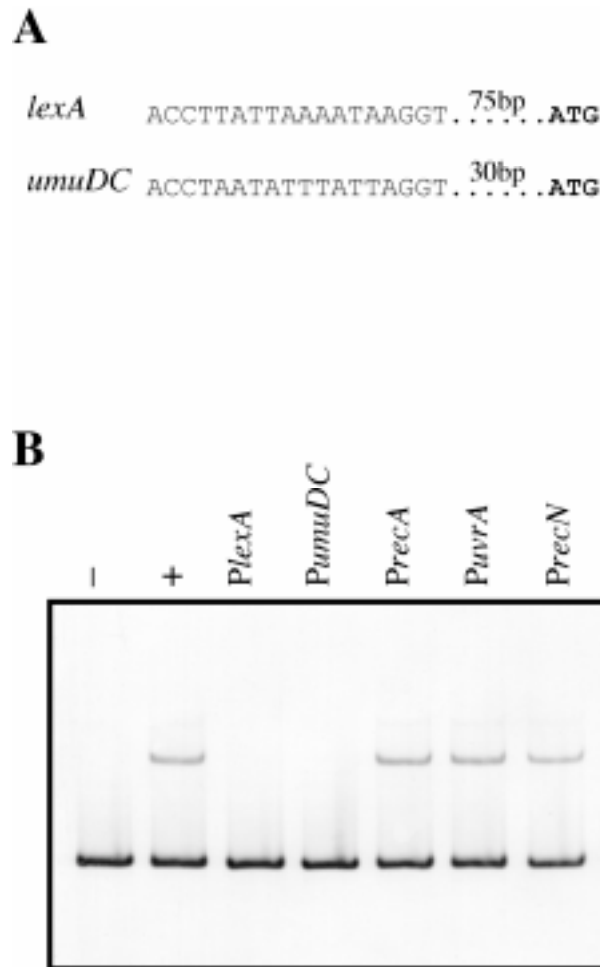
Així, esgrimint tots els resultats obtinguts, varem poder dictaminar com a seqüència que conformava la caixa SOS a MC-1, la que segueix: CCTTATTAATAAGG.



**Figura 2.11.** Determinació de la caixa SOS present en el promotor del gen *lexA* de la soca MC-1. La tècnica de mutagènesi dirigida mitjançant PCR fou usada per a canviar els nucleòtids en la regió compresa entre la posició -93 i la -76 (en negreta) en relació al codó d'inici de traducció (en cursiva) del gen *lexA* de *Magnetococcus sp.* soca MC-1. Els assaigs EMSA varen permetre determinar l'afinitat de la proteïna LexA purificada per mitjà dels fragments marcats amb DIG. Les fletxes indiquen el nucleòtid usat per a reemplaçar el present en el fragment salvatge, i la posició relativa respecte l'ATG es mostra entre parèntesis. El fragment LexA3 sense cap canvi introduït es mostra com a control positiu (W). El percentatge residual d'unió del LexA després de cada un dels canvis, fou quantificat per mitjà del programa ImageQuant 1.2 (Unió). Els nucleòtids clau, les mutacions dels quals causaven una disminució de la unió major al 50% (en relació al fragment salvatge control) es mostren en la part inferior de la figura (Bases).

### 2.2.3. El reguló *lexA* a *Magnetococcus sp* soca MC-1

A *E. coli*, un mínim de 40 gens estan directament regulats pel repressor LexA (Courcelle *et al.*, 2001; Khil *et al.*, 2002; Fernández de Henestrosa *et al.*, 2000). Aquest és amb diferència, el bacteri on s'han detectat major quantitat de gens amb la presència de la caixa SOS en la seva regió promotora. Partint de la premissa de la homologia, modulable, en la seqüència primària de les proteïnes amb funció similar, varem realitzar una cerca TBLASTN cercant aquests gens en el genoma incomplet de *Magnetococcus sp.* soca MC-1. Tot i que es detectaren força gens homòlegs als gens SOS d'*E. coli*, altres com *dinI*, *dinG*, *polB*, *sbmC*, *yebG*, *ydjM*, *yjiW* i *yfbE* no foren detectats, degut potser a que encara no es disposa de la seqüència de tot el genoma d' MC-1, que els productes proteics amb funció homòloga presenten una seqüència aminoacídica divergent o que simplement aquella funció no es realitzada per cap gen en aquest microorganisme.



**Figura 2.12.** A) Seqüències operadores situades amunt dels gens *lexA* i *umuDC* de la soca MC-1 que presenten regions potencials d'unió de la proteïna LexA. En la figura es mostren aquests potencials punts d'unió, els hipotètics codons d'inici de traducció així com la distància relativa entre ells. B) Assaigs EMSA realitzats amb el fragment LexA3 incubat amb proteïna LexA purificada d'MC-1 en presència de diferents DNAs competidors. El fragment LexA3 fou incubat en absència (-) i presència (+) de proteïna LexA pura. També es mostren les reaccions que contenen el fragment LexA3 marcat amb DIG incubat amb proteïna LexA pura junt amb un dels cinc DNAs competidors no marcats (en una concentració 100 cops superior), que respectivament contenen les regions promotores i els potencials punts d'unió del LexA als gens *lexA* (*PlexA*), *umuDC* (*PumuDC*), *recA* (*PrecA*), *uvrA* (*PuvrA*) i *recN* (*PrecN*), prèviament caracteritzats com regulats per LexA en la subclasse Alfa dels Proteobacteris.

Una anàlisi exhaustiva de les regions promotores dels gens detectats i de la totalitat del genoma disponible, realitzat mitjançant el programa EditSeq (versió 4.05) del paquet informàtic DNASTar, ens va revelar la presència d'una seqüència situada a -30 pb (figura 2.12A) en relació al codó d'inici de transcripció del gen *umuDC*: ACCTAATATTTATTAGGT. En la resta de gens analitzats, no es va trobar cap seqüència a la que concedir el benefici del dubte a fi de ser estudiada.

Per a comprovar si la seqüència de 18 pb ACCTAATATTTATTAGGT, es reconeguda per LexA, varem procedir a realitzar els assajos EMSA corresponents. Al mateix temps que

comprovarem la unió a aquest promotor, varem poder estudiar la presència, potser no detectada anteriorment, d'aquest motiu en els promotors dels gens *recA*, *uvrA* i *recN*. Tots aquests fragments, junt amb el corresponent al gen *umuDC*, foren usats com a competidors en les reaccions d'unió del LexA amb el LexA3. L'elecció d'aquests gens, sense tenir el compte els *umuDC* per motius evidents, es realitzar en base a estudis anteriors. Així per exemple, els gens *recA* i *uvrA*, junt amb el *ssb*, són comuns en la subclasse Alfa dels Proteobacteris (Fernández de Henestrosa *et al.*, 1998; Tapias i Barbé, 1999). Pel que fa a *recN*, es considerat un gen SOS canònic.

En els assaigs realitzats, cap d'ells excepte el fragment corresponent al promotor del gen *umuDC*, fou capaç de fer desaparèixer la banda d'unió, pel que podem afirmar que la seqüència detectada en aquest promotor es diana de la proteïna LexA d'aquest microorganisme, i per tant, controla negativament l'expressió de la polimerasa V o polimerasa tendent a error (figura 2.12B). Amb la intenció de conèixer quines eren les bases importants en aquest motiu, varem procedir a realitzar un procés similar a l'anteriorment descrit: mutagènesi dirigida puntual amb els primers JC57-JC74. Dels experiments EMSA realitzats, en varem poder inferir que hi ha diferents residus clau per la unió, però els més crítics serien la C de la posició -46, l'A de la posició -34 i la G de la posició -33 en relació al codó d'inici de traducció (figura 2.13).

Amb el mateix programa anterior, ImageQuant 1.2, varem poder determinar que cada nucleòtid jugava un paper significatiu en la unió (figura 2.13), que junt amb les dades dels nucleòtids primordials, ens mostraven la coincidència de resultats amb els estudis del promotor *lexA*, corroborant encara més, la seqüència CCTN<sub>10</sub>AGG com l'específica per la unió del LexA en la soca MC-1.

El nombre de gens sota el control de la proteïna LexA és extremadament variable en cada espècie. Així, és a *E. coli* on s'ha descrit un nombre més elevat de gens que formen part d'aquest reguló (Courcelle *et al.*, 2001; Fernández de Henestrosa *et al.*, 2000; Khil i Camerini-Otero, 2002). En altres espècies, com *Xylella fastidiosa* (Campoy *et al.*, 2002) o el mateix *Magnetococcus sp.* soca MC-1, les cerques informàtiques realitzades en els seus genomes, tant sols ens han reportat tres gens sota el control del LexA, de manera semblant a *D. ethenogenes* on tant sols se'n varen poder identificar dos. Per tant, en tots tres microorganismes, la majoria de gens que a *E. coli* romanen sota el control negatiu del LexA, escapen de la seva regulació (Campoy *et al.*, 2002; Courcelle *et al.*, 2001; Fernández de Henestrosa *et al.*, 2000; Khil i Camerini-Otero, 2002). L'explicació, segurament rau en la necessitat d'adaptació a cada un dels habitats que colonitzen les diferents espècies, fet que implicaria diferents organitzacions òptimes. Aquest fet explicaria les divergències en el contingut dels gens en el sistema SOS de cada una de les espècies, ja que si les condicions ambientals són molt exigents i per tant trobem la cèl·lula sotmesa a un estrès continu, serà més eficient que alguns gens de reparació s'expressin de forma constitutiva.



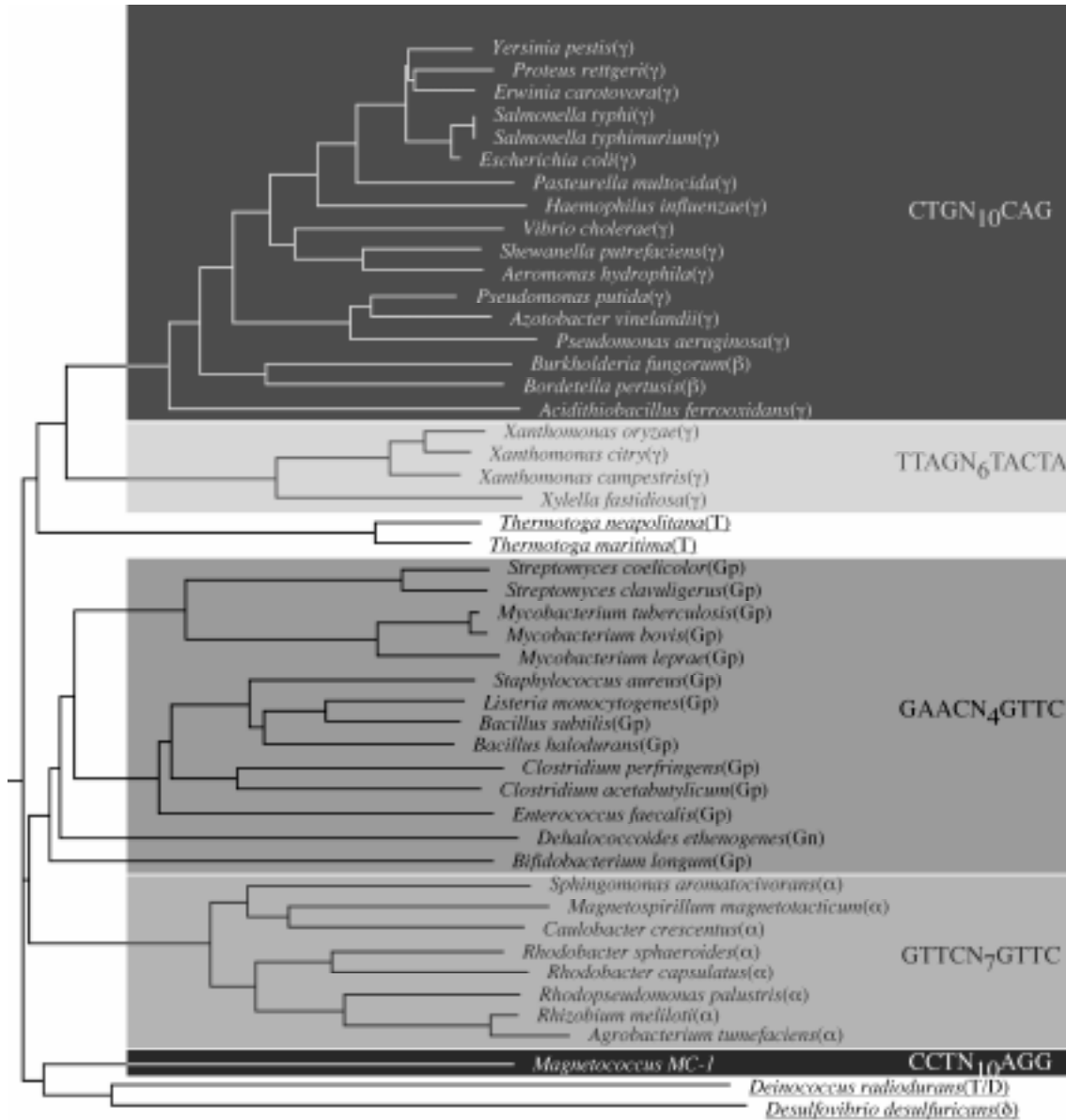
En aquest grup filogenètic, inferit de les dades que es disposava, els gens *recA*, *uvrA* i *ssb* havien estat considerats com a integrants del reguló *lexA* (Fernández de Henestrosa *et al.*, 1998; Tapias i Barbé, 1999). No obstant, en aquest estudi s'ha provat que el LexA no s'uneix a la regió promotora de cap d'ells, no exercint per tant, de regulador negatiu. En el cas del gen *ssb*, no podem assegurar-ho totalment, ja que en no disposar de la totalitat del genoma, el contig on es va detectar el marc obert de lectura corresponent, era d'una longitud molt limitada, fet que podria implicar l'absència de la regió reguladora. Per altra banda, la caixa SOS detectada i provada en *Magnetococcus sp.* sota MC-1, no es correspon amb la repetició directa descrita en la subclasse Alfa dels Proteobacteris. Per tant, semblava que els nostres resultats no encaixaven amb la classificació taxonòmica actual basada en estudis amb rRNA.

La realització d'un dendrograma usant la regió N-terminal d'unió al DNA del LexA de 48 proteïnes LexA pertanyents a organismes de diferents grups filogenètics, corroboraren les nostres apreciacions, ja que la proteïna LexA de la sota MC-1, divergia clarament de la branca principal de la subclasse Alfa dels Proteobacteris (figura 2.14). Malgrat la proteïna LexA d'MC-1 mostrava una elevada identitat amb la dels bacteris grampositius, aquesta es mostra com una branca independent separada d'aquests, estretament relacionada amb els bacteris gramnegatius de la subclasse Alfa dels Proteobacteris.

Una altra possibilitat seria el possible origen viral del gen *lexA*, ja que tant els residus aminoacídics implicats en l'autotrencament com el mecanisme d'hidròlisis són comuns en els repressors lítics dels bacteriòfags (Walker, 1984) i en les proteïnes LexA. De fet, aquesta és una explicació proposada de l'origen de la segona còpia del gen *lexA* present en *D. radiodurans* (Narumi *et al.*, 2001). En aquest bacteri el gen *lexA* es troba ubicat a continuació d'un fag temperat defectiu integrat en el cromosoma I (Morgan *et al.*, 2002), fet que fa força plausible aquesta explicació. Tanmateix aquesta hipòtesis no sembla que sigui extrapolable al cas que ens concerneix, ja que no es detectaren gens relacionats amb bacteriòfags situats prop del gen *lexA* o de la unitat transcripcional *umuDC*, que com es comprovar, formen part del reguló *lexA* ([http://www.jgi.doe.gov/JGI\\_microbial/html/magnetococcus/magneto\\_homepage.html](http://www.jgi.doe.gov/JGI_microbial/html/magnetococcus/magneto_homepage.html)).

Per mitjà de la possibilitat que ens conferia l'eina CLUSTAL W, en tant la capacitat d'associar una nova proteïna LexA a un grup en base a la seva seqüència d'unió diana, varem obtenir una sèrie de resultats, no en perfecte comunió amb els grups de filiació descrits fins el moment. Espècies pertanyents a diferents grups filogenètics com són *Burkholderia fungorum* i *Bordetella pertusis* (de la subclasse Beta dels Proteobacteris), apareixien agrupades junt amb membres de la subclasse gamma que s'unien a la mateixa caixa LexA: CTGN<sub>10</sub>CAG (figura 2.14). La proteïna LexA de *D. ethenogenes* apareixia relacionada amb els bacteris grampositius (figura 2.14), degut a la seva capacitat d'unir-se a la caixa DinR. Espècies del gènere *Xylella* i *Xanthomonas*, tot i la seva ubicació dins el grup de la subclasse Gamma dels Proteobacteris

(figura 2.14), com s'havia provat recentment, haurien d'agrupar-se en un grup independent des del punt de vista de la seva caixa SOS: TTAN<sub>6</sub>TACTA (Campoy *et al.*, 2002).



**Figura 2.14.** Dendrograma construït mitjançant el programa CLUSTAL W (Mac Vector, versió 6.5) usant el domini N-terminal de la proteïna LexA de 48 espècies diferents pertanyents a diferents grups filogenètics (des del primer residu aminoacídic fins al punt de trencament Ala-Gly). La matriu usada fou BLOSUM30. Els grups es mostren ombrejats amb una escala de grisos, segons el motiu d'unió del LexA detectat en la regió promotora dels respectius gens *lexA*. Els organismes que no mostren cap de les cinc caixes LexA descrites fins el moment estan subratllats. El grup filogenètic al que pertanyen es troba entre parèntesis. α, β, γ i δ, correponen al grup Alfa, Beta, Gamma i Delta dels Proteobacteris respectivament; Gp, grampositius; Gn, bacteris verds no sulfurosos; T, *Thermotogales*; T/D, grup *Thermus/Deinococcus*.

La descripció d'aquest motiu d'unió a la soca MC-1, augmentava fins a cinc el nombre total de motius en el domini *Bacteria*. Possiblement, aquest nombre es veurà incrementat, des del moment en que altres espècies tals com *Desulfovibrio desulfuricans* (subclasse Delta dels Proteobacteris) o *Thermotoga maritima* (*Thermotogales*) no contenen cap d'aquests cinc motius, però si que presenten *lexA* (<http://igweb.integratedgenomics.com/GOLD/index.cgi?want=Prokaryotic+Ongoin+Genomes>), el que suggereix la utilització d'una seqüència no descrita fins el moment per a la regulació negativa del LexA. El seu coneixement, pot aportar més llum a quina ha estat la evolució de la xarxa del sistema SOS, així com la del seu regulador.

### **2.3. Estudi del reguló *lexA* a *Leptospira interrogans***

#### **2.3.1. Caracterització de l'organització transcripcional del gen *lexA***

Per a determinar els gens que formen part dels regulons bacterians, primerament és necessari determinar quina és la seqüència d'unió de la proteïna reguladora en cada un d'ells. Totes les proteïnes LexA caracteritzades i estudiades en el món dels bacteris, presentaven la capacitat de regular la seva pròpia transcripció, en alguns casos en unitats transcripcionals policistròniques. L'estudi *in silico* de la regió cromosòmica que contenia l'hipòtetic gen *lexA* de *L. interrogans*, ens revelar la presència de dos marcs oberts de lectura (LA1445 i LA1446) amunt del gen *lexA* (Figura 2.15A).

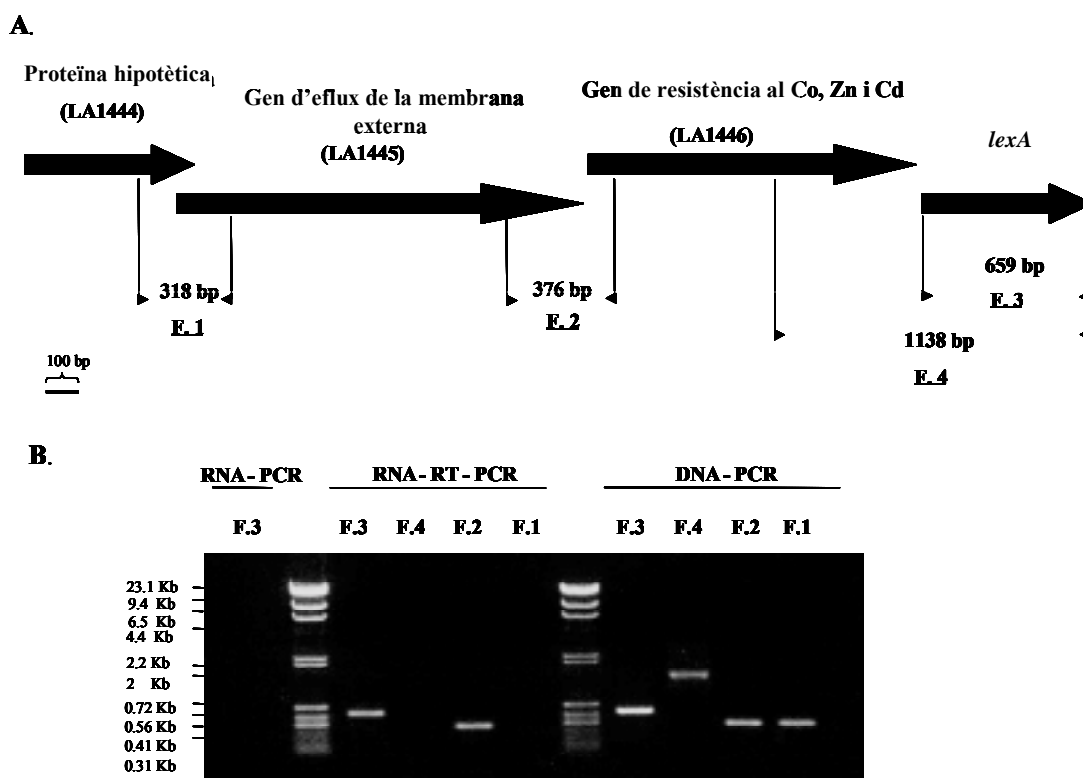
Amb la finalitat d'establir el punt on es trobava la caixa LexA de *L. interrogans* serovar Lai, era necessari coomprovar si aquests gens es presentaven en un sol transcrit, o per contra, cada un representava una unitat transcripcional independent. Per a obtenir tal informació, es va realitzar una anàlisi per mitjà d'una PCR amb la transcriptasa inversa (RT-PCR) en la totalitat de l'RNA extret de cèl·lules crescudes en medi EMJH, obtingut per mitjà del procediment convencional en aquest organisme (Cullen *et al.*, 2002). Prèviament, es van dissenyar uns primers que ens proporcionarien uns fragments de longitud coneguda en el cas que *lexA*-LA1446 o *lexA*-LA1446-LA1445, es trobessin transcripcionalment lligats. Els resultats obtinguts, ens descartaren aquesta possibilitat (figura 2.15B), però ens mostraren que LA1446 i LA1445 si que formaven una única unitat transcripcional (figura 2.15B).



### 2.3.2. Localització i caracterització de la seqüència d'unió del LexA

Els resultats obtinguts ens van permetre concloure que la hipotètica caixa LexA a *L. interrogans*, es trobava localitzada entre el final del marc obert de lectura LA1446 i el codó d'inici de traducció de *lexA*.

Per a confirmar aquesta premissa, el gen *lexA* fou amplificat per mitjà de PCR usant el primer NdeI-*lexA*, que presentava la diana de restricció *NdeI* solapada amb l'ATG d'inici de traducció del gen en qüestió. La parella de primers es completava amb el BamHI-*lexA*, l'extrem 5' del qual era complementari amb el nucleòtid ubicat a 30 pb avall del codó d'aturada de la traducció de *lexA*.



**Figura 2.15.** A) Organització genètica de la regió del gen *lexA* de *L. interrogans*. La longitud dels diferents fragments obtinguts (del F1 al F4), es mostra en la part superior de cada un dels noms. B) Les anàlisis RT-PCR dels hipotètics transcrits localitzats en els fragments F1-F4 de *L. interrogans* (RNA-RT-PCR). Com a controls, es realitzaren PCRs, usant els mateixos primers i com a motlle RNA (RNA-PCR) o DNA (DNA-PCR), però sense transcriptasa inversa. El marcador mol·lecular es mostra a l'esquerra de la imatge.

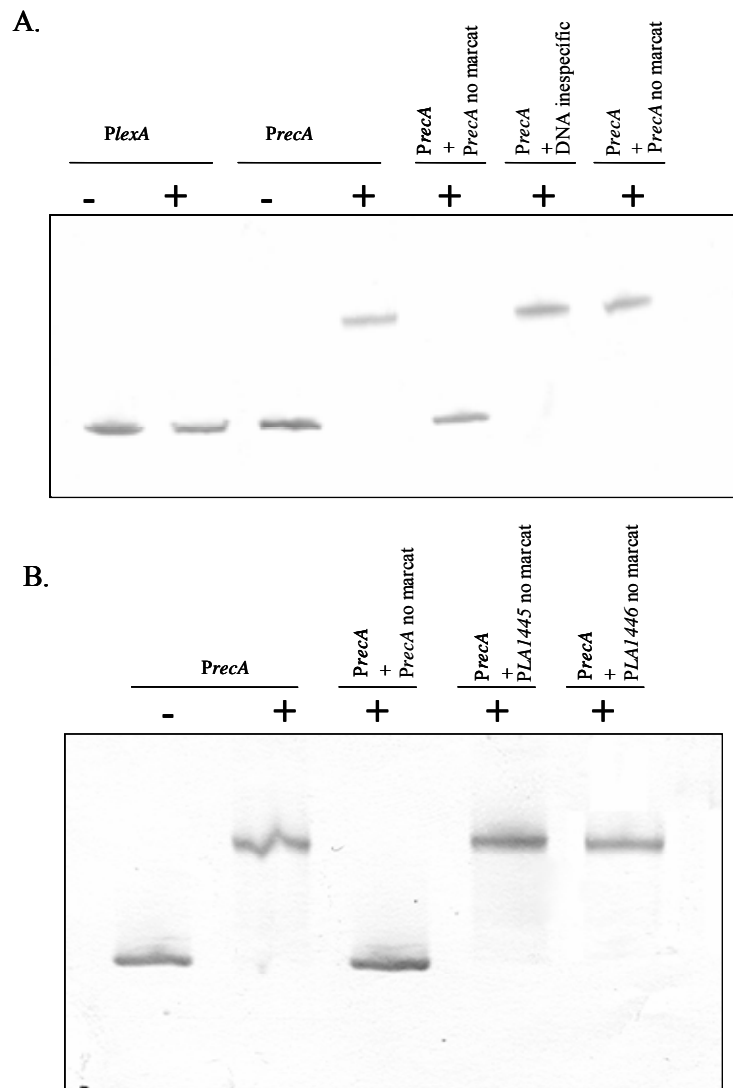
El fragment de 659 pb obtingut per mitjà de la PCR fou clonat en el pGEM-T, i a partir d'aquest i per mitjà de la digestió amb els esmentats enzims de restricció, en el vector d'expressió pET15b. El vector derivat d'aquesta última clonació, fou transformat a cèl·lules competents d'*E. coli* BL21 (DE3) CodonPlus (Stratagene) per a procedir a la sobreexpressió de la proteïna LexA, subseqüentment purificada usant el *Talon metal affinity resin kit* (Clontech) per mitjà del protocol ja descrit (Mazón *et al.*, 2004). El percentatge de puresa amb que es va obtenir fou determinat gràcies a la tinció amb blau de Coomassie emprada en un gel d' SDS-PAGE sotmès a un camp electroforètic, usant la metodologia estàndard (Laemmli, 1970).

Paral·lelament, per PCR i usant els primers LpLexAup i LpLexAdw, es va obtenir un fragment de 490 pb, que contenia la regió de davant de l'origen de traducció del *lexA*, junt amb 202 nucleòtids pertanyents a l'extrem 3' del marc obert de lectura LA1446.

Aquest fragment fou usat en assaigs de mobilitat electroforètica junt amb la proteïna LexA pura, esperant l'aparició d'una banda de retard, no essent així en cap dels casos en que es repetí l'experiment en pro de la corroboració dels resultats inicials.

Com alternativa i amb la premissa de que un gran quantitat dels gens *recA* bacterians estan negativament regulats per LexA, varem passar a comprovar la possibilitat de que la proteïna LexA si que s'unís a la regió promotora d'aquest gen. Com s'indica en la figura 2.16B, la proteïna LexA de *L. interrogans* s'uneix al fragment de 225 pb que conté la regió corrent amunt del gen *recA*. Aquesta interacció DNA-proteïna es provar que era específica, ja que l'addició de DNA específic (el mateix fragment que mostrava unió però sense marcar amb DIG, en concentració molt superior) abolia la banda de retard, no essent així, quan el que s'hi afegia era DNA inespecífic (figura 2.16A). En comunió amb els resultats negatius prèviament obtinguts, l'addició de fragment no marcat que contenia la regió promotora del gen *lexA*, no eliminava el canvi de mobilitat electroforètica del fragment (figura 2.16A).

El fet de que fins el moment tots els gens *lexA* presentessin la capacitat d'autoregulació, ens portar a plantejar-nos possibles explicacions a la no detecció del motiu d'unio en el nostre cas. Així, una primera explicació en relació amb l'organització transcripcional de *lexA*, podria ser una molt baixa expressió d'aquest gen, que faria factible la presència de la caixa en la regió de davant del LA1445 o bé del LA1446. Per a provar aquesta possibilitat, dos fragments de DNA no marcats que contenien les esmentades regions, foren usats com a competidors específics en les reaccions que contenien la regió promotora de *recA* més el LexA leptospiral purificat. Els resultats confirmaren la no presència de la caixa d'unio en cap dels fragments (figura 2.16B). Totes aquestes dades validen la conclusió que el *lexA* de *L. interrogans* no s'autoregula, però en canvi si que ho fa a altres gens, tals com *recA*.

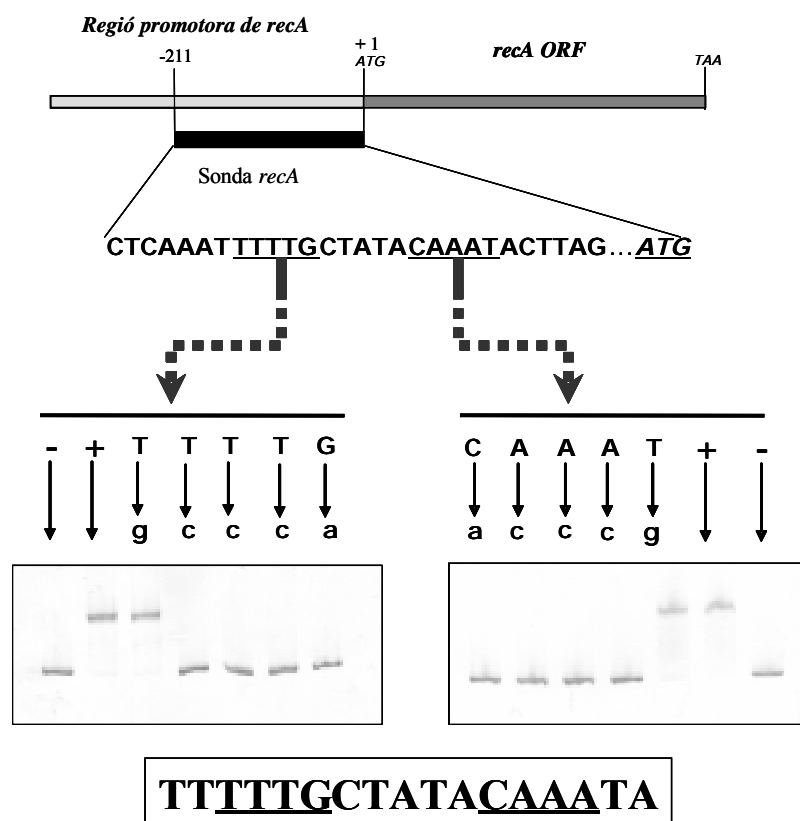


**Figura 2.16.** A) Assaigs EMSA realitzats amb el fragment que contenia l'operador del gen *lexA* i del *recA* en absència (-) i presència (+) de 40 ng de proteïna LexA purificada. Efecte de l'addició del fragment no marcat del operador *recA* (*PrecA*), *lexA* (*PlexA*) i de DNA inespecífic en una concentració 300 cops superior a la del *PrecA* marcat amb DIG. B) Efecte de l'addició amb una concentració 300 cops superior a la del *PrecA* marcat amb DIG dels fragments no marcats corresponents a la regió promotora de LA1445 i LA1446, en presència de 40 ng del LexA purificat de *L. interrogans*. Com a controls es va analitzar la mobilitat en el gel del fragment del promotor *recA* en absència (-) i presència (+) del LexA.

La seqüència on s'uneix el LexA de *L. interrogans* fou localitzada de manera precisa per mitjà dels experiments de *footprinting* realitzats amb el fragment portador del promotor del *recA*, la proteïna LexA purificada i el seqüenciador *ALF sequencer* (Pharmacia Biotech) per mitjà del protocol descrit en treballs anteriors (Campoy *et al*, 2003). El resultat dels assaigs, ens van mostrar que el nucli de la regió protegida, era de 20 nucleòtids (AATTTTTGCTATACAAATAC), trobant-se ubicat a 206 parells de bases davant de l'hipotètic codó d'inici de traducció del gen *recA*. Aquesta seqüència apareixia protegida de l'acció de la DNAsaI tant si s'usava la cadena codificant com la no codificant (figura 2.17).

La inspecció visual d'aquesta seqüència va revelar la presència d'un palíndrom perfecte, els submotius del qual eren TTTG i CAAA.

Per a confirmar els nostres resultats i determinar si aquest palíndrom és la seqüència reconeguda pel LexA, es varen introduir mutacions puntuals en cada un dels nucleòtids, així com en les regions flanquejants. Els canvis en el fragment portador de la regió promotora de *recA*, foren introduïts per mitjà de mutagènesi dirigida, emprant la metodologia convencional (Campoy *et al.*, 2003), i la presència dels canvis confirmada per seqüenciació. Després dels assaigs EMSA, varem poder afirmar que la unió requeria de les dues meitats del palíndrom, derivat de que es podia observar la disminució de la formació del complex DNA-LexA (figura 2.18).



**Figura 2.18.** Efectes de les mutacions puntuals introduïdes en el palíndrom TTTTGGCTATACAAATA present en la regió promotora de *recA* junt amb la presència de proteïna LexA purificada. Com a control es va usar el fragment amb la regió promotora salvatge de *recA* en absència (-) i presència (+) de proteïna LexA.

### 2.3.3. El reguló *lexA* a *L. interrogans*

Com esperàvem, no es va trobar cap còpia del palíndrom en la regió predecessora del gen *lexA* amb la cerca realitzada amb el programa informàtic RCGScanner, no sols això, aquest motiu no es va poder localitzar en cap dels gens categoritzats com a canònics del sistema SOS (*uvrAB*, *ruvAB*, *ssb* i *recN*) (Erill *et al.*, 2003; Erill *et al.*, 2004) ni en cap d'altre. Tot plegat, ens va portar a afirmar que l'únic gen directament regulat per LexA a *L. interrogans* era *recA*.

Altrament, estudis en el genoma de *L. interrogans* Copenhageni Fiocruz L1-130 (Nascimento *et al.*, 2004), llavors recentment seqüenciat, reafirmaren les nostres conclusions, ja que el palíndrom TTTGN<sub>5</sub>CAAA fou detectat davant del gen *recA*, però no del *lexA*.

El fet que l'únic gen sota el control directe del LexA en *L. interrogans* sigui el *recA*, podria indicar que es tracta d'un pas entremig en el procés de reducció genòmica que sembla que té lloc en les espiroquetes. Per tant, la presència del gen *lexA* a *L. interrogans*, se'ns mostra com un pas previ a la pèrdua d'aquest gen en les espiroquetes patògenes tals com *Treponema pallidum*, *Treponema denticola* i *Borrelia burgdorferi*, els genomes de les quals són significativament menors que el del nostre microorganisme en qüestió.

Anant un pas més enllà, es necessari destacar que l'absència de gen *lexA* en aquestes tres espiroquetes, pot augmentar el seu *fitness* cel·lular en resposta a la presència continua d'agents potencialment nocius i alterants del DNA, com poden ser els elements de defensa de l'hoste. El principal mecanisme de reparació de les lesions presents en els cromosomes dels bacteris patògens és la recombinació, del que es deriva que l'expressió constitutiva dels gens pertanyents a aquesta via (*recA*, *ruvAB* i *recN*), per la seva no regulació, segurament incrementa l'adaptació a aquets habitats (Kline *et al.*, 2003).

## **4. CONCLUSIONS**

### 3) CONCLUSIONS

- 1) La caracterització de la caixa SOS de *D. ethenogenes* (AGAACN<sub>4</sub>GTTCT), representant del filum dels bacteris verds no sulfurosos, revela una seqüència homòloga a la *Cheo box* (CGAACATATGTTTCG), fet que mostra un estret vincle evolutiu amb els bacteris grampositius.
- 2) L'estudi de la composició gènica del sistema SOS a *D. ethenogenes* demostra que tant sols es detecta una caixa LexA en el gen *uvrA*.
- 3) Tot i la ubicació de *Magnetococcus sp.* soca MC-1 en el grup dels Alfa Proteobacteris, la seqüència de la seva caixa SOS evidencia que aquest apareix com una branca independent dels mateixos, tot i que estretament relacionada.
- 4) A *Magnetococcus sp* soca MC-1, la caixa LexA caracteritzada no es detecta en cap dels gens considerats com a canònics del grup dels Proteobacteris Alfa. En canvi sí que es pot afirmar que la proteïna LexA sí controla l'expressió dels gens *umuDC*, fet que fiança la ubicació d'aquest microorganisme en una branca independent dels Proteobacteris Alfa.
- 5) La seqüència d'unió de la proteïna LexA descrita a *L. Interrogans*, tant sols es present davant del gen *recA*. Aquest és el primer cas descrit d'un gen *lexA* que no s'autoregula.
- 6) La caixa LexA descrita a *L. Interrogans* serovar Lai, es pot afirmar que és la representativa de les espiroquetes gràcies a la seva detecció per mitjà d'anàlisis informàtiques, en altres integrants del gènere.

## **5. BIBLIOGRAFIA**



#### 4. BIBLIOGRAFIA

**Anderson, D.G., i S.C. Kowalczykowski.** 1998. Reconstitution of an SOS response pathway: derepression of transcription in response to DNA breaks. *Cell* **95**: 975 – 979.

**Au, N., E. Kuester-Schoeck, V. Mandava, L.E. Bothwell, S.P. Canny, K. Chachu and nineteen more authors.** 2005. Genetic composition of the *Bacillus subtilis* SOS system. *J. Bacteriol.* **187**:7655 - 7666.

**Barabote, R.D., i M.H. Jr. Saier.** 2005. Comparative genomic analyses of the bacterial phosphotransferase system. *Microbiol. Mol. Biol. Rev.* **69**: 608 - 34.

**Battista, J.R., T. Nohmi, C.E. Donnelly, i G.C. Walker.** 1990a. Genetic analyses of cellular functions required for UV mutagenesis in *Escherichia coli*. *Basic Life Sci.* **52**: 269 - 75.

**Battista, J.R., J. Ohta, T. Nohmi, W. Sun, i G.C. Walker.** 1990b. Dominant negative *umuD* mutations decreasing RecA-mediated cleavage suggest roles for intact UmuD in modulations of SOS mutagenesis. *Proc. Natl. Acad. Sci. USA* **87**: 7190 – 7194.

**Brent, R., i M. Ptashne.** 1981. Mechanism of action of the *lexA* gene product. *Proc. Natl. Acad. Sci. USA* **78**: 4204 - 4208.

**Bonacossa De Almeida, C., G. Coste, S. Sommer, i A. Bailone.** 2002. Quantification of RecA protein in *Deinococcus radiodurans* reveals involvement of RecA, but not LexA, in its regulation. *Mol. Genet. Genomics.* **268**: 28 - 41.

**Brooks, P.C., F. Movahedzadeh, i E.O. Davis.** 2001. Identification of some DNA damage-inducible genes of *Mycobacterium tuberculosis*: apparent lack of correlation with LexA binding. *J. Bacteriol.* **183**: 4459 – 4467.

**Calero, S., X. Garriga, i J. Barbé.** 1991. One-step cloning for isolation of bacterial *lexA*-like genes. *J. Bacteriol.* **173**: 7345 – 7350.

**Campoy, S., G. Mazon, A.R. Fernández de Henestrosa, M. Llagostera, R.B. Monteiro, i J. Barbé.** 2002. A new regulatory DNA motif of the gamma subclass Proteobacteria:

identification of the LexA binding site of the plant pathogen *Xylella fastidiosa*. *Microbiology*. **148**: 3583 - 3597.

**Campoy, S., M. Fontes, S. Padmanabhan, P. Cortés, M. Llagostera, i J. Barbé.** 2003. LexA-independent DNA damage-mediated induction of gene expression in *Myxococcus xanthus*. *Mol. Microbiol.* **49**: 769 – 781

**Campoy, S., G. Mazon, A.R. Fernandez de Henestrosa, M. Llagostera, P.B. Monteiro, i J. Barbé.** 2004. A new regulatory DNA motif of the gamma subclass Proteobacteria: identification of the LexA protein binding site of the plant pathogen *Xylella fastidiosa*. *Microbiology*. **148**: 3583 - 3597.

**Campoy, S., N. Salvador, P. Cortes, I. Erill, i J. Barbé.** 2005. Expression of canonical SOS genes is not under LexA repression in *Bdellovibrio bacteriovorus*. *J. Bacteriol.* **187**: 5367 - 5375.

**Cheo, D.L., K.W. Bayles, i R.E. Yasbin.** 1991. Cloning and characterization of DNA damage-inducible promoter regions from *Bacillus subtilis*. *J. Bacteriol.* **173**: 1696 –1703.

**Churchill, J.J., D.G. Anderson, i S.C. Kowalczykowski.** 1999. The RecBC enzyme loads RecA protein onto ssDNA asymmetrically and independently of chi, resulting in constitutive recombination activation. *Genes Dev.* **13**: 901 - 11.

**Collyn, F., H. Fukushima, C. Carnoy, M. Simonet, i P. Vincent.** 2005. Linkage of the horizontally acquired *ypm* and *pil* genes in *Yersinia pseudotuberculosis*. *Infect. Immun.* **73**: 2556 - 2568.

**Courcelle, J., A. Khodursky, B. Peter, P.O. Brown, i P.C. Hanawalt.** 2001. Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics.* **158**: 41 – 64.

**Cox, M.M.** 2000. Recombinational DNA repair in bacteria and the RecA protein. *Progress Nucl. Acid. Res.* **63**: 311 – 366.

**Cox, M.M.** 2003. The bacterial RecA protein as a motor protein. *Annu. Rev. Microbiol.* **57**: 551 - 577.

- Cullen, P.A., S.J. Cordwell, D.M. Bulach, D.A. Haake, i B. Adler.** 2002. Global analysis of outer membrane proteins from *Leptospira interrogans* serovar Lai. *Infect. Immun.* **70**: 2311 - 2318.
- Davis, E.O., E.M. Dullaghan, i L. Rand.** 2002. Definition of the Mycobacterial SOS box and use to identify LexA-regulated genes in *Mycobacterium tuberculosis*. *J. Bacteriol.* **184**: 3287 - 3295.
- del Rey, A., J. Diestra, A.R. Fernández de Henestrosa, i J. Barbé.** 1999. Determination of the *Paracoccus denitrificans* SOS box. *Microbiology.* **145**: 577 - 584.
- Demple, B.** 1996. Redox signaling and gene control in the *Escherichia coli* *soxRS* oxidative stress regulon. *Gene.* **7**: 53 - 57.
- Demple, B.** 1997a. Exploitation of oxygen. *Methods.* **11**: 265 - 276.
- Demple, B.** 1997b. Study of redox-regulated transcription factors in prokaryotes. *Methods.* **11**: 267 - 278.
- Drees, J.C., S.L. Lusetti, i M.M. Cox.** 2004. Inhibition of RecA protein by the *Escherichia coli* RecX protein: modulation by the RecA C terminus and filament functional state. *J. Biol. Chem.* **279**: 52991 - 52997.
- Dronkert, M.L., i R. Kanaar.** 2001. Repair of DNA interstrand cross-links. *Mutat. Res.* **486**: 217 - 247.
- Duarte, F.T., F.M. Carvalho, U. Bezerra e Silva, K.C. Scortecci, C.A. Blaha, L.F. Agnez-Lima, i A.S. Medeiros.** DNA repair in *Chromobacterium violaceum*. *Genet. Mol. Res.* **31**: 167 - 180.
- Dudás, A., i M. Chovanec.** 2004. DNA double-strand break repair by homologous recombination. *Mutat. Res.* **566**: 131 - 167.
- Egelman, E.H.** 1998. Bacterial helicases. *J. Struct. Biol.* **124**: 123 - 128.

**Eisen, J.A.** 1995. The RecA protein as a model molecule for molecular systematic studies of bacteria: comparison of trees of RecAs and 16S rRNAs from the same species. *J. Mol. Evol.* **41**: 1105 - 1123.

**Ekaterina, G.F., J. Hauser, A.S. Levine, i R. Woodgate.** 1993. Targeting of the UmuD, UmuD', and MucA' mutagenesis proteins to DNA by RecA protein. *Proc. Natl. Acad. Sci. USA* **90**: 8169 – 8173.

**Erill, I., M. Escribano, S. Campoy, i J. Barbé.** 2003. In silico analysis reveals substantial variability in the gene contents of the Gamma Proteobacteria LexA regulon. *Bioinformatics* **19**: 2225 – 2236

**Erill, I., M. Jara, N. Salvador, M. Escribano, S. Campoy, i J. Barbé.** 2004. Differences in LexA regulon structure among Proteobacteria through in vivo assisted comparative genomics. *Nucleic Acids Res.* **32**: 6617 - 6626.

**Fernández de Henestrosa, A.R., E. Rivera, A. Tapias, i J. Barbé.** 1998. Identification of the *Rhodobacter sphaeroides* SOS box. *Mol. Microbiol.* **28**: 991 – 1003.

**Fernandez de Henestrosa, A.R., T. Ogi, S. Aoyagi, D. Chafin, J.J. Hayes, H. Ohmori, i R. Woodgate.** 2000. Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. *Mol. Microbiol.* **35**: 1560 - 1572.

**Fogh, R.H., G. Oettleben, H. Ruterjans, M. Schnarr, R. Boelens, i R. Kaptein.** 1994. Solution structure of the LexA repressor DNA binding domain determined by <sup>1</sup>H NMR spectroscopy. *EMBO J.* **13**: 3936 - 3944.

**Friedberg, E.C., G.C. Walker, i W. Siede.** 1995. DNA repair and mutagenesis. American Society for Microbiology Press, American Society for Microbiology, Washington D. C.

**Fyfe, J.A, i J.K. Davies.** 1990. Nucleotide sequence and expression in *Escherichia coli* of the *recA* gene of *Neisseria gonorrhoeae*. *Gene* **93**: 151 – 156.

**Garriga, X., S. Calero, i J. Barbé.** 1992. Nucleotide sequence analysis and comparison of the *lexA* genes from *Salmonella typhimurium*, *Erwinia carotovora*, *Pseudomonas aeruginosa* and *Pseudomonas putida*. *Mol. Gen. Genet.* **236**: 125 – 134.

- Geoghegan, C.M., i J.A. Houghton.** 1987. Molecular cloning and isolation of a cyanobacterial gene which increases the UV and methyl methanesulphonate survival of *recA* strains of *Escherichia coli* K12. *J. Gen. Microbiol.* **133**: 119 - 26.
- George, H., I. Kuraoka, D.A. Nauman, W.R. Kobertz, R.D. Wood, i S.C. West.** 2000. RuvAB-mediated branch migration does not involve extensive DNA opening within the RuvB hexamer. *Curr. Biol.* **10**: 103 - 116.
- González, M., i R. Woodgate.** 2002. The “tale” of UmuD and its role in SOS mutagenesis. *BioEssays.* **24**: 141 – 148.
- Groban, E.S., M.B. Johnson, P. Banky, P.G. Burnett, G.L. Calderon and thirteen more authors..** 2005. Binding of the *Bacillus subtilis* LexA protein to the SOS operator. *Nucleic Acids Res.* **33**: 6287 - 6295.
- Gupta, R.S.** 1998. Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaeobacteria, eubacteria, and eukaryotes. *Microbiol. Mol. Biol. Rev.* **62**: 1435 – 1491.
- Gupta, R.S.** 2000. The phylogeny of Proteobacteria: relationship to other eubacterial phyla and eukaryotes. *FEMS Microbiol. Rev.* **24**: 367 – 402.
- Gupta, R.S., i E. Griffiths.** 2002. Critical issues in bacterial phylogeny. *Theor. Popul. Biol.* **61**: 423 – 434
- Hajjema, B.J., D. van Sinderen, K. Winterling, J. Koositra, G. Venema, i L.W. Hamoen.** 1996. Regulated expression of the *dinR* and *recA* genes during competence development and SOS induction in *Bacillus subtilis*. *Mol. Microbiol.* **22**: 75 – 85.
- Hamoen, L.W., B. Hajjema, J.J. Bijlsma, G. Venema, i C.M. Lovett.** 2001. The *Bacillus subtilis* competence transcription factor, ComK, overrides LexA-imposed transcriptional inhibition without physically displacing LexA. *J. Biol. Chem.* **276**: 42901 – 42907.
- Harmon, F.G., W.M. Rehrauer, i S.C. Kowalczykowski.** 1996. Interaction of *Escherichia coli* RecA protein with LexA repressor II. Inhibition of DNA strand exchange by the uncleavable LexA S119A repressor argues that recombination and SOS induction are competitive processes. *J. Biol. Chem.* **271**: 238.

**Heide, L., M. Melzer, M. Siebert, A. Bechthold, J. Schroder, i K. Severin.** 1993. Clarification of the *Escherichia coli* genetic map in the 92-minute region containing the *ubiCA* operon and the *plsB*, *dgk*, *lexA*, and *dinF* genes. *J. Bacteriol.* **175**: 5728 - 9.

**Horii, T., T. Ogawa, T. Nakatani, T. Hase, H. Matsubara, i H. Ogawa.** 1981. Regulation of SOS functions: purification of *E. coli* LexA protein and determination of its specific site cleaved by the RecA protein. *Cell.* **27**: 515 - 522.

**Hugenholtz, P., B.M. Goebel, i N.R. Pace.** 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J. Bacteriol.* **180**: 4765 – 4774.

**Huisman, O., i R. d'Ari.** 1981. An inducible DNA replication-cell division coupling mechanism in *Escherichia coli*. *Nature* **290**: 797 – 799.

**Hurstel, S., M. Granger-Schnarr, M. Daune, i M. Schnarr.** 1986. In vitro binding of LexA repressor to DNA: evidence for the involvement of the amino-terminal domain. *EMBO J.* **5**: 793 - 798.

**Ivancic-Bace, I., P. Peharec, S. Moslavac, N. Skrobot, E. Salaj-Smic, i K. Brcic-Kostic.** 2002. RecFOR function is required for DNA repair and recombination in a RecA loading-deficient *recB* mutant of *Escherichia coli*. *Genetics* **163**: 485 - 94.

**Jara, M., C. Nunez, S. Campoy, A.R. Fernandez de Henestrosa, D.R. Lovley, i J. Barbé.** 2002. *Geobacter sulfurreducens* has two autoregulated *lexA* genes whose products do not bind the *recA* promoter: differing responses of *lexA* and *recA* to DNA damage. *J. Bacteriol.* **185**: 2493 - 2502.

**Karlin, S., G.M. Weinstock, i V. Brendel.** 1995. Bacterial classifications derived from RecA protein sequence comparisons. *J. Bacteriol.* **177**: 6881 – 6893.

**Karlin, S., i L. Brocchieri.** 1996. Evolutionary conservation of RecA genes in relation to protein structure and function. *J. Bacteriol.* **178**: 1881 – 1894.

**Kim, B., i J.W. Little.** 1992. Dimerization of a specific DNA-binding protein on the DNA. *Science*. **255**: 203 – 206.

**Kim, B., i J.W. Little.** 1993. LexA and lambda CI repressors as enzymes: specific cleavage in an intermolecular reaction. *Cell*. **73**: 1165 - 1173.

**Khil, P.P., i R.D. Camerini-Otero.** 2002. Over 1000 genes are involved in the DNA damage response of *Escherichia coli*. *Mol. Microbiol.* **44**: 89 – 105.

**Kline, K.A., E.V. Sechman, E.P. Skaar, H.S. Seifert.** 2003. Recombination, repair and replication in the pathogenic *Neisseriae*: the 3 R's of molecular genetics of two human-specific bacterial pathogens. *Mol. Microbiol.* **50**: 3-13.

**Knegtel, M.A., R.H. Fogh, G. Oettleben, H. Rüterjans, P. Dumoulin, M. Schnarr, R. Boelens, i R. Kaptein.** 1995. A model for the LexA repressor DNA complex. *Proteins*. **21**: 226 – 236.

**Koch, W.H., i R. Woodgate.** 1998. The SOS response DNA damage and repair: DNA repair in prokaryotes and lower eukaryotes. Eds. J. A. Nickoloff i M. F. Hoekstra, 1a. ed. Humana Press, Totowa, New Jersey. 107 – 134.

**Kogoma, T.** 1997. Stable DNA replication: interplay between DNA replication, homologous recombination and transcription. *Microbiol. Mol. Biol. Rev.* **61**: 212 – 238.

**Kojima, M., M. Suzuki, T. Morita, T. Ogawa, H. Ogawa, i M. Tada.** 1990. Interaction of RecA protein with pBR322 DNA modified by N-hydroxy-2-acetylaminofluorene and 4-hydroxyaminoquinoline 1-oxide. *Nucleic Acids Res.* **18**: 2707 - 2714.

**Kolodner, R., S.D. Hall, i C. Luisi-DeLuca.** 1994. Homologous pairing proteins encoded by the *Escherichia coli* *recE* and *recT* genes. *Mol. Microbiol.* **11**: 23 - 30.

**Konforti, B.B., i R.W. Davis.** 1992. ATP hydrolysis and the displaced strand are two factors that determine the polarity of RecA-promoted DNA strand exchange. *J. Mol. Biol.* **227**: 38 - 53.

**Krawczyk, B., K. Lewandowski, i J. Kur.** 2002. Comparative studies of the *Acinetobacter* genus and the species identification method based on the *recA* sequences. *Mol. Cell Probes*. **16**: 1 - 11.

**Kulaeva, O.I., J.C. Wootton, A.S. Levine, i R. Woodgate.** 1995. Characterization of the *umu*-complementing operon from R391. *J. Bacteriol.* **177**: 2737 - 2743.

**Kunin, V., L. Goldovsky, N. Darzentas, i C.A. Ouzounis.** 2005. The net of life: reconstructing the microbial phylogenetic network. *Genome Res.* **15**: 954 - 9.

**Kusano, K., N.K. Takahashi, H. Yoshikura, i I. Kobayashi.** 1994. Involvement of RecE exonuclease and RecT annealing protein in DNA double-strand break repair by homologous recombination. *Gene*. **138**: 17 - 25.

**Kuzminov, A.** 1999. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage  $\lambda$ . *Microbiol. Mol. Biol. Rev.* **63**: 751 – 813.

**Labazi, M., A. del Rey, A.R. Fernández de Henestrosa, i J. Barbé.** 1999. A consensus sequence for the *Rhodospirillaceae* SOS operators. *FEMS Microbiol. Lett.* **171**: 37 – 42.

**Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227**: 680 - 685.

**Landini, P., i M. Volkert.** 2000. Regulatory responses of the adaptive response to alkylation damage: a simple regulon with complex regulatory features. *J. Bacteriol.* **182**: 6543 – 6549.

**Lewis, L.K., G.R. Harlow, L.A. Gregg-Jolly, i D.W. Mount.** 1994. Identification of high affinity binding sites for LexA which define new DNA damage-inducible genes in *Escherichia coli*. *J. Mol. Biol.* **241**: 507 - 523.

**Little, J.W.** 1984. Autodigestion of LexA and phage lambda repressors. *Proc. Natl. Acad. Sci. USA*. **81**: 1375 – 1379.

**Little, J.W.** 1991. Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease. *Biochimie*. **73**: 411 – 421.



**Little, J.W.** 1993. LexA cleavage and other self-processing reactions. *J. Bacteriol.* **175**: 4943 – 4950.

**Love, P.E., M.J. Lyle, i R.E. Yasbin.** 1985. DNA damage-inducible (*din*) loci are transcriptionally activated in competent *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA.* **82**: 6201 – 6205

**Lovley, D. R.** 2000. Fe(III) and Mn(IV) reduction, p. 3-30. *In* D. R. Lovley (ed), enviromental microbe-*metal* interactions. ASM Press, Washington, D.C.

**Lusetti, S.L., O.N. Voloshin, R.B. Inman, R.D. Camerini-Otero, i M.M. Cox.** 2004a. The DinI protein stabilizes RecA protein filaments. *J. Biol. Chem.* **279**: 30037 -30046.

**Lusetti, S.L., J.C. Drees, E.A. Stohl, H.S. Seifert, i M.M. Cox.** 2004b. The DinI and RecX proteins are competing modulators of RecA function. *J. Biol. Chem.* **279**: 55073 -55079.

**Luo, Y., R.A. Pfuetzner, S. Mosimann, M. Paetzel, E.A. Frey, M. Cherney, B. Kim, J.W. Little, i N.C. Strynadka.** 2001. Crystal structure of LexA: a conformational switch for regulation of self-cleavage. *Cell* **106**: 585 – 594.

**Margulis, L.** 1996. Archaeal eubacterial mergers in the origin of Eukarya: phylogenetic classification of life. *Proc. Natl. Acad. Sci U S A.* **93**: 1071 - 1076.

**Mazón, G., I. Erill, S. Campoy, P. Cortes, E. Forano, i J. Barbe.** 2004.Reconstruction of the evolutionary history of the LexA-binding sequence. *Microbiology.* **150**: 3783-95.

**Michán, C., M. Machado, G. Dorado, i C. Pueyo.** 1999. *In vivo* transcription of the *Escherichia coli oxyR* regulon as a function of growth phase and in response to oxidative stress. *J. Bacteriol.* **181**: 2759 – 2764.

**Mohana-Borges, R., A.B. Pacheco, F.J. Sousa, D. Foguel, D.F. Almeida, i J.L. Silvas.** 2000. LexA repressor forms stable dimers in solution. *J. Biol. Chem.* **275**: 4708 – 4712.

**Moolenaar, G.F., R. Visse, M. Ortiz-Buysse, N. Goosen, i P. van de Putte.** 1994. Helicase motifs V and VI of the *Escherichia coli* UvrB protein of the UvrABC endonuclease are essential for the formation of the preincission complex. *J.Mol.Biol.* **240**: 294 – 307.

**Moolenaar, G.F., C. Moorman, i N. Goosen.** 2000. Role of the *Escherichia coli* nucleotide excision repair proteins in DNA replication. *J.Bacteriol.* **182**: 5706 – 5714.

**Moolenaar, G.F., L. Höglund, i N. Goosen.** 2001. Clue to damage recognition by UvrB: residues in the  $\beta$ -hairpin structure prevent binding to non-damaged DNA. *EMBO J.* **21**: 6140 – 6149.

**Moolenaar G.F., S. van Rossum-Fikkert, M. van Kesteren, i N. Goosen.** 2002. Cho, a second endonuclease involved in *Escherichia coli* nucleotide excision repair. *Proc. Natl. Acad. Sci. U S A.* **99**:1467 - 1472.

**Movahedzadeh, F., M.J. Colston, i E.O. Davis.** 1997a. Characterization of *Mycobacterium tuberculosis* LexA: recognition of a *Cheo* (*Bacillus*-type SOS) box. *Microbiology.* **143**: 929 – 936.

**Movahedzadeh, F., M.J. Colston, i E.O. Davis.** 1997b. Determination of DNA sequences required for regulated *Mycobacterium tuberculosis* RecA expression in response to DNA damaging agents suggests that two modes of regulation exist. *J. Bacteriol.* **179**: 3509 – 3518.

**Murphy, R.C., D.A. Bryant, R.D. Porter, i N.T. de Marsac.** 1987. Molecular cloning and characterization of the *recA* gene from the cyanobacterium *Synechococcus sp.* strain PCC 7002. *J. Bacteriol.* **169**: 2739 - 2747.

**Mustard, J.A., i J.W. Little.** 2000. Analysis of *Escherichia coli* RecA interactions with LexA,  $\lambda$  CI, and UmuD by site-directed mutagenesis of *recA*. *J. Bacteriol.* **182**: 1659 –1670.

**Nascimento, A.L., S. Verjovski-Almeida, M.A. Van Sluys, C.B. Monteiro-Vitorello, i L.E. Camargo and thirteen more authors.** 2004. Genome features of *Leptospira interrogans* serovar Copenhageni. *Braz. J. Med. Biol. Res.* **37**: 459 - 477.

**Narumi, I., K. Satoh, M. Kikuchi, T. Funayama, T. Yanagisawa, Y. Kobayashi, H. Watanabe, i K. Yamamoto.** 2001. The LexA protein from *Deinococcus radiodurans* is not involved in RecA induction following gamma irradiation. *J. Bacteriol.* **183**: 6951-956.

**Nascimento, A.L., A.I. Ko, E.A. Martins, C.B. Monteiro-Vitorello, i P.L. Ho and forty-six more authors.** 2004. Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. *J. Bacteriol.* **186**: 2164 - 72.

- Neher, S.B., J.M. Flynn, R.T. Sauer, i T.A Baker.** 2003. Latent ClpX-recognition signals ensure LexA destruction after DNA damage. *Genes. Dev.* **17**: 1084 - 089.
- Nohmi, T., J.R. Battista, L.A. Dodson, i G.C. Walker.** 1988. RecA-mediated cleavage activates UmuD mutagenesis: mechanistic relationship between transcriptional derepression and posttranslational activation. *Proc. Natl. Acad. Sci. USA* **85**: 1816 –1820.
- Morgan, G. J., G. F. Hatfull, S. Casjens, i R.W. Hendrix.** 2002. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. *J. Mol. Biol.* **317**: 337 - 359
- Norioka, N., M.Y. Hsu, S. Inouye, i M. Inouye.** 1995. Two *recA* genes in *Myxococcus xanthus*. *J. Bacteriol.* **177**: 4179 - 4182.
- Novichkov, P.S., M.V. Omelchenko, M.S. Gelfand, A.A. Mironov, Y.I. Wolf, E.V. Koonin.** 2004. Genome-wide molecular clock and horizontal gene transfer in bacterial evolution. *J. Bacteriol.* **186**: 6575 - 6585.
- Oertel-Buchheit, P., R.M. Lamerichs, M. Schnarr, i Granger-Schnarr, M.** 1990. Genetic analysis of the LexA repressor: isolation and characterization of LexA(Def) mutant proteins. *Mol. Gen. Genet.* **223**: 40 – 48.
- Oertel-Buchheit, P., D. Porte, M. Schnarr, i M. Granger-Schnarr.** 1992. Isolation and characterization of LexA mutant repressors with enhanced DNA binding affinity. *J. Mol. Biol.* **225**: 609 – 620.
- Oertel-Buchheit, P., J. Reinbolt, M. John, M. Granger-Schnarr, i M. Schnarr.** 1998. A LexA mutant repressor with a relaxed inter-domain linker. *Protein Sci.* **7**: 512 – 515.
- Owtrim, G.W., i J.R. Coleman.** 1987. Molecular cloning of a *recA*-like gene from the cyanobacterium *Anabaena variabilis*. *J. Bacteriol.* **169**: 1824 - 1829.
- Radman, M.** 1975. SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic. Life Sci.* **5A**: 355 - 367.

- Rand, L., J. Hinds, B. Springer, P. Sander, R.S. Buxton, i E.O. Davis.** 2003. The majority of inducible DNA repair genes in *Mycobacterium tuberculosis* are induced independently of RecA. *Mol. Microbiol.* **50**: 1031 – 1042.
- Rangarajan, S., R. Woodgate, i M.F. Goodman.** 2002. Replication restart in UV-irradiated *Escherichia coli* involving pols II, III, V, PriA, RecA and RecFOR proteins. *Mol. Microbiol.* **43**: 617 - 28
- Rehrauer, W.M., i S.C. Kowalczykowski.** 1996. The DNA binding site(s) of the *Escherichia coli* RecA protein. *J. Biol. Chem.* **271**: 11996 - 12002.
- Reuven, N.B., G. Arad, A.Z. Stasiak, A. Stasiak, i Z. Livneh.** 2001. Lesion bypass by the *Escherichia coli* DNA polymerase V requires assembly of a RecA nucleoprotein filament. *J. Biol. Chem.* **276**: 5511 – 5517.
- Riera, J., i J. Barbé.** 1993. Sequence of the *Providencia rettgeri* *lexA* gene and its control region. *Nucleic Acids Res.* **25**: 2256.
- Riera, J., i J. Barbé.** 1995. Cloning, sequence and regulation of expression of the *lexA* gene of *Aeromonas hydrophila*. *Gene* **154**: 71 - 75.
- Rivera, M.C., i J.A. Lake.** 2004. The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature.* **431**: 152 - 155.
- Roca, A.I., i M.M. Cox.** 1990. The RecA protein: structure and function. *CRC Crit. Rev. Biochem. Molec. Biol.* **25**: 415 – 456.
- Roland, K.L., M.H. Smith, J.A. Rupley, i J.W. Little.** 1992. In vitro analysis of mutant LexA proteins with an increased rate of specific cleavage. *J. Mol. Biol.* **228**: 395 - 408.
- Rostas, K., S.J. Morton, S.M. Picksley, i R.G. Lloyd.** 1987 Nucleotide sequence and LexA regulation of the *Escherichia coli* *recN* gene. *Nucleic Acids Res.* **15**: 5041 - 5049.
- Sassanfar, M., i J.W. Roberts.** 1990. Nature of the SOS-inducing signal in *Escherichia coli*. The involvement of DNA replication. *J. Mol. Biol.* **212**: 79 – 96.

- Sattin, B.D., i M.C. Goh.** 2004. Direct observation of the assembly of RecA/DNA complexes by atomic force microscopy. *Biophys. J.* **87**: 3430-6.
- Sauer, R.T., R.R. Yocum, R.F. Doolittle, M. Lewis, i C.O. Pabo.** 1982. Homology among DNA binding proteins suggests use of a conserved super-secondary structure. *Nature* **298**: 447 – 451.
- Schnarr, M., P. Oertel-Buchheit, M. Kazmaier, i M Granger-Schnarr.** 1991. DNA binding properties of the LexA repressor. *Biochimie.* **73**: 423 – 431.
- Schuler, D., S. Spring , i D.A. Bazyliniski.** 1999. Improved technique for the isolation of magnetotactic spirilla from a freshwater sediment and their phylogenetic characterization. *Syst. Appl. Microbiol.* **22**: 466 - 471.
- Silberstein, Z., Y. Tzfati, i A. Cohen.** 1995. Primary products of break-induced recombination by *Escherichia coli* RecE pathway. *J. Bacteriol.* **177**: 1692 - 1698
- Spring, S., i D. A. Bazyliniski.** 2000. Magnetotactic bacteria. In M. Dworkin, S. Falkow, E. Rosenberg, K.H. Schleifer, and E. Stackebrandt (ed.), *The prokaryotes: an evolving electronic resource for the microbiological community.*
- Stamm, L.V. i N.W. Charon.** 1988. Sensitivity of pathogenic and free-living *Leptospira sp.* to UV radiation and mitomycin C. *Appl. Environ. Microbiol.* **54**: 728 -33.
- Stamm, L.V., E.A. Parrish, i F.C. Gherardini.** 1991. Cloning of the *recA* gene from a free-living leptospire and distribution of RecA-like protein among spirochetes. *Appl. Environ. Microbiol.* **57**: 183 - 189.
- Starr, M.P. i R.J. Seidler.** 1971. The Bdellovibros. *Annu. Rev. Microbiol.* **25**: 649 -678.
- Story, R.M., I.T. Weber, i T.A Steitz.** 1992. The structure of the *E. coli* RecA protein monomer and polymer. *Nature.* **355**: 318 - 325.
- Sutton, M., B.T. Smith, V.G. Godoy, i G.C. Walker.** 2000. The SOS response: recent insights into *umuDC*-dependent mutagenesis and DNA damage tolerance. *Annu. Rev. Genet.* **34**: 479 – 497.

- Sutton, M.N., M. Kim, i G.C. Walker.** 2001. Genetic and biochemical characterization of a novel *umuD* mutation: insights into a mechanism for umuD self-cleavage. *J. Bacteriol.* **183**: 347 – 357.
- Tang, M., I. Bruck, R. Eritja, J. Turner, E.G. Frank, R. Woodgate, M. O'Donnell, i M.F. Goodman.** 1998. Biochemical basis of SOS-induced mutagenesis in *Escherichia coli*: reconstitution of *in vitro* lesion bypass dependent on the UmuD'<sub>2</sub>C mutagenic complex and RecA protein. *Proc. Natl. Acad. Sci. U S A.* **95**: 9755 - 9760
- Tang, M., X. Shen, E.G. Frank, M. O'Donnell, R. Woodgate, i M.F. Goodman.** 1999. UmuD'<sub>2</sub>C is an error-prone DNA polymerase, *Escherichia coli* polIV. *Proc. Natl. Acad. Sci. USA.* **96**: 8919 – 8924.
- Tapias, A., A.R. Fernández de Henestrosa, i J. Barbé.** 1997. Characterization of the promoter of the *Rhizobium etli recA* gene. *J. Bacteriol.* **179**: 1573 – 1579.
- Tapias, A., i J. Barbé.** 1999. Regulation of divergent transcription from the *uvrA-ssb* promoters in *Sinorhizobium meliloti*. *Mol. Gen. Genet.* **262**: 121 – 130.
- Tapias, A., S. Campoy, i J. Barbé.** 2000. Analysis of the expression of the *Rhodobacter sphaeroides lexA* gene. *Mol. Gen. Genet.* **263**: 957 – 965.
- Tapias, A., S. Fernández, J.C. Alonso, i J. Barbé.** 2002. *Rhodobacter sphaeroides* LexA dual activity: optimising and repressing *recA* gene transcription. *Nucleic Acids Res.* **30**: 1539 – 1546.
- Thliveris, A.T., i D.W. Mount.** 1992. Genetic identification of the DNA binding domain of *Escherichia coli* LexA protein. *Proc. Natl. Acad. Sci U S A.* **89**: 4500 - 4504.
- Van Houten, B. i A. Snowden.** 1993. Mechanism of action of the *Escherichia coli* UvrABC nuclease: clues to the damage recognition problem. *Bioessays.* **15**: 51 - 59.
- van Waasbergen, L.G., D.L. Balkwill, F.H. Crocker, B.N. Bjornstad, i R.V. Miller.** 2000. Genetic diversity among *Arthrobacter* species collected across a heterogeneous series of terrestrial deep-subsurface sediments as determined on the basis of 16S rRNA and *recA* gene sequences. *Appl. Environ. Microbiol.* **66**: 3454 - 3463.

**Walker, G.C.** 1996. The SOS response of *Escherichia coli*. F. C. Neidhardt, R. Curtis III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, i H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, 2a. ed. American Society of Microbiology Press, Washington D. C. 1400 - 1416.

**Winterling, K.W., A.S. Levine, R.E. Yasbin, i R. Woodgate.** 1997. Characterization of DinR, the *Bacillus subtilis* SOS repressor. *J. Bacteriol.* **179**: 1698 – 1703.

**Winterling, K.W., D. Chafin, J.J. Hayes, J. Sun, A.S. Levine, R.E. Yasbin, i R. Woodgate.** 1998. The *Bacillus subtilis* DinR binding site: redefinition of the consensus sequence. *J. Bacteriol.* **180**: 2201 – 2211.

**Woese, C.R.** 1987. Bacterial evolution. *Microbiol. Rev.* **51**: 221 – 271.

**Woese, C.R.** 2000. Interpreting the universal phylogenetic tree. *Proc. Natl. Acad. Sci. USA* **97**: 8392 – 8396.

**Wojciechowski, M.F., K.R. Peterson, P.E. Love.** 1991. Regulation of the SOS response in *Bacillus subtilis*: evidence for a LexA repressor homolog. *J. Bacteriol.* **173**: 6489 - 6498.

**Wolf, Y.I., L. Aravind, N.V. Grishin, i E.V. Koonin.** 1999. Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* **9**: 689 - 710.

**Yasbin, R.E., D.L. Cheo, i K.W. Bayles.** 1992. Inducible DNA repair and differentiation in *Bacillus subtilis*. Interactions between global regulons. *Mol. Microbiol.* **6**: 1263 – 1270.

**Yasuda, T., K. Morimatsu, R. Kato, J. Usukura, M. Takahashi, i H. Ohmori.** 2001. Physical interactions between DinI and RecA nucleoprotein filament for the regulation of SOS mutagenesis. *EMBO J.* **20**: 1192 – 1202.

**Yu, X., i E.H. Egelman.** 1993. The LexA repressor binds within the deep helical groove of the activated RecA filament. *J. Mol. Biol.* **231**: 29 - 40

**Zheng, M., B. Doan, T.D. Schneider, i G. Stolz.** 1999. OxyR and SoxRS regulation of *fur*. *J. Bacteriol.* **181**: 4639 – 4643.

**Zou, Y., R. Walker, H. Bassett, N.E. Geacintov, i B. Van Houten.** 1997. Formation of DNA repair intermediates and incision by the ATP-dependent UvrB-UvrC endonuclease. *J. Biol. Chem.* **272**: 4820 - 4827.



## **6. ANNEXES**

## **ARTICLE I**

## A Green Nonsulfur Bacterium, *Dehalococcoides ethenogenes*, with the LexA Binding Sequence Found in Gram-Positive Organisms

Antonio R. Fernández de Henestrosa,<sup>1\*</sup> Jordi Cuñé,<sup>1</sup> Ivan Erill,<sup>2</sup> Jon K. Magnuson,<sup>3</sup>  
and Jordi Barbé<sup>1,4\*</sup>

*Departament Genètica i Microbiologia*<sup>1</sup> and *Centre de Recerca en Sanitat Animal*,<sup>4</sup> *Universitat Autònoma de Barcelona*, and  
*Biomedical Applications Group, Centro Nacional de Microelectrónica*,<sup>2</sup> *Bellaterra, 08193 Barcelona, Spain*, and  
*Pacific Northwest National Laboratory, Richland, Washington 99352*<sup>3</sup>

Received 6 May 2002/Accepted 5 August 2002

***Dehalococcoides ethenogenes* is a member of the physiologically diverse division of green nonsulfur bacteria. Using a TBLASTN search, the *D. ethenogenes* *lexA* gene has been identified, cloned, and expressed and its protein has been purified. Mobility shift assays revealed that the *D. ethenogenes* LexA protein specifically binds to both its own promoter and that of the *uvrA* gene, but not to the *recA* promoter. Our results demonstrate that the *D. ethenogenes* LexA binding site is GAACN<sub>4</sub>GTTC, which is identical to that found in gram-positive bacteria. In agreement with this fact, the *Bacillus subtilis* DinR protein binds specifically to the *D. ethenogenes* LexA operator. This constitutes the first non-gram-positive bacterium exhibiting a LexA binding site identical to that of *B. subtilis*.**

In *Escherichia coli*, the SOS regulon consists of a pool of at least 40 DNA damage-inducible genes whose functions are involved in DNA replication, DNA repair, mutagenesis, and control of the cell cycle (3, 5, 8, 11). This regulon is controlled by two key proteins, RecA and LexA, the positive and negative regulators, respectively (1, 12, 15, 22). Under normal conditions LexA binds specifically to an operator region located within the promoters of the SOS genes, repressing their expression (8, 22). When DNA is damaged or replication is stalled, RecA protein forms a nucleofilament with the resulting single-stranded DNA, achieving its activated conformation (RecA\*). RecA\* mediates the autocatalytic cleavage of LexA between residues Ala<sup>84</sup> and Gly<sup>85</sup>, resulting in two parts, the N-terminal DNA binding domain and the C-terminal domain, which is involved in dimerization of the repressor (13, 14). In this process, the catalytic site formed by Ser<sup>119</sup> and Lys<sup>156</sup> is responsible for the hydrolysis of the Ala-Gly bond through a biochemical reaction characteristic of serine proteases (16). As a consequence of the cleavage, LexA no longer binds DNA, the SOS genes are derepressed, and their products participate in repairing the DNA lesions to guarantee cell survival. This process is known as the SOS response and was first suggested in the 1970s by Miroslav Radman (19). Once DNA is repaired, RecA\* loses its active conformation, LexA autohydrolysis ceases, and the repressor accumulates again, inhibiting the expression of the SOS genes.

The *lexA* gene seems to be widespread in the *Bacteria* domain. However, *lexA* is not found in the fully sequenced genomes of several bacterial species, such as *Aquifex aeolicus*,

*Borrelia burgdorferi*, *Chlamydia pneumoniae*, *Mycoplasma pneumoniae*, *Campylobacter jejuni*, *Helicobacter pylori*, and *Porphyromonas gingivalis*. Moreover, it appears that a *lexA*-like gene does not exist in the *Archaea* domain, as *lexA* homologues are absent in the 19 archaeal genomes currently available (4).

Recently, the three-dimensional structure of the *E. coli* LexA protein has been solved (16). In *E. coli*, LexA binds specifically to a DNA motif known as the SOS-box or LexA-box (22). Comparative analysis of the 31 *E. coli* LexA binding sites shows a 16-bp consensus sequence that responds to the motif CTGN<sub>10</sub>CAG (5). Besides the *E. coli* SOS-box, two more LexA-boxes have been identified so far in bacteria. The directed repeat GTTCN<sub>7</sub>GTTC is recognized by LexA protein of members of the alpha-class *Proteobacteria*, such as *Rhodobacter sphaeroides* and *Rhizobium etli* (6, 21). In gram-positive bacteria such as *Bacillus subtilis* and *Mycobacterium tuberculosis*, the consensus sequence CGAACRNRYGTTYG is the target for this repressor (2, 24).

*Dehalococcoides ethenogenes* is an anaerobic bacterium capable of dechlorinating tetrachloroethene, one of the most common groundwater contaminants, to ethene (17). *D. ethenogenes* is rather enigmatic taxonomically. Its cell wall composition is unlike that of gram-positive or gram-negative bacteria, instead more closely resembling the cell wall of archaea (17). However, on the basis of 16S rRNA gene sequence comparisons, *D. ethenogenes* has been assigned to the green nonsulfur bacteria, a division of physiologically diverse species with relatively few cultured representatives (10). The unusual phylogenetic position, structural properties, and metabolic capabilities of *D. ethenogenes* make it an interesting organism for further biological characterization. Here, we report the cloning of the key regulatory gene *lexA* from *D. ethenogenes*, purification of the protein, and characterization of its recognition site. Prior to this report, there were no data available pertaining to the motif recognized by LexA in this bacterial division. Finally,

\* Corresponding author. Mailing address: Universitat Autònoma de Barcelona, Departament de Genètica i Microbiologia, Unitat de Microbiologia, Edifici C, Bellaterra, 08193 Barcelona, Spain. Phone: 34 93 581 1837. Fax: 34 93 581 23 87. E-mail for Antonio R. Fernández de Henestrosa: antonio.rodriguez@uab.es. E-mail for Jordi Barbé: jordi.barbe@uab.es.

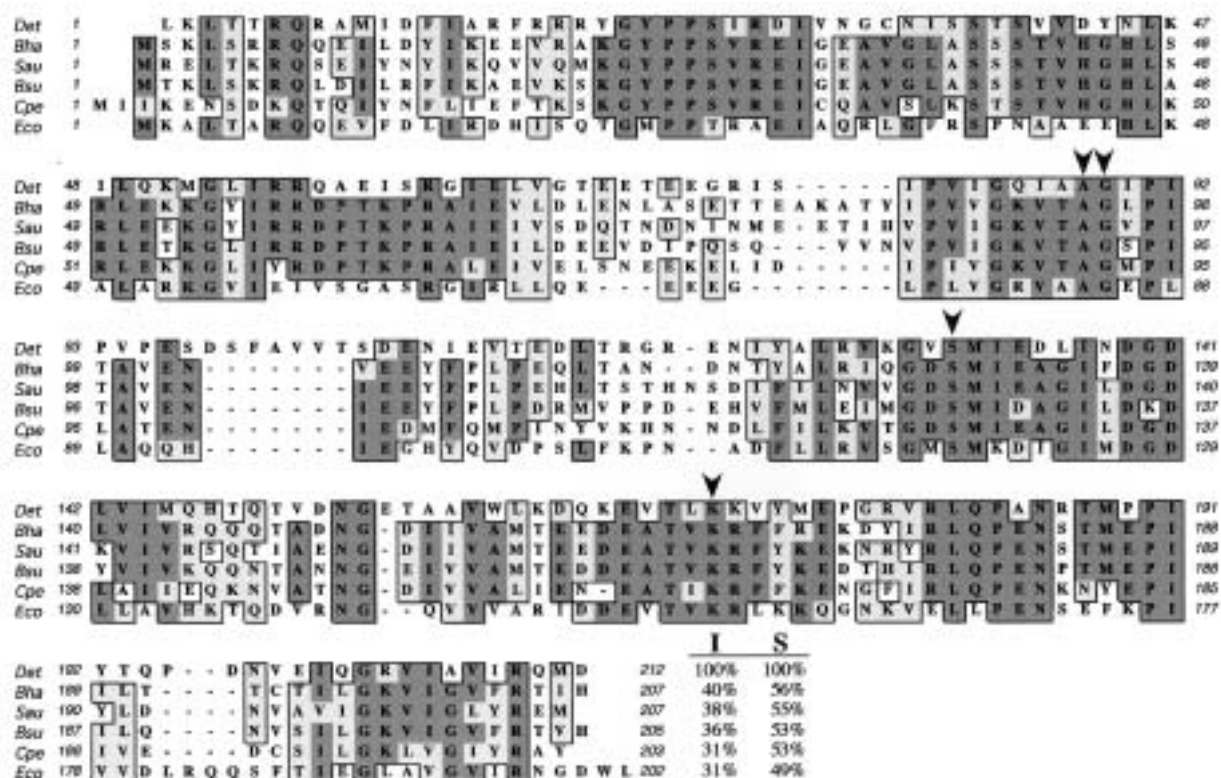


FIG. 1. CLUSTAL W alignment performed with Mac Vector 6.5 (Oxford Molecular) with LexA proteins from *Dehalococcoides ethenogenes* (Det), *Bacillus halodurans* (Bha), *Staphylococcus aureus* (Sau), *Bacillus subtilis* (Bsu), *Clostridium perfringens* (Cpe), and *Escherichia coli* (Eco). Identities have a dark gray background, and similarities are light gray. The perfectly conserved residues Ala<sup>88</sup> and Gly<sup>89</sup> forming the LexA cleavage site and the amino acids Ser<sup>131</sup> and Lys<sup>170</sup> required for protein hydrolysis are indicated with arrowheads on the *D. ethenogenes* LexA sequence. At the end of each LexA sequence appears the percent identity (I) and percent similarity (S) that each protein shows with the *D. ethenogenes* LexA repressor. Accession numbers in the Entrez protein database at NCBI are as follows: *B. halodurans*, Q9KAD3; *S. aureus*, BAB57501; *B. subtilis*, P31080; *C. perfringens*, BAB80867; and *E. coli*, P03033.

we report a genomic analysis of the composition of the *D. ethenogenes* LexA network.

**Identification and cloning of *D. ethenogenes* LexA gene.** The *D. ethenogenes* strain 195 genome sequence is currently unassembled, but contigs are available for performing BLAST analysis at the Institute for Genomic Research and at the National Center of Biotechnology Information ([http://www.ncbi.nlm.nih.gov/Microb\\_blast/unfinishedgenome.html](http://www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html)).

The *Escherichia coli* LexA amino acid sequence was used to query the *D. ethenogenes* genome with the TBLASTN program, revealing a protein homologue of 212 amino acids. Although *D. ethenogenes* is not a gram-positive organism, its LexA protein is highly homologous to the LexA proteins belonging to several members of this bacterial phylum, such as *Bacillus halodurans*, *Staphylococcus aureus*, and *Clostridium perfringens*. Specifically, *D. ethenogenes* LexA protein shows the highest level of identity with *Bacillus halodurans* LexA (40%), while the *D. ethenogenes* repressor is only 31% identical to *E. coli* LexA (Fig. 1). Furthermore, *D. ethenogenes* LexA contains all of the conserved residues involved in repressor autocleavage (Ala<sup>88</sup>, Gly<sup>89</sup>, Ser<sup>331</sup>, and Lys<sup>170</sup>) as deduced from CLUSTAL W alignment of different LexA proteins (Fig. 1). It is important to note that nine codons downstream of the proposed initiation site for the 212-amino-acid open reading

frame there is a 203-amino-acid protein starting with methionine. However, this truncated version of *D. ethenogenes* LexA would lack several N-terminal residues likely to be essential to the LexA structure, since it would lack important residues belonging to the first  $\alpha$ -helix characterized by nuclear magnetic resonance spectroscopy in *E. coli* LexA (7).

The BLAST analysis returned 1 kb of DNA flanking the *lexA* coding sequence, which was used to design the primers AR31 and AR34 (Table 1). These primers were employed to PCR amplify the entire *D. ethenogenes* *lexA* and its promoter region with genomic DNA obtained from a 10-liter tetrachloroethene-methanol-fed anaerobic enrichment culture. The resulting 866-bp DNA fragment was then cloned into pGEM-T (Promega). To confirm that no mutation was introduced during the PCR, the sequence of the insert was determined by labeling DNA samples with the fmol DNA cycle sequencing system (Promega) and analyzed on an ALF sequencer (Amersham Pharmacia Biotech). This plasmid, called pUA969, was used as a template for further studies.

**Purification of *D. ethenogenes* LexA protein.** A glutathione S-transferase (GST)-LexA fusion protein was generated. The primers AR61 and AR34 (Table 1) were used to amplify the *D. ethenogenes* *lexA* gene from pUA969. The oligonucleotide AR61 incorporates an *Eco*RI restriction site upstream of the

TABLE 1. Oligonucleotide primers used in this work<sup>a</sup>

Primer	Sequence	Position	Application
AR31	5'-CTCAACCAAGCCGCTAAATAACTCC-3'	-121	Upper primer for cloning <i>D. ethenogenes lexA</i> gene and LexA1 probe
AR33	5'-AGAGGTAGAGCTGATATTACAGCCG-3'	+117	Lower primer to amplify unlabeled specific DNA competitors
AR33DIG	5'-DIG-AGAGGTAGAGCTGATATTACAGCCG-3'	+117	Lower primer to amplify digoxigenin-labeled probe fragments of <i>lexA</i> promoter
AR34	5'-GTCGACCTAATCCATCTGGCGATTACAGC-3'	+636	Lower primer for cloning <i>D. ethenogenes lexA</i> gene
AR35	5'-TACCTGCCTGAATGACTTGACACC-3'	-295	Upper primer to amplify <i>D. ethenogenes recA</i> promoter
AR36	5'-CTTAGTTTCATTATTGAACCTTGC-3'	+85	Lower primer to amplify <i>D. ethenogenes recA</i> promoter
AR38	5'-TATGCCGGGACAGAGTTGAC-3'	-15	Upper primer to amplify LexA4 probe
AR39	5'-TTATGTAGAACAAGTGTCTTATGC-3'	-35	Upper primer to amplify LexA3 probe
AR40	5'-TTGTGATTGACAGGGAAGG-3'	-68	Upper primer to amplify LexA2 probe
AR61	5'-GAATTCATGAACTGACCACCAGACAGCG-3'	+1	Upper primer for directed cloning of <i>D. ethenogenes lexA</i> in pGEX4T1
AR62	5'-GCCGAGGCATGACCCAGCTGGGGC-3'	-515	Upper primer to amplify <i>uvrA</i> promoter region
AR63	5'-TTTTGAGGTTGTGTTCCGCGGCACC-3'	+51	Lower primer to amplify <i>uvrA</i> promoter region
AR64	5'-TAGTGGAGAAATACCACTCTCAGGC-3'	-531	Upper primer to amplify <i>nuvA</i> promoter region
AR65	5'-TTGCCGGGACAGAGTTGAC-3'	+40	Lower primer to amplify <i>nuvA</i> promoter region
AR66	5'-GGGTGAACTGGTAGACAAGTTCCGC-3'	-525	Upper primer to amplify <i>nuvB</i> promoter region
AR67	5'-GGCCAGTTTCCCGGATATAAGACGC-3'	+42	Lower primer to amplify <i>nuvB</i> promoter region
AR68	5'-AGCCGGTACCGCCGGATGCACCTGC-3'	-532	Upper primer to amplify <i>recN</i> promoter region
AR69	5'-CCAGGTGATATCCTCAATAATGCCG-3'	+53	Lower primer to amplify <i>recN</i> promoter region
AR81	5'-ACTCATGATCATAACCTCCAACAGG-3'	-161	Upper primer to amplify <i>B. subtilis dinR</i> promoter region
AR82	5'-GGACGCAAGCCGACAGCCTCTCCG-3'	+116	Lower primer to amplify <i>B. subtilis dinR</i> promoter region

<sup>a</sup> Added restriction sites are shown in italics. The position of either the 5' or 3' end of the oligonucleotide with respect to the proposed translational starting point of each *D. ethenogenes* gene is shown except for AR81 and AR82, where the position of either the 5' or 3' end of the oligonucleotide with respect to the translational starting point of the *B. subtilis dinR* gene is shown. Primers AR41 to AR60 and AR76 are derivatives of AR39 but with the desired nucleotide changes.

*lexA* translational start codon, permitting fusion of the protein in frame with the GST protein. The 650-bp PCR fragment was cloned into pGEM-T, generating plasmid pUA970. Afterwards, pUA970 was digested with *EcoRI* and *SalI*, and the resulting 0.65-kb fragment was cloned into pGEX4T1 (Amersham Pharmacia Biotech), creating the GST-LexA fusion.

This plasmid, designated pUA971, was transformed into *Escherichia coli* BL21 Codon Plus cells (Stratagene) to overproduce the protein. An overnight culture of the BL21 Codon Plus strain containing plasmid pUA971 was diluted in one-half liter of Luria-Bertani (LB) medium and incubated at 37°C until it reached an optical density at 600 nm of 0.8. At this time isopropylthiogalactopyranoside (IPTG) was added (1 mM final concentration), and the culture was incubated for 4 more hours. Afterwards, cells were collected by centrifugation for 15 min at 5,000 rpm.

The bacterial pellet was resuspended in phosphate-buffered saline (PBS) (10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.7 mM KH<sub>2</sub>PO<sub>4</sub>, 137 mM NaCl, 2.7 mM KCl [pH 7.4]) containing complete Mini protease inhibitor cocktail (Roche) and sonicated to break the cells. The cell lysate was centrifuged at 15,000 rpm for 30 min, and the supernatant containing the soluble GST-LexA protein fusion was incubated for 2 h at 4°C with glutathione-Sepharose 4B beads (Amersham Pharmacia Biotech), previously equilibrated in PBS. The beads were washed twice with PBS-0.1% Triton X-100 and three more times with PBS to eliminate contaminant proteins. The washed glutathione-Sepharose beads containing bound GST-LexA protein were equilibrated in 0.1 M Tris-HCl (pH 8)-0.1 M NaCl buffer, and GST-LexA was eluted in the same buffer containing 20 mM glutathione (Sigma), resulting in about 85% pure GST-LexA (Fig. 2).

The GST-LexA fusion protein has the amino acid sequence

Leu-Val-Pro-Arg-Gly-Ser between GST and LexA. This motif is recognized by the thrombin protease, which cleaves the Arg-Gly peptide bond. In this case, proteolysis results in *D. ethenogenes* LexA protein with the peptide Gly-Ser-Pro-Glu-Phe attached to the N terminus. GST-LexA bound to glutathione-Sepharose was incubated at room temperature for 16 h with 50 U of thrombin protease (Amersham Pharmacia Biotech) in 1 ml of PBS. The liquid phase containing LexA was collected, and a portion was separated by electrophoresis on a denaturing 13% polyacrylamide gel. The purity of the *D. ethenogenes* LexA protein was greater than 95% (Fig. 2).

**Determination of *D. ethenogenes* LexA binding motif.** Analysis of the DNA contig containing *D. ethenogenes lexA* revealed that it was located between the *rpsT* and *fucA* genes (Fig. 3A). From the TAA translational stop codon of *rpsT* to the putative TTG start codon of *lexA*, there are 104 bp that presumably contain the LexA binding site. Hence, a PCR fragment containing this sequence was amplified with primers AR31 and AR33DIG (Table 1). The resulting digoxigenin-labeled probe (LexA1) was used as a target to observe whether purified *D. ethenogenes* LexA was able to bind its own promoter in electrophoretic mobility shift assay experiments carried out as previously described (6).

Basically, 20- $\mu$ l reaction mixtures containing 10 ng of digoxigenin-labeled DNA probe and 25 ng (50 nM) of *D. ethenogenes* LexA were incubated in binding buffer: 10 mM HEPES NaOH (pH 8), 10 mM Tris-HCl (pH 8), 5% glycerol, 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 1  $\mu$ g of bulk carrier DNA, and 50  $\mu$ g of bovine serum albumin per ml. After 30 min of incubation at 30°C, the mixture was loaded onto a 5% non-denaturing Tris-glycine-polyacrylamide gel. Thus, *D. ethenogenes* LexA was able to shift the mobility of the LexA1 fragment (Fig.



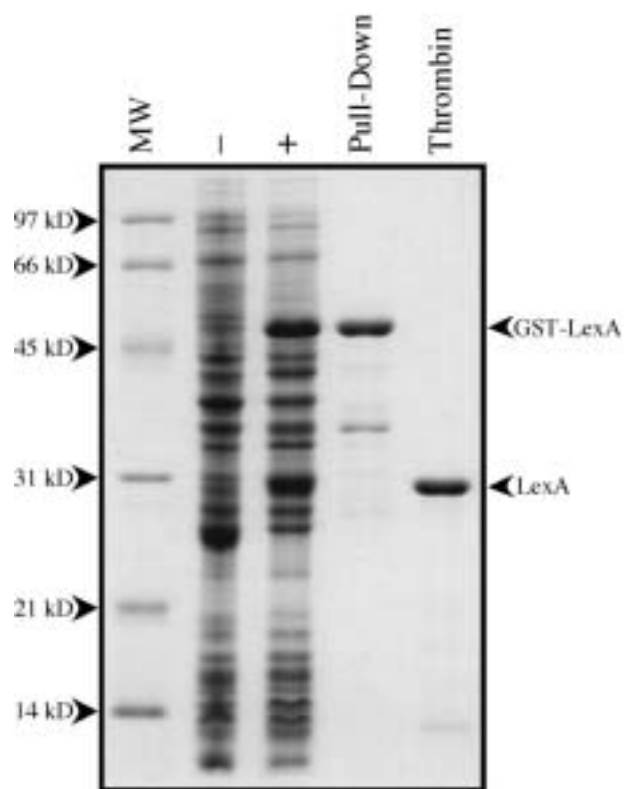


FIG. 2. Purification of the *D. ethenogenes* LexA protein. *D. ethenogenes* LexA was purified to greater than 95%. Each line shown in the denaturing 13% polyacrylamide gel represents each of the different protein purification steps employed: crude extract of BL21 Codon Plus/pUA971 (-); crude extract of BL21 Codon Plus/pUA971 induced with 1 mM IPTG (+); *D. ethenogenes* GST-LexA affinity purified with glutathione-Sepharose 4B (pull-down); purified *D. ethenogenes* LexA protein after thrombin protease digestion of GST-LexA fusion (thrombin). Lane MW, molecular size markers.

3B). Furthermore, the retarded band was abolished when unlabeled LexA1 was used as a competitor (100-fold molar excess over the digoxigenin-labeled probe). However, the shifted band could not be eliminated when nonspecific DNA competitors were used (Fig. 3B), suggesting that the binding of the *D. ethenogenes* LexA protein to its promoter is specific. It is important to note that the GST-LexA fusion protein retained its DNA binding activity in electrophoretic mobility shift assay experiments (Fig. 3B).

To establish the bounds of the LexA binding site, serial deletion fragments of the upstream promoter region were analyzed in electrophoretic mobility shift assay experiments with purified LexA. The DNA probes LexA1, LexA2 and LexA3 exhibited retarded bands in the presence of LexA protein, but the mobility of the LexA4 DNA was not shifted (Fig. 3C). These results indicate that the LexA binding motif is located within the sequence 35 bp upstream of the *lexA* translation start codon. An examination of this region revealed the presence of the inverted repeat GAACAAGTGTTTC 28 bp upstream of the putative TTG translational start codon (Fig. 3A). This sequence matches the DinR recognition motif (formerly Cheo-box), CGAACN<sub>4</sub>GTTTCG (23). Changing the GAAC submotif to TTTT in the LexA3 fragment abolished LexA

protein binding, demonstrating that this tetranucleotide is part of the *D. ethenogenes* LexA binding site (Fig. 3C).

The importance of the rotational orientation of the inverted repeats relative to one another was examined by introducing additional nucleotides into the center of the GAACN<sub>4</sub>GTTTC motif, lengthening the linker region between the inverted repeats (Fig. 3A). With modified LexA3 PCR primers (AR58 to AR60, Table 1), one to three adenines were inserted at position -21 relative to the translational start codon (Fig. 3A). The introduction of a single adenine resulted in a sequence that retained some residual binding of the LexA protein (Fig. 3D). However, the introduction of two or three nucleotides into the linker region completely abolished LexA binding (Fig. 3D), consistent with studies of LexA binding sites in other microorganisms (6).

To gain insight into the importance of each of the 12 bases of this imperfect palindrome and the neighboring nucleotides in the LexA binding process, single nucleotide changes were introduced into the LexA3 probe by PCR amplification with primers AR41 to AR57 and AR76 (Table 1). The changes introduced were the least permissive based on the thorough characterization of the DinR binding site performed by Winterling et al. (23). Changes introduced in the main motifs GAAC and GTTC severely decreased or eliminated LexA binding (Fig. 4A). Also critical for binding were the flanking bases of the extended direct repeat, the adenine residues at positions -29 and -24, and the thymine residues at positions -16 and -21 (relative to the TTG translational start codon). These results indicated that the tetranucleotide GAAC and the inverted repeat GTTC are critical to LexA binding, consistent with previous studies of the DinR recognition site (23). Collectively, these results demonstrate that the sequence AGAACN<sub>4</sub>GTTCT is the specific recognition site for the *D. ethenogenes* LexA repressor.

**The *Bacillus subtilis* DinR repressor recognizes the *D. ethenogenes* LexA binding site.** The electrophoretic mobility shift assay experiments performed with the wild-type and mutated variants of the *D. ethenogenes* *lexA* operator revealed that A and T at positions 1 and 14, respectively, of the AGAACN<sub>4</sub>GTTCT sequence are important bases for recognition by LexA. Nevertheless, the consensus LexA-box sequence established for *B. subtilis* and other gram-positive bacteria is CGAACN<sub>4</sub>GTTTCG (23). Indeed, these two terminal nucleotides are part of only a few gram-positive LexA-boxes, especially the T at position 14, which is observed in only 1 of 20 boxes, the *Streptomyces lividans* *recA* operator (23).

In order to confirm that the LexA-box of *D. ethenogenes* is the same as that of gram-positive species, we analyzed the capacity of the DinR repressor to bind to the *D. ethenogenes* *lexA* operator. To examine binding, digoxigenin-labeled LexA1 DNA was incubated in the presence of purified *B. subtilis* DinR protein or *D. ethenogenes* LexA with or without DNA competitors containing the *B. subtilis* DinR- or *D. ethenogenes* LexA-box. The results (Fig. 4B) show that the DinR protein bound to the *D. ethenogenes* *lexA* operator and that binding was abolished by competition with unlabeled DNA containing either the *D. ethenogenes* LexA-box or the *B. subtilis* DinR-box. All these results clearly indicate that *D. ethenogenes* LexA protein specifically recognizes the motif GAACN<sub>4</sub>GTTTC, as does DinR. Hence, *D. ethenogenes* is the first bacterium outside of

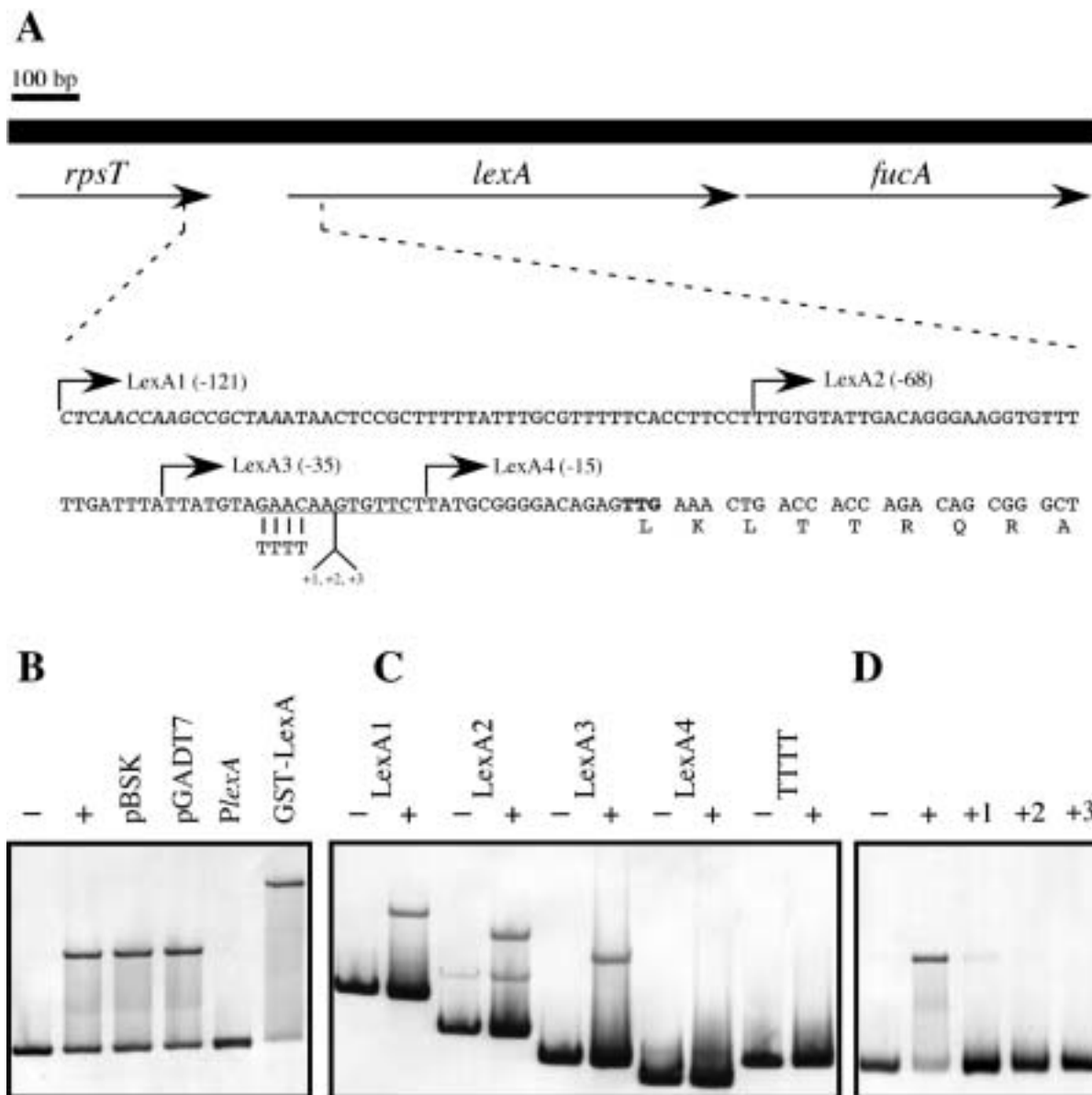


FIG. 3. (A) Schematic representation of the *D. ethenogenes* chromosomal region containing *lexA*, which is located between the *rpsT* and *fucA* genes. The sequence of the *lexA* promoter region and the first nine amino acid residues of LexA are shown. The 3' end of the *rpsT* gene is in italic type, the putative TTG translational start codon of *lexA* is in bold type, and the homologue of the DinR-box is underlined. The start points of each fragment of the *lexA* promoter (LexA1, LexA2, LexA3, and LexA4) used in electrophoretic mobility shift assay experiments are indicated by arrows, and the relative distances to the TTG are in parentheses. (B) Specific binding of *D. ethenogenes* LexA protein. Electrophoretic mobility shift assay experiments were performed with different DNAs as competitors. LexA1 fragment was incubated at 30°C for 30 min in the absence (-) and in the presence (+) of 25 ng of pure LexA protein (50 nM). To show the specificity of LexA binding, the LexA1 fragment was incubated in the presence of *D. ethenogenes* LexA protein and the following unlabeled DNA competitors: 2 µg of pBSK, 2 µg of pGADT7, or 1 µg of unlabeled LexA1 DNA fragment (*PlexA*). In the last lane, LexA1 probe was incubated in the presence of the *D. ethenogenes* GST-LexA fusion protein. (C) Establishment of the limits of the *D. ethenogenes* LexA binding site. LexA1, LexA2, LexA3, LexA4, and a derivative of the LexA3 fragment where GAAC was changed to TTTT (TTTT) were incubated in the absence (-) or presence (+) of purified LexA. (D) Electrophoretic mobility shift assays of LexA3 fragments where one (+1), two (+2), or three (+3) adenine residues were inserted at position -21 with respect to the initiation start codon. Different probes were incubated with LexA protein and loaded on a native Tris-glycine-5% polyacrylamide gel. The LexA3 probe alone (-) and the same fragment incubated in the presence of LexA (+) acted as negative and positive controls, respectively.

the gram-positive phylum whose LexA repressor binds to the DinR-box.

**The LexA regulon is different in *B. subtilis* than in *D. ethenogenes*.** In *Escherichia coli*, at least 40 genes are directly under the negative control of the LexA repressor (3, 5). The amino

acid sequences of each of these *E. coli* SOS genes were used as queries in BLAST searches of the unfinished *D. ethenogenes* genome to seek homologous genes in this bacterial species. Homologues of *recA*, *uvrA*, *ruvA*, *ruvB*, *uvrB*, *uvrD*, *ssb*, *dinB*, and *recN* were identified. However, homologues to the rest of





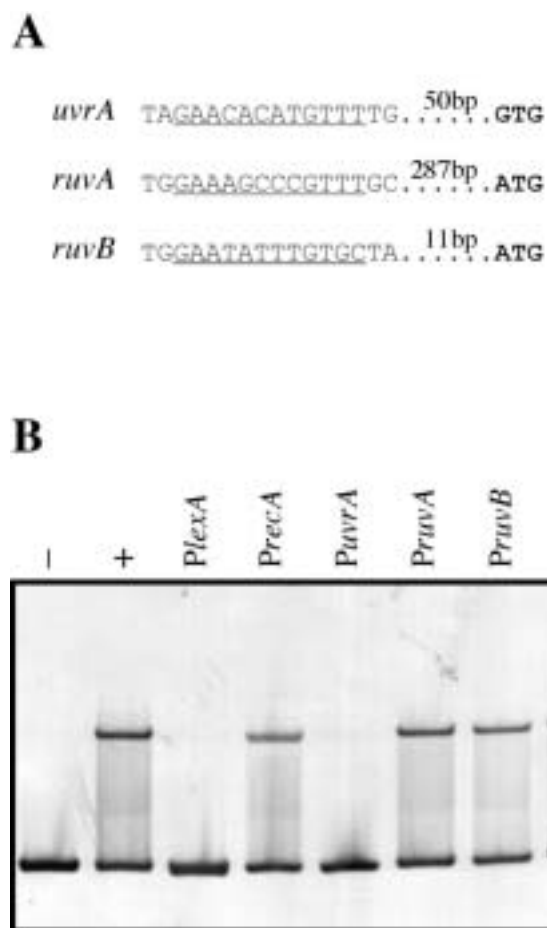


FIG. 5. (A) Upstream sequences of *D. ethenogenes* *uvrA*, *ruvA*, and *ruvB* genes presenting putative LexA binding sites. Potential LexA binding sites are underlined, and the putative translation initiation codons are depicted in bold letters, with the distance between them indicated. (B) Electrophoretic mobility shift assay of the LexA1 fragment incubated with purified *D. ethenogenes* LexA in the presence of different DNA competitors. LexA1 fragment was incubated at 30°C for 30 min in the absence (-) or presence (+) of pure LexA protein. Simultaneous reactions contained the LexA1 fragment, *D. ethenogenes* LexA protein, and one of six unlabeled DNA competitors (100-fold molar excess) containing the promoter regions and potential LexA binding sites of genes *lexA* (*PlexA*), *recA* (*PrecA*), *uvrA* (*PuvrA*), *ruvA* (*PruvA*), *ruvB* (*PruvB*), and *recN* (*PrecN*).

the *E. coli* LexA-regulated genes could not be identified, possibly because the *D. ethenogenes* genome has not been completely sequenced or because they simply do not exist in this microorganism. Subsequently, the promoters of the genes identified were analyzed to find potential LexA binding sites.

One clear motif, similar to the 12 bp found in the *lexA* promoter, was identified 64 bp upstream of the putative GTG translational start codon of the *uvrA* gene (Fig. 5A). Potential motifs were also found in the promoter regions of the *ruvA* and *ruvB* genes 301 and 25 bp upstream of their putative translation initiation codons, respectively (Fig. 5A). Examination of the remaining SOS gene homologues failed to reveal any potential LexA binding sites. Strikingly, at first glance, *recA* does not seem to be regulated by LexA because no LexA binding motif has been identified in its promoter region. To determine

if these potential LexA binding sites were functional, PCR-amplified fragments of the promoter regions from *uvrA*, *ruvA*, *ruvB*, *recN*, and *recA* were used as competitors in band shift assays with the LexA1 DNA fragment as the labeled probe. While a 100-fold molar excess of DNA containing the *uvrA* promoter was able to abolish the band belonging to the DNA-LexA complex, the shifted band did not disappear when the same amount of DNA containing the *recA*, *ruvA*, *ruvB*, or *recN* promoter was used as the competitor (Fig. 5B). This result clearly indicates that LexA binds the *uvrA* promoter, suggesting that the motif AGAACACATGTTTT is indeed a *D. ethenogenes* LexA recognition site.

The lack of a visible LexA binding motif in the *recA* promoter sequence is consistent with the inability of the *recA* promoter fragment to abolish the retarded band. This confirms that no LexA binding motif is present in the *recA* promoter region and indicates that the *recA* gene in *D. ethenogenes* is not under the direct control of the LexA repressor. This has also been observed for *Deinococcus radiodurans*, where the *recA* gene is induced by gamma radiation in a LexA-independent manner. However, it has been reported that *D. radiodurans* possesses a second *lexA* gene that might be involved in the direct regulation of *recA* (18). Our results demonstrate that *D. ethenogenes* LexA does not bind to the *recA* promoter, thus constituting the first report of a *recA* gene that is not regulated directly by the LexA repressor.

On the basis of the signature sequences found in various proteins, it has been recently proposed that the green nonsulfur bacterial phylum, to which *D. ethenogenes* belongs, evolved from a universal ancestor later than the gram-positive bacteria but before the cyanobacteria (9). The *Deinococcus-Thermus* group seems to have appeared before the green nonsulfur bacteria, although this order has not been definitively fixed (9). Our results demonstrating that the LexA-box sequences of *B. subtilis* and *D. ethenogenes* are nearly identical and that the DinR repressor from a gram-positive organism binds to the *D. ethenogenes* LexA box give support to the hypothesis that a close phylogenetic relationship exists between the gram-positive and green nonsulfur bacteria. However, conclusions about the order of evolution of the green nonsulfur bacteria and the *Deinococcus-Thermus* group cannot be made on this basis due to the current lack of information on the LexA binding motif for members of the latter group.

Despite having similar LexA binding sequences, gram-positive bacteria and *D. ethenogenes* display significant differences in the composition of their LexA regulons. A BLAST search of GenBank indicates that genes with an upstream LexA binding sequence are more numerous in the gram-positive bacteria *B. subtilis*, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, and *Clostridium acetobutylicum* than in *D. ethenogenes* (Table 2). The reasons for these differences are not clear but could be related to the environmental conditions in which each of these organisms has evolved.

Finally, performing a TBLASTN analysis with *D. ethenogenes* LexA protein as a query at the DOE Joint Genome Institute, a LexA homologue was found in the unfinished genome of the photosynthetic bacterium *Chloroflexus aurantiacus*, which is also a member of the green nonsulfur bacteria. The putative *lexA* gene is found at the 5' end of a short contig (694), and thus the N terminus of the protein and the potential

TABLE 2. Comparison of the presence of a LexA binding sequence upstream of different genes putatively belonging to the LexA regulon from several gram-positive bacteria, *D. ethenogenes*, and *E. coli*<sup>a</sup>

Gene	<i>B. subtilis</i>	<i>S. aureus</i>	<i>M. tuberculosis</i>	<i>C. acetobutylicum</i>	<i>D. ethenogenes</i>	<i>E. coli</i>
<i>recA</i>	+	+	+	+	–	+
<i>lexA</i>	+	+	+	+	+	+
<i>ruvAB</i>	+	–	+	–	– <sup>b</sup>	+
<i>uvrA</i>	–	–	–	–	+	+
<i>uvrB</i>	+	–	–	+	–	+
<i>uvrD</i>	+	–	–	–	–	+
<i>ssb</i>	–	+	–	–	–	+
<i>recN</i>	–	–	–	–	–	+

<sup>a</sup> The presence or absence of a LexA binding site is represented as + and –, respectively. Searches of the whole genomes of several bacteria were performed with the RCGScanner program, which uses a recursive method (20) to identify a given nucleotide sequence as well as its neighboring open reading frames (Erill et al., submitted).

<sup>b</sup> The *ruvA* and *ruvB* genes do not form an operon in *D. ethenogenes*.

LexA-box could not be identified. However, the truncated LexA protein of *C. aurantiacus* contains the helix-turn-helix motif, which is involved in DNA binding (7), showing 87% similarity with the *D. ethenogenes* helix-turn-helix domain. This fact suggests that *C. aurantiacus* LexA might also recognize the GAACN<sub>4</sub>GTTC as in gram-positive bacteria. The availability of the assembled genome of this organism will be very useful for understanding the evolutionary and phylogenetic history of the green nonsulfur bacteria, for confirming the homogeneity of this phylum with respect to the LexA regulatory sequence, and for determining the composition of the LexA regulon. These data will also be useful for testing the hypothesis, cited above, that a relationship exists between environmental conditions and the genes which are under control of the LexA protein in bacteria.

This work was funded by grants BMC2001-2065 from the Ministerio de Ciencia y Tecnología (MCyT) de España and 2001SGR-206 from the Departament d'Universitats, Recerca i Societat de la Informació (DURSI) de la Generalitat de Catalunya. Antonio R. Fernández de Henestrosa is a recipient of a postdoctoral reincorporation contract from the MCyT.

We acknowledge Joan Ruiz, Susana Escribano, and Pilar Cortés for excellent technical assistance. Preliminary *D. ethenogenes* sequence data were obtained from the Institute for Genomic Research website at <http://www.tigr.org>. The *Chlorefexus aurantiacus* partial *lexA* sequence was provided free by the U.S.-DOE Joint Genome Institute (<http://www.jgi.doe.gov/>). We are deeply grateful to Roger Woodgate and Kevin Winterling for generously providing us with *B. subtilis* genomic DNA, *B. subtilis* DinR, and *E. coli* LexA protein.

#### REFERENCES

- Brent, R., and M. Ptashne. 1981. Mechanism of action of the *lexA* gene product. Proc. Natl. Acad. Sci. USA **78**:4204–4208.
- Brooks, P. C., F. Movahedzadeh, and E. O. Davis. 2001. Identification of some DNA damage-inducible genes of *Mycobacterium tuberculosis*: apparent lack of correlation with LexA binding. J. Bacteriol. **183**:4459–4467.
- Courcelle, J., A. Khodursky, B. Peter, P. O. Brown, and P. C. Hanawalt. 2001. Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. Genetics **158**:41–64.
- Eisen, J. A., and P. C. Hanawalt. 1999. A phylogenomic study of DNA repair genes, proteins, and processes. Mutat. Res. **435**:171–213.
- Fernandez de Henestrosa, A. R., T. Ogi, S. Aoyagi, D. Chafin, J. J. Hayes, H. Ohmori, and R. Woodgate. 2000. Identification of additional genes belonging to the LexA-regulon in *Escherichia coli*. Mol. Microbiol. **35**:1560–1572.
- Fernandez de Henestrosa, A. R., E. Rivera, A. Tapias, and J. Barbe. 1998. Identification of the *Rhodobacter sphaeroides* SOS box. Mol. Microbiol. **28**:991–1003.
- Fogh, R. H., G. Oettleben, H. Ruterjans, M. Schnarr, R. Boelens, and R. Kaptein. 1994. Solution structure of the LexA repressor DNA binding domain determined by <sup>1</sup>H NMR spectroscopy. EMBO J. **13**:3936–3944.
- Friedberg, E. C., G. C. Walker, and W. Siede. 1995. DNA repair and mutagenesis. American Society for Microbiology, Washington, D.C.
- Gupta, R. S. 2000. The natural evolutionary relationships among prokaryotes. Crit. Rev. Microbiol. **26**:111–131.
- Hugenholtz, P., B. M. Goebel, and N. R. Pace. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. **180**:4765–4774.
- Khil, P. P., and R. D. Camerini-Otero. 2002. Over 1000 genes are involved in the DNA damage response of *Escherichia coli*. Mol. Microbiol. **44**:89–105.
- Koch, W. H., and R. Woodgate. 1998. The SOS response, p. 107–134. In J. A. Nickoloff and M. F. Hoekstra (ed.), DNA damage and repair: DNA repair in prokaryotes and lower eukaryotes, 1st ed. Humana Press, Totowa, N.J.
- Little, J. W. 1993. LexA cleavage and other self-processing reactions. J. Bacteriol. **175**:4943–4950.
- Little, J. W., S. H. Edmiston, L. Z. Pacelli, and D. W. Mount. 1980. Cleavage of the *Escherichia coli* LexA protein by the RecA protease. Proc. Natl. Acad. Sci. USA **77**:3225–3229.
- Little, J. W., D. W. Mount, and C. R. Yanisch-Perron. 1981. Purified LexA protein is a repressor of the *recA* and *lexA* genes. Proc. Natl. Acad. Sci. USA **78**:4199–4203.
- Luo, Y., R. A. Pfuetzner, S. Mosimann, M. Paetzel, E. A. Frey, M. Cherney, B. Kim, J. W. Little, and N. C. Strynadka. 2001. Crystal structure of LexA: a conformational switch for regulation of self-cleavage. Cell **106**:585–594.
- Maymó-Gatell, X., Y. Chien, J. M. Gossett, and S. H. Zinder. 1997. Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene. Science **276**:1568–1571.
- Narumi, I., K. Satoh, M. Kikuchi, T. Funayama, T. Yanagisawa, Y. Kobayashi, H. Watanabe, and K. Yamamoto. 2001. The LexA protein from *Deinococcus radiodurans* is not involved in RecA induction following gamma irradiation. J. Bacteriol. **183**:6951–6956.
- Radman, M. 1974. Phenomenology of an inducible mutagenic DNA repair pathway in *Escherichia coli*: SOS repair hypothesis, p. 128–142. In L. Prakash, F. Sherman, M. Miller, C. W. Lawrence, and H. W. Tabor (ed.), Molecular and environmental aspects of mutagenesis. Charles C Thomas, Springfield, Ill.
- Rodionov, D. A., A. A. Mironov, and M. S. Gelfand. 2001. Transcriptional regulation of pentose utilisation systems in the *Bacillus/Clostridium* group of bacteria. FEMS Microbiol. Lett. **205**:305–314.
- Tapias, A., and J. Barbe. 1999. Regulation of divergent transcription from the *uvrA-ssb* promoters in *Sinorhizobium meliloti*. Mol. Gen. Genet. **262**:121–130.
- Walker, G. C. 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. Microbiol. Rev. **48**:60–93.
- Winterling, K. W., D. Chafin, J. J. Hayes, A. S. Levine, R. E. Yasbin, and R. Woodgate. 1998. The *Bacillus subtilis* DinR binding site: redefinition of the consensus sequence. J. Bacteriol. **180**:2201–2211.
- Winterling, K. W., A. S. Levine, R. E. Yasbin, and R. Woodgate. 1997. Characterization of DinR, the *Bacillus subtilis* SOS repressor. J. Bacteriol. **179**:1698–1703.

## **ARTICLE II**

## Characterization of a New LexA Binding Motif in the Marine Magnetotactic Bacterium Strain MC-1

Antonio R. Fernández de Henestrosa,<sup>1\*</sup> Jordi Cuñé,<sup>1</sup> Gerard Mazón,<sup>1</sup> Bradley L. Dubbels,<sup>2</sup> Dennis A. Bazylinski,<sup>2</sup> and Jordi Barbé<sup>1,3\*</sup>

*Departament Genètica i Microbiologia*<sup>1\*</sup> and *Centre de Recerca en Sanitat Animal (CRESA)*,<sup>3</sup>  
*Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain, and*  
*Department of Microbiology, Iowa State University, Ames, Iowa 50011*<sup>2</sup>

Received 6 December 2002/Accepted 23 April 2003

**MC-1 is a marine, magnetotactic bacterium that is phylogenetically associated with the alpha subclass of the *Proteobacteria* and is the first and only magnetotactic coccus isolated in pure culture to date. By using a TBLASTN search, a *lexA* gene was identified in the published genome of MC-1; it was subsequently cloned, and the protein was purified to >90% purity. Results from reverse transcription-PCR analysis revealed that the MC-1 *lexA* gene comprises a single transcriptional unit with two open reading frames encoding proteins of unknown function and with a *rumA*-like gene, a homologue of the *Escherichia coli umuD* gene. Mobility shift assays revealed that this LexA protein specifically binds both to its own promoter and to that of the *umuDC* operon. However, MC-1 LexA does not bind to the promoter regions of other genes, such as *recA* and *uvrA*, that have been previously reported to be regulated by LexA in bacterial species belonging to the alpha subclass of the *Proteobacteria*. Site-directed mutagenesis of both the *lexA* and *umuDC* operator regions demonstrated that the sequence CCTN<sub>10</sub>AGG is the specific target motif for the MC-1 LexA protein.**

The term “SOS response” was introduced in the early 1970s by Miroslav Radman, who described a network of genes induced as a consequence of DNA damage (29). In *Escherichia coli*, this regulon consists of more than 40 genes involved in DNA replication, DNA repair, mutagenesis, and control of the cell cycle (6, 11, 15, 17). Under normal conditions the SOS network remains repressed by the product of the *lexA* gene, its negative regulator (3, 19, 23, 36). LexA binds specifically to an operator region located at the promoters of the SOS genes, reducing their expression to basal level (15, 36). As a conse-

quence of DNA damage or stalling of the replication fork, the positive regulator of the SOS regulon, RecA, achieves its active conformation (RecA\*) after binding to the generated single-stranded DNA, thereby creating a nucleofilament. RecA\* mediates the autohydrolysis of the LexA repressor between residues Ala<sup>84</sup> and Gly<sup>85</sup>, giving rise to two nonfunctional products, the N-terminal DNA binding domain and the C-terminal domain, the latter of which is involved in the dimerization of the repressor (21, 22). During this autocatalytic cleavage, residues Ser<sup>119</sup> and Lys<sup>156</sup> are responsible for the

TABLE 1. Bacterial strains and plasmids used in this work

Strain or plasmid	Relevant features	Source or reference
<b>Strains</b>		
MC-1	Cocoid, marine, magnetotactic bacterium	7
<i>Escherichia coli</i>		
DH5 $\alpha$	<i>supE4 <math>\Delta</math>lacU169</i> ( $\phi$ 80 <i>lacZ</i> $\Delta$ M15) <i>hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	Clontech
BL21(DE3) Codon plus	<i>E. coli</i> B F <sup>-</sup> <i>ompT hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) dcm<sup>+</sup> Tc<sup>r</sup> gal<sup>+</sup> (DE3) endA Hte[<i>argU ileY leuW Cam<sup>r</sup></i>]</i>	Stratagene
<b>Plasmids</b>		
pGEM-T	PCR cloning vector; Ap <sup>r</sup>	Promega
pET15b	His <sub>6</sub> tag expression vector; Ap <sup>r</sup>	Novagen
pUA1011	pGEM-T derivative carrying the 678-bp PCR fragment containing the coding region of the MC-1 <i>lexA</i> gene; amplified with primers JC5 and JC6	This work
pUA1013	pET15b derivative carrying the 678-bp <i>NdeI-BamHI</i> fragment containing the coding region of MC-1 <i>lexA</i>	This work

\* Corresponding author. Mailing address for Jordi Barbé: Universitat Autònoma de Barcelona, Departament de Genètica i Microbiologia, Unitat de Microbiologia, Edifici C, Bellaterra, 08193 Barcelona, Spain. Phone: 34 93 581 1837. Fax: 34 93 581 2387. E-mail: jordi.barbe@uab.es. Present address for Antonio R. Fernández de Henestrosa: Centro de Biología Molecular Severo Ochoa (CSIC), Cantoblanco, 28049 Madrid, Spain. Phone: 34 91 397 3981. Fax: 34 91 397 4799. E-mail: arfernandez@cbm.uam.es.

TABLE 2. Oligonucleotide primers used in this work

Primer	Sequence <sup>a</sup>	Position <sup>b</sup>	Application
JC5	5'-CATATGAAACCAGGACGGAGACCCACAGAAGGG-3'	+1	Upper primer for cloning coding region of <i>Magnetococcus</i> <i>lexA</i> gene and amplifying <i>lexA</i> fragment in RT-PCR experiments
JC6	5'-GGATCCGTCGACTTATTCATTATTGTTGGCGGCCCGTGG-3'	+672	Lower primer for cloning coding region of <i>Magnetococcus</i> <i>lexA</i> gene and amplifying <i>lexA</i> fragment in RT-PCR experiments
JC7	5'-GAGGTGGAAGGCCATTCTGTTGCC-3'	-261	Upper primer to amplify LexA1 probe
JC8DIG	5'-DIG-CCAGTTTCTGCACCTGTTTCATGGGC-3'	+468	Lower primer to amplify DIG-labeled probe fragments of <i>lexA</i> promoter
JC8	5'-CCAGTTTCTGCACCTGTTTCATGGGC-3'	+468	Lower primer to amplify unlabeled probe fragments of <i>lexA</i> promoter
JC13	5'-CAGATCCATATTGGTACTTACCG-3'	-178	Upper primer to amplify LexA2 probe
JC14	5'-AACTTACACCTTATTAATAAGG-3'	-100	Upper primer to amplify LexA3 probe
JC15	5'-TAAAATAAGGTCACAAGAGAG-3'	-86	Upper primer to amplify LexA4 probe
JC16	5'-AACTTACGGGTATTAATAAGG-3'	-100	JC14 with ACCT-to-GGGG change in LexA3 probe
JC22	5'-AACTTACACCTTATTAATAAGGTCAC-3'	-100	JC14 with addition of an A at position -84 in LexA3 probe
JC23	5'-AACTTACACCTTATTAATAAGGTCAC-3'	-100	JC14 with addition of two A's at position -84 in LexA3 probe
JC24	5'-AACTTACACCTTATTAATAAGGTCAC-3'	-100	JC14 with addition of three A's at position -84 in LexA3 probe
JC26	5'-AACTTACGCCTTATTAATAAGG-3'	-100	JC14 with A-to-G change at position -93 in LexA3 probe
JC27	5'-AACTTACAGCTTATTAATAAGG-3'	-100	JC14 with C-to-G change at position -92 in LexA3 probe
JC28	5'-AACTTACACCTTATTAATAAGG-3'	-100	JC14 with C-to-A change at position -91 in LexA3 probe
JC29	5'-AACTTACACCTTATTAATAAGG-3'	-100	JC14 with T-to-A change at position -90 in LexA3 probe
JC30	5'-AACTTACACCTTATTAATAAGG-3'	-100	JC14 with T-to-G change at position -89 in LexA3 probe
JC31	5'-AACTTACACCTTATTAATAAGG-3'	-100	JC14 with A-to-G change at position -88 in LexA3 probe
JC32	5'-AACTTACACCTTAGTAAATAAGG-3'	-100	JC14 with T-to-G change at position -87 in LexA3 probe
JC33	5'-AACTTACACCTTATTAATAAGG-3'	-100	JC14 with T-to-G change at position -86 in LexA3 probe
JC34	5'-AACTTACACCTTATTAATAAGGTCAC-3'	-100	JC14 with A-to-G change at position -85 in LexA3 probe
JC35	5'-AACTTACACCTTATTAATAAGGTCAC-3'	-100	JC14 with A-to-G change at position -84 in LexA3 probe
JC36	5'-AACTTACACCTTATTAATAAGGTCAC-3'	-100	JC14 with A-to-G change at position -83 in LexA3 probe
JC37	5'-AACTTACACCTTATTAATAAGGTCAC-3'	-100	JC14 with A-to-G change at position -82 in LexA3 probe
JC38	5'-AACTTACACCTTATTAATAAGGTCAC-3'	-100	JC14 with T-to-G change at position -81 in LexA3 probe
JC39	5'-AACTTACACCTTATTAATAAGGTCACAAGAGAG-3'	-100	JC14 with A-to-G change at position -80 in LexA3 probe
JC40	5'-AACTTACACCTTATTAATAAGGTCACAAGAGAG-3'	-100	JC14 with A-to-G change at position -79 in LexA3 probe
JC41	5'-AACTTACACCTTATTAATAAGGTCACAAGAGAG-3'	-100	JC14 with G-to-C change at position -78 in LexA3 probe
JC42	5'-AACTTACACCTTATTAATAAGGTCACAAGAGAG-3'	-100	JC14 with G-to-C change at position -77 in LexA3 probe
JC43	5'-AACTTACACCTTATTAATAAGGTCACAAGAGAG-3'	-100	JC14 with T-to-G change at position -76 in LexA3 probe
JC11	5'-GGTTTCCGTGAAAAGTCTCATAACCGCG-3'	-308	Upper primer to amplify <i>umuDC</i> promoter region
JC12	5'-AGCCGGTACCGCCGATGCACCTGC-3'	+228	Lower primer to amplify <i>umuDC</i> promoter region
JC46	5'-ACATCCCATCAATCCGACCATGG-3'	-342	Upper primer to amplify <i>recA</i> promoter region
JC47	5'-CAGTGCTTGTCTTATCTTATCC-3'	+32	Lower primer to amplify <i>recA</i> promoter region
JC48	5'-CCAATGGGGCTCAACTGCGTGGGG-3'	-271	Upper primer to amplify <i>uvrA</i> promoter region
JC49	5'-CAGGTTGTTCCTCCGGCACCCAGG-3'	+48	Lower primer to amplify <i>uvrA</i> promoter region
JC50	5'-AACCGCCCATTTGCGGTGCCCGCG-3'	-270	Upper primer to amplify <i>recN</i> promoter region
JC51	5'-CGACATGCTTCAATGGTGAGTTGACG-3'	+30	Lower primer to amplify <i>recN</i> promoter region
JC19	5'-ACCGGATACCTAATTTATTAGG-3'	-55	Upper primer to amplify <i>umuD</i> probe
JC12DIG	5'-AGCCGGTACCGCCGATGCACCTGC-3'	+228	Lower primer to amplify DIG-labeled <i>umuD</i> probe
JC57	5'-ACCGGATACCTAATTTATTAGG-3'	-55	JC19 with A-to-G change at position -48 in <i>umuD</i> probe
JC58	5'-ACCGGATACCTAATTTATTAGG-3'	-55	JC19 with C-to-G change at position -47 in <i>umuD</i> probe
JC59	5'-ACCGGATACCTAATTTATTAGG-3'	-55	JC19 with C-to-G change at position -46 in <i>umuD</i> probe
JC60	5'-ACCGGATACCTAATTTATTAGG-3'	-55	JC19 with T-to-G change at position -45 in <i>umuD</i> probe
JC61	5'-ACCGGATACCTAATTTATTAGG-3'	-55	JC19 with A-to-G change at position -44 in <i>umuD</i> probe
JC62	5'-ACCGGATACCTAATTTATTAGG-3'	-55	JC19 with A-to-G change at position -43 in <i>umuD</i> probe
JC63	5'-ACCGGATACCTAATTTATTAGG-3'	-55	JC19 with T-to-G change at position -42 in <i>umuD</i> probe
JC64	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with A-to-G change at position -41 in <i>umuD</i> probe
JC65	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with T-to-G change at position -40 in <i>umuD</i> probe
JC66	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with T-to-G change at position -39 in <i>umuD</i> probe
JC67	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with T-to-G change at position -38 in <i>umuD</i> probe
JC68	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with A-to-G change at position -37 in <i>umuD</i> probe
JC69	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with T-to-G change at position -36 in <i>umuD</i> probe
JC70	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with T-to-G change at position -35 in <i>umuD</i> probe
JC71	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with A-to-G change at position -34 in <i>umuD</i> probe
JC72	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with G-to-C change at position -33 in <i>umuD</i> probe
JC73	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with G-to-C change at position -32 in <i>umuD</i> probe
JC74	5'-ACCGGATACCTAATTTATTAGGTATCGCTGC-3'	-55	JC19 with T-to-G change at position -31 in <i>umuD</i> probe
637rv	5'-CCATCCATGTACTACAGGTC-3'	+1542	Lower primer to amplify <i>lexA</i> -ORF637 fragment in RT-PCR experiments
637fwd	5'-CTTGGATTGGGTGAAGCTTTGCC-3'	+1542	Upper primer to amplify ORF637- <i>rumA</i> and ORF637-ORF640 fragments in RT-PCR experiments
639rv	5'-ATCGGCTCCAATGATCGATCC-3'	+3120	Lower primer to amplify ORF637- <i>rumA</i> fragment in RT-PCR experiments
640rv	5'-ATGAAGGTATTCAACCCGGCCC-3'	+3787	Lower primer to amplify ORF637-ORF640 fragment in RT-PCR experiments

<sup>a</sup> Added restriction sites are italicized; nucleotide changes are underlined.

<sup>b</sup> Position of the 5' end of the oligonucleotide with respect to the proposed translational starting point of each *Magnetococcus* MC-1 gene.



hydrolysis of the Ala-Gly bond through a biochemical reaction characteristic of serine proteases (24). As a consequence of the cleavage, LexA no longer represses the SOS genes, whose expression is enhanced, and the SOS proteins participate directly in repairing DNA lesions, thereby aiding cell survival. Recently, another regulator has been added to the SOS system. It has been demonstrated that the product of the *dinI* gene is involved in the arrest of RecA activation, preventing its binding to single-stranded DNA (35). Once DNA damage is repaired, RecA\* loses its active conformation, LexA hydrolysis stops, and the repressor accumulates again, thus inhibiting the transcriptional expression of the SOS network.

The three-dimensional structure of the *E. coli* LexA protein has recently been determined (24). Two different *E. coli* LexA conformations exist: a cleavable form and a noncleavable conformation in which the catalytic center is distant from the cleavage site. In this model, RecA\* is suggested to be involved in the stabilization of the cleavable conformation (24).

The presence of the *lexA* gene represents a distinct difference between species of the domains *Archaea* and *Bacteria*. The occurrence of the *lexA* gene is widespread among members of the domain *Bacteria* and is exclusive to it. However, no *lexA* homologues have been found by BLAST queries in the completely sequenced genomes of several bacterial species including *Aquifex aerolicus*, *Borrelia burgdorferi*, *Chlamydia pneumoniae*, *Mycoplasma pneumoniae*, *Campylobacter jejuni*, *Helicobacter pylori*, *Porphyromonas gingivalis*, and others. *lexA*-like genes appear to be completely absent in members of the domain *Archaea*: no *lexA* homologues have been identified in the 21 archaeal genomes currently available for examination (9).

In *E. coli*, LexA binds specifically to a DNA motif known as the SOS box or LexA box (36). Both comparative analyses of the 31 *E. coli* LexA binding sites and site-directed mutagenesis experiments confirm that a 16-bp consensus sequence, CTGN<sub>10</sub>CAG, is the target for LexA in this organism (11, 37). This sequence also represents the LexA box in other species of the *Enterobacteriaceae* (16), as well as in other members of the gamma subclass of the *Proteobacteria* including *Vibrionaceae*, *Pasteurellaceae*, and *Pseudomonaceae* (16, 30). Nuclear magnetic resonance (NMR) studies of the N-terminal DNA binding domain of *E. coli* LexA suggest that the amino acids Asn<sup>41</sup>, Glu<sup>44</sup>, and Glu<sup>45</sup> are directly involved in the recognition of the SOS box (13). In addition to the *E. coli* SOS box, three more LexA boxes have been identified in the *Bacteria* domain. The direct repeat GTTCN<sub>7</sub>GTTC is recognized by LexA proteins of members of the alpha subclass of the *Proteobacteria* including *Rhodobacter sphaeroides*, *Paracoccus denitrificans*, and *Rhizobium etli* (8, 12, 34). The consensus sequence CGAAC RNRYGTTYG is the target for LexA in gram-positive bacteria such as *Bacillus subtilis* and *Mycobacterium tuberculosis*, among others (4, 38). Strikingly, the same motif has been described recently for *Dehalococcoides ethenogenes*, the first non-gram-positive bacterial species whose LexA repressor specifically recognizes the DinR box (10). The third LexA box was found in a member of the gamma subclass of the *Proteobacteria*; the sequence TTAGN<sub>6</sub>TACTA was identified as the target for the LexA repressor in *Xylella fastidiosa* (5).

Strain MC-1 is an incompletely characterized, unnamed, obligately microaerophilic gram-negative bacterium that represents the first and only magnetotactic coccus isolated in pure

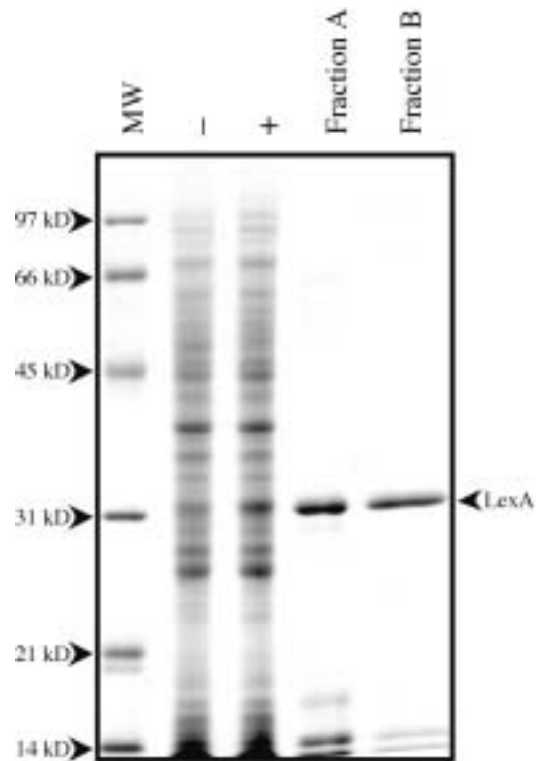


FIG. 1. Purification of the MC-1 LexA protein. MC-1 His-LexA was purified >90% by Co<sup>2+</sup> affinity chromatography using a TALON purification kit (see Materials and Methods for the detailed protocol). Each lane shown in the sodium dodecyl sulfate–13% polyacrylamide gel represents one of the different protein purification steps employed: MW, molecular weight marker; –, crude extract of BL21 Codon plus/pUA1013; +, crude extract of BL21 Codon plus/pUA1013 induced with 1 mM IPTG; Fraction A, MC-1 His-LexA eluted with 150 mM imidazole; Fraction B, purified MC-1 His-LexA protein after elution with 200 mM imidazole.

culture (26). Phylogenetically, this microorganism is affiliated with the alpha subclass of the *Proteobacteria*, in which it and the other (all uncultured) magnetotactic cocci appear to form a distinct group (7, 32). Although the genome of strain MC-1 is fully sequenced and available for examination ([http://www.jgi.doe.gov/JGI\\_microbial/html/magnetococcus/magneto\\_homepage.html](http://www.jgi.doe.gov/JGI_microbial/html/magnetococcus/magneto_homepage.html)), it has not been completely annotated and mapped. Sequencing of the DNA repair genes under LexA regulation in the alpha subclass of the *Proteobacteria* (such as *recA*, *uvrA*, and *ssb*) (12, 34) did not reveal the presence either of the direct repeat GTTCN<sub>7</sub>GTTC (the LexA box of this phylogenetic group, as mentioned above) or of the other LexA box sequences described previously. In order to gain insight into the composition of the MC-1 LexA regulon and to characterize its LexA binding site, we identified, cloned, and purified the product of its *lexA* gene.

#### MATERIALS AND METHODS

**Bacterial strains, plasmids, primers, and general techniques.** The bacterial strains and plasmids used in this study are listed in Table 1 together with their relevant features. Strain MC-1 was grown autotrophically with 10 mM sodium thiosulfate as the electron donor as previously described (14). *E. coli* strains were grown in Luria-Bertani medium (Difco) at 37°C. Ampicillin at 100 µg/ml and 1

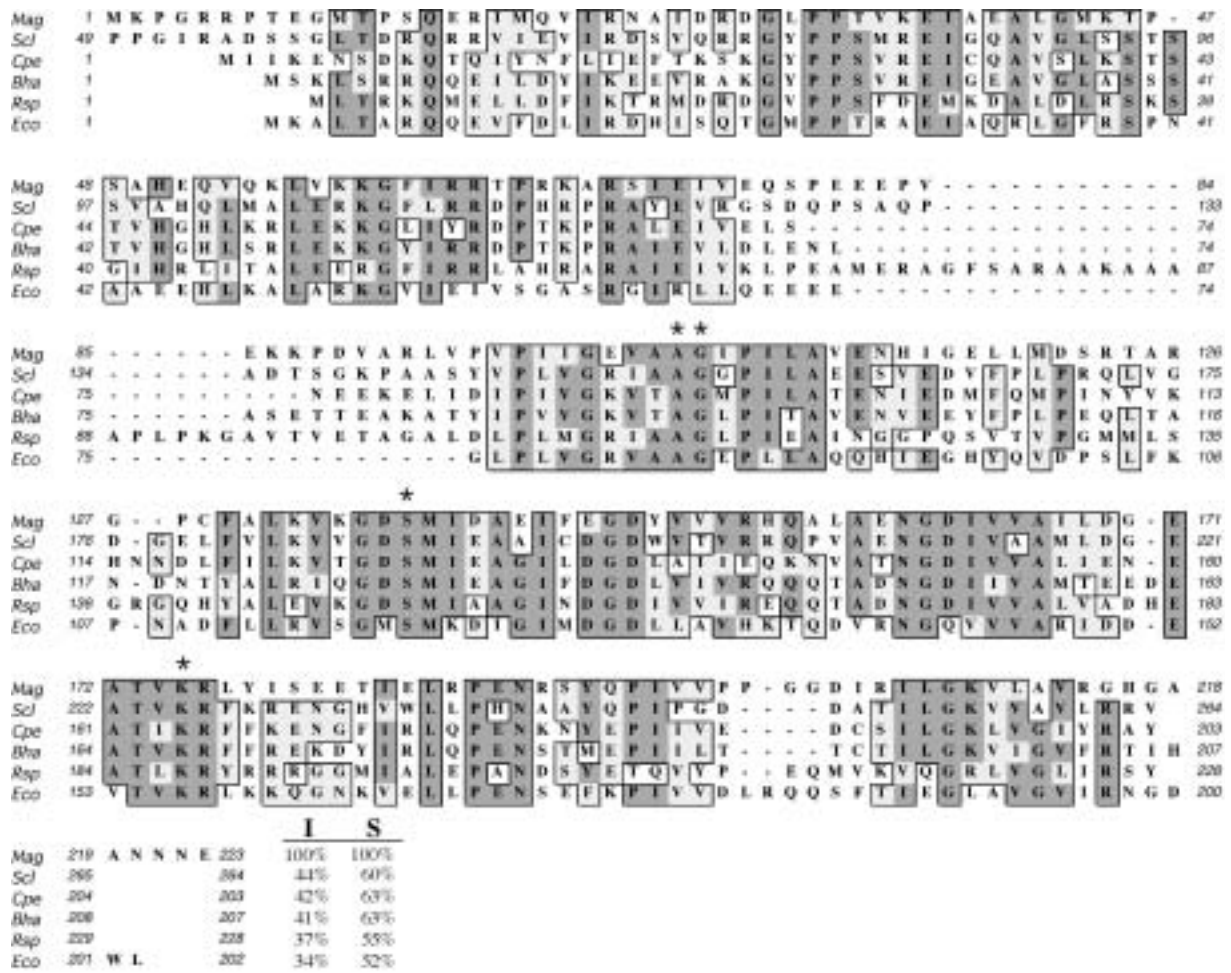


FIG. 2. CLUSTAL W alignment performed with Mac Vector (version 6.5; Oxford Molecular) comparing LexA proteins from strain MC-1 (Mag), *S. clavuligerus* (Scl) (residues 49 to 264), *C. perfringens* (Cpe), *B. halodurans* (Bha), *R. sphaeroides* (Rsp), and *E. coli* (Eco). Dark shading, identical conserved amino acids; light shading, similar conserved residues. I and S, percentages of identity and similarity, respectively, that each LexA sequence shares with the MC-1 LexA repressor. For better visualization of the figure, the first 49 residues of the *S. clavuligerus* LexA protein are not included in the alignment, because none of them matched in the comparison. Stars indicate the Ala, Gly, Ser, and Lys residues involved in the autocatalytic cleavage of the LexA protein. Accession numbers in the Entrez protein database at NCBI are as follows: *S. clavuligerus*, CAA12169; *C. perfringens*, BAB80867; *B. halodurans*, Q9KAD3; *R. sphaeroides*, Q9ZFA4; *E. coli*, P03033.

mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Roche, Mannheim, Germany) were added to growth media when required.

Cells of MC-1 used for genomic DNA extraction were harvested by centrifugation at 10,000 × g for 20 min at 4°C, resuspended in 20 mM Tris-HCl in a dilute artificial seawater (pH 7.0) (1), and re-centrifuged. Cells were then suspended in TE buffer (10 mM EDTA–10 mM Tris-HCl [pH 8.0]) for DNA extraction. Genomic DNA from strain MC-1 was obtained by using a modification of the method of Marmur (25) as described elsewhere (18).

Synthetic oligonucleotides (Tib Molbiol or Genosys) used in this study for PCR amplification of different DNA fragments are listed in Table 2. When required, PCR fragments were cloned into the pGEM-T vector (Promega) and both strands of the insert were sequenced by the dideoxynucleotide method, by labeling DNA samples with the fmol DNA Cycle Sequencing system (Promega) and using an ALF Sequencer (Amersham Pharmacia Biotech).

**Cloning of the MC-1 *lexA* gene.** The MC-1 *lexA* gene sequence was identified by performing a TBLASTN search of its unfinished genome at the National Center for Biotechnology Information (NCBI) ([http://www.ncbi.nlm.nih.gov/Microb\\_blast/unfinishedgenome.html](http://www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html)) with *E. coli* LexA protein as the query. The comparison yielded a region containing significant homology in part of contig 371, suggesting the presence of a 672-bp *lexA* gene in strain MC-1 (223 amino acid residues with an estimated molecular size of 24,453 kDa). Using BLAST analyses we also obtained the sequence of about 1 kb of DNA flanking each side of the *lexA* coding sequence. With this information, primers JCS and

JC6 (Table 2) were designed and used to PCR amplify the entire MC-1 *lexA* gene. The temperature profile employed was 35 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. The resulting DNA fragment was then cloned into pGEM-T (Promega). This plasmid, pUA1011, was used as a template for further studies.

**RT-PCR analysis of the *lexA* gene region.** To determine the transcriptional organization of the MC-1 *lexA* region, reverse transcriptase (Roche) was used to generate cDNA by reverse transcription-PCR (RT-PCR) using total RNA from MC-1 strain as a template and the pairs of primers indicated in Table 2. These oligonucleotides were designed to amplify PCR products of 1,542 and 1,578 bp if *lexA* and the three open reading frames found immediately downstream constituted a single transcription unit.

Total RNA from strain MC-1 was obtained as reported elsewhere (31). The RNA extracted was treated with RNase-free DNase I (Roche) to ensure the absence of contaminating DNA. The concentration and integrity of the RNA were determined by *A*<sub>260</sub> measurements and 1% formaldehyde-agarose gel electrophoresis, respectively. In all RT-PCR experiments, the absence of contaminating DNA in RNA samples after treatment with RNase-free DNase I was confirmed by carrying out PCR amplification without reverse transcriptase.

**Mobility shift assays.** Electrophoretic mobility shift assays (EMSA) were performed as previously described (12). Basically, probes were prepared by PCR amplification from MC-1 genomic DNA with one of the primers labeled at its 5' end with digoxigenin (DIG) (Table 2), and the products were purified in a 2 to

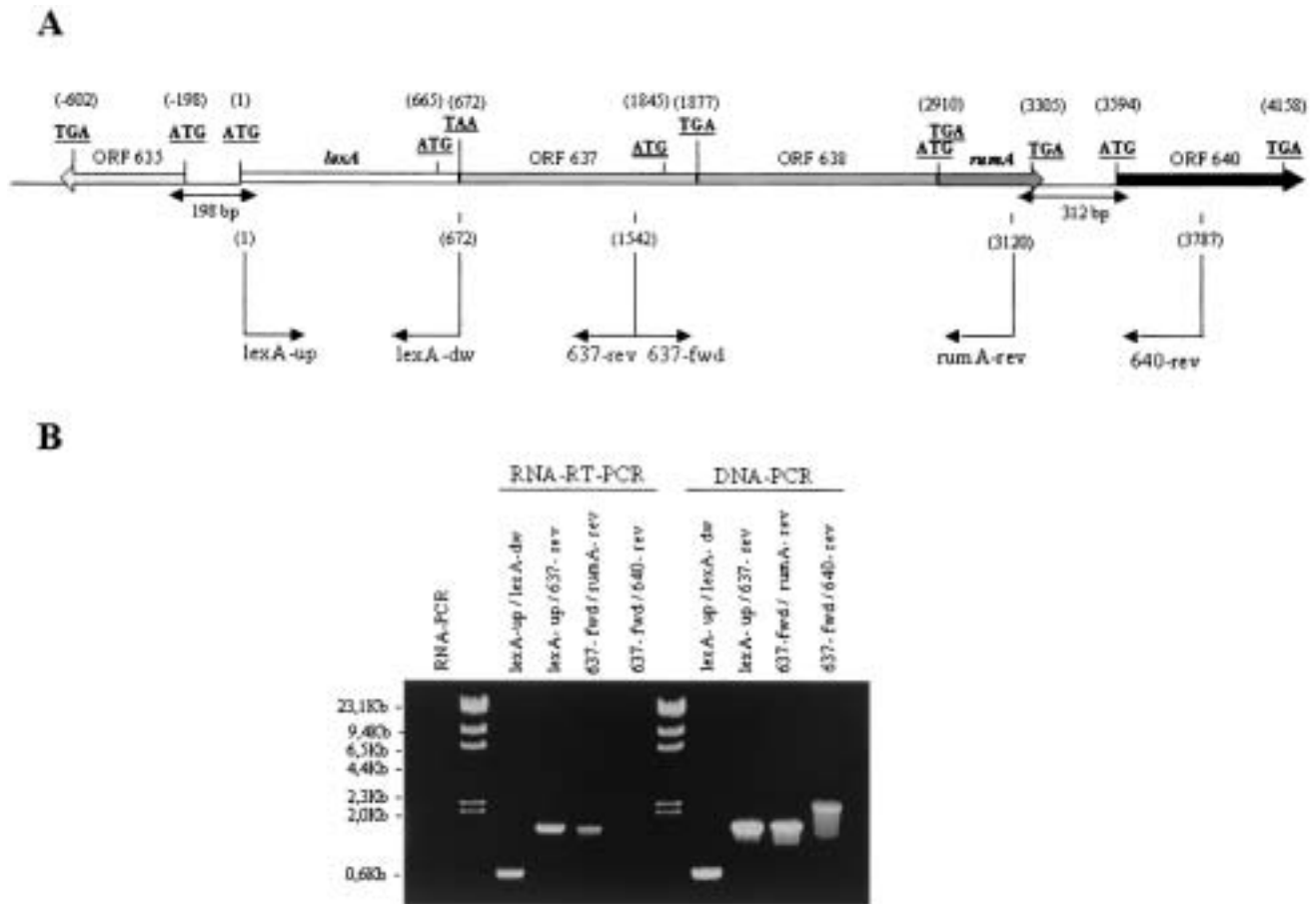


FIG. 3. (A) Genetic organization of the MC-1 *lexA* region. The proposed translational starting point and stop codon of each of these genes are boldfaced and underlined. Arrows indicate positions of primers used to identify the transcripts. Numerical positions refer to the putative *lexA* translational start codon. (B) RT-PCR transcriptional analysis of the region surrounding the *lexA* gene using total RNA from MC-1 cells (RNA-RT-PCR). As a control, PCR experiments were carried out with the same primers but without reverse transcriptase and with either RNA (RNA-PCR) or DNA (DNA-PCR) as a template. The band sizes of the molecular mass marker used (*Hind*III-digested  $\lambda$  DNA) are shown at the left of the gel.

3% low-melting-point agarose gel depending on DNA size. The thermal cycler profile was 35 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 30 s. Reaction mixtures (20  $\mu$ l) containing 10 ng of a DIG-labeled DNA probe and 15 ng (final concentration, approximately 25 to 30 nM) of pure MC-1 LexA were incubated in a binding buffer containing 10 mM HEPES NaOH (pH 8), 10 mM Tris-HCl (pH 8), 5% glycerol, 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 1  $\mu$ g of bulk carrier DNA, and 50  $\mu$ g of bovine serum albumin/ml. After 30 min of incubation at 30°C, the mixture was loaded onto a 5 to 6% nondenaturing Tris-glycine polyacrylamide gel (prerun for 30 min at 10 V/cm in 25 mM Tris-HCl [pH 8]–250 mM glycine–1 mM EDTA). DNA-protein complexes were separated at 150 V for 2.5 h in a 20-cm-long gel, followed by transfer to a Biodine B nylon membrane (Pall Gelman Laboratory). DIG-labeled DNA-protein complexes were detected by following the manufacturer's protocol (Roche). Experiments were repeated a minimum of three times to ensure the reproducibility of the results.

**Purification of MC-1 LexA protein.** MC-1 LexA protein was purified with the TALON purification kit (Stratagene), by  $\text{Co}^{2+}$  affinity chromatography taking advantage of the histidine tag placed in its N terminus. By use of primers JC5 and JC6, the coding sequence of the MC-1 *lexA* gene was amplified and cloned into the pGEM-T vector, generating plasmid pUA1011. To facilitate the cloning of *lexA* in the expression vector, primers JC5 and JC6 contain incorporated *Nde*I and *Bam*HI restriction sites, respectively (Table 2). Subsequently, pUA1011 was digested with *Nde*I-*Bam*HI, and the 0.7-kb DNA fragment was cloned into pET15b (Novagene) to express LexA with a hexahistidine tag at the N-terminal end (pUA1013). The latter plasmid was transformed into BL21 Codon plus RIL cells (Stratagene) to overproduce the protein. An overnight culture of BL21/

pUA1013 was diluted 1/100 in 1 liter of Luria-Bertani medium and incubated at 37°C until an optical density at 600 nm of 0.5 was reached. At that time, IPTG was added to the culture to a final concentration of 1 mM, and the culture was incubated for three additional hours. Cells were recovered by centrifugation at  $8,000 \times g$  for 15 min and resuspended in 20 ml of extraction-wash buffer (pH 7.0) according to the recommendations of the manufacturer (Stratagene). The cell suspension was then sonicated for 7 min at 40 W by using a Braun LabsonicU (Braun Biotech International) and centrifuged at  $18,000 \times g$  for 30 min. The supernatant containing the soluble His-LexA protein was incubated for 2 h at 4°C in TALON metal affinity resin previously equilibrated in extraction-wash buffer containing 0.1% Triton X-100. The resin was washed three times with extraction-wash buffer–0.1% Triton X-100 and twice more with extraction-wash buffer without Triton X-100. Next, resin containing attached His-LexA protein was placed in a 2-ml column, and after a wash with 2 bed volumes of extraction-wash buffer containing 10 mM imidazole, the protein was eluted with 2 ml of elution buffer containing 150 mM imidazole (fraction A) and then with 2 more ml of elution buffer containing 200 mM imidazole (fraction B) (Fig. 1). MC-1 LexA was purified to approximately 80% in fraction A, whereas LexA in fraction B was judged to be more than 90% pure based on a sodium dodecyl sulfate–13% polyacrylamide gel (Fig. 1).

## RESULTS AND DISCUSSION

**Cloning of the MC-1 *lexA* gene.** Although the genome sequence of strain MC-1 is not completely annotated, BLAST



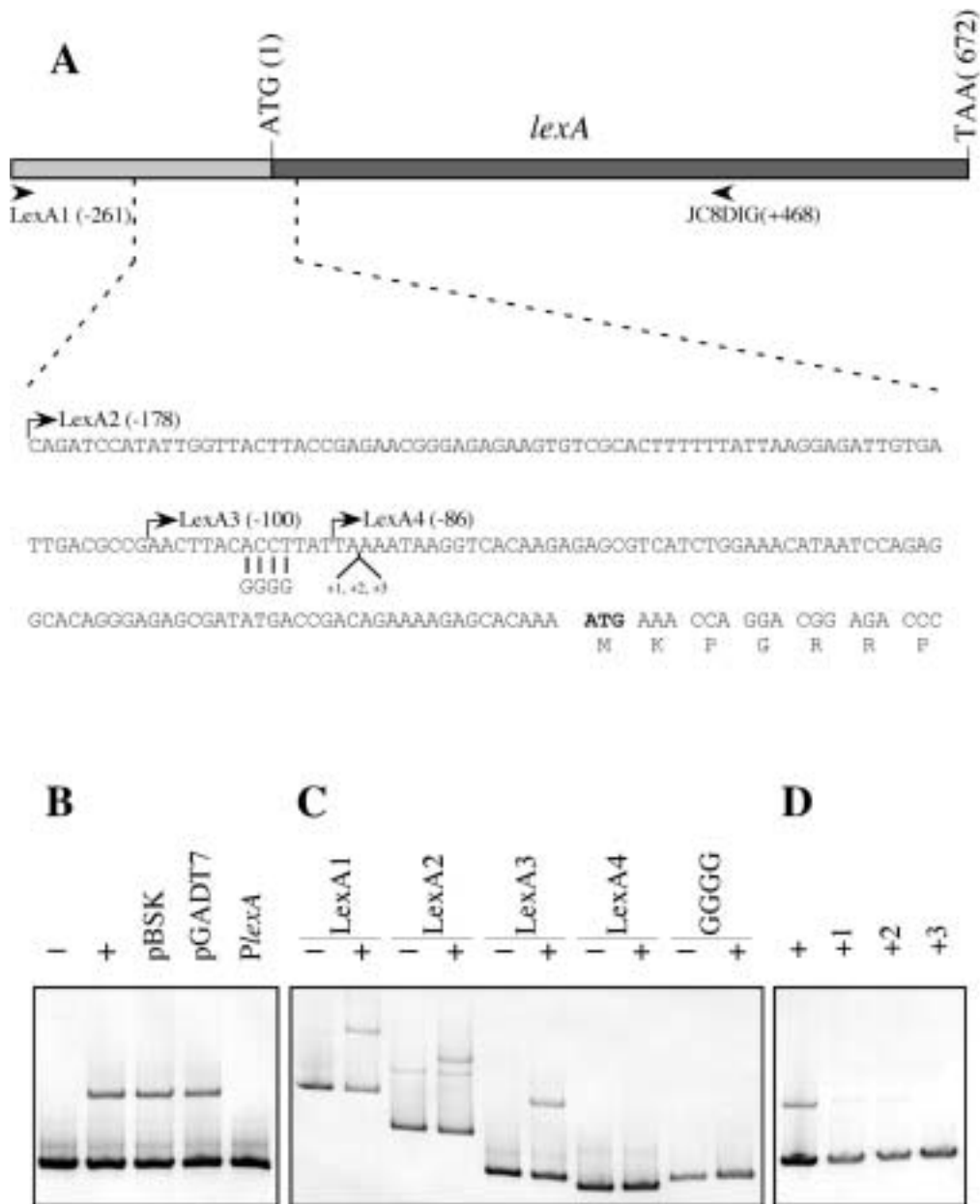


FIG. 4. (A) Sequence of the MC-1 *lexA* gene and promoter. The sequence of the *lexA* promoter, the first 7 amino acid residues of LexA, and the putative initiation and stop codons are shown. The start points of each fragment of the *lexA* promoter (LexA1, LexA2, LexA3, and LexA4) used in EMSA experiments are indicated by arrows, and the relative distances to the ATG are given in parentheses. (B) Study of the specific binding of the MC-1 LexA protein. EMSA experiments were performed using different DNAs as competitors. The LexA1 fragment was incubated in the absence (-) and in the presence (+) of pure LexA protein (see Materials and Methods for details). To demonstrate the specificity of LexA binding, the LexA1 fragment was incubated with pure MC-1 LexA protein and several unlabeled DNA competitors in the same reaction mixture containing 3  $\mu$ g of pBSK, 3  $\mu$ g of pGADT7, or a 100-fold molar excess (1  $\mu$ g) of the unlabeled LexA1 DNA fragment (*PlexA*). (C) Setting the bounds of the MC-1 LexA binding site. LexA1, LexA2, LexA3, LexA4, and a derivative of the LexA3 fragment in which the ACCT tetranucleotide was changed to GGGG (lanes marked "GGGG" above the gel) were incubated in the absence (-) or presence (+) of purified LexA from MC-1. (D) EMSA of LexA3-derived fragments where one (+1), two (+2), or three (+3) adenine residues were inserted at position -84 with respect to the ATG. Different probes were incubated in the presence of the LexA protein of MC-1 and then loaded onto a native Tris-glycine polyacrylamide (5%) gel, as described in Materials and Methods. The wild-type LexA3 fragment incubated in the presence of LexA (+) acted as the positive control.

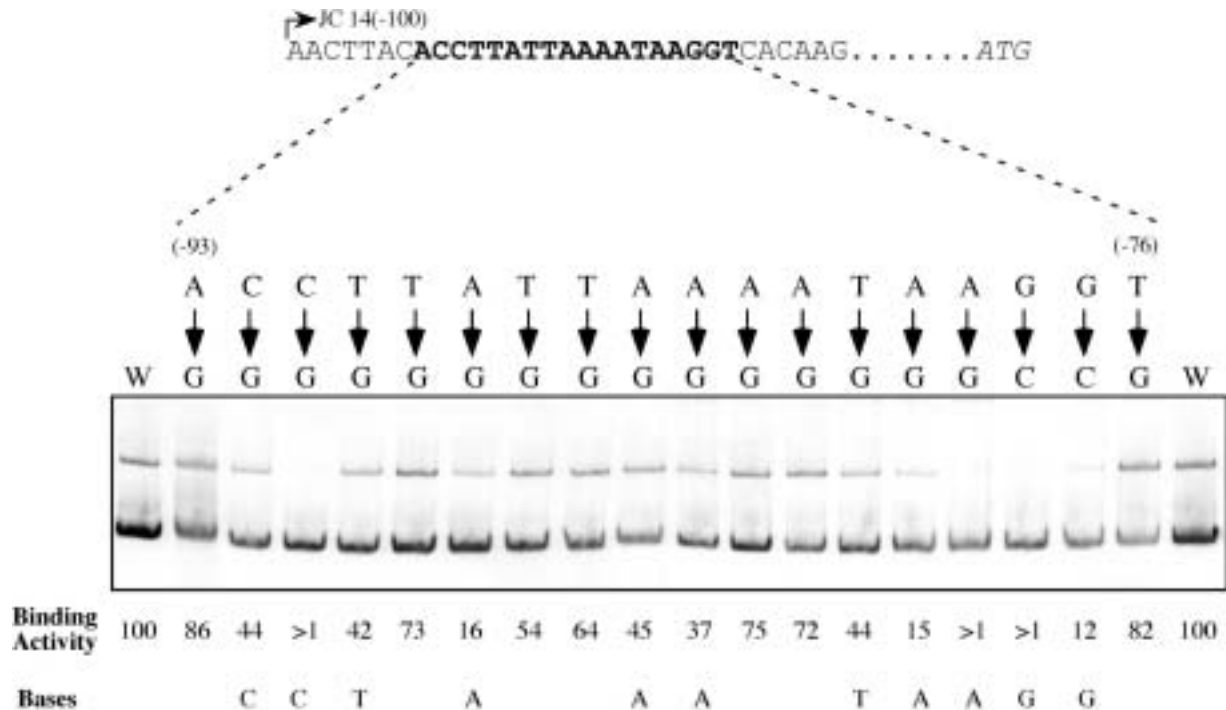


FIG. 5. Determination of the specific LexA binding target in the MC-1 *lexA* promoter. PCR-directed mutagenesis was employed to change single nucleotides in the region from  $-93$  to  $-76$  (boldfaced) relative to the start codon of the MC-1 *lexA* gene (italicized). EMSA experiments allowed determination of the affinity of the purified LexA protein for the resulting DIG-labeled probes. Arrows point to the nucleotide used to replace the native nucleotide, and the relative position of the nucleotide with respect to the translation initiation codon is given in parentheses. The LexA3 fragment with no change introduced as a positive control (W). The percentage of LexA binding activity remaining after each change was assessed using ImageQuant 1.2 software (Binding activity). Key nucleotides mutations of which caused a decrease of  $>50\%$  in LexA binding activity (relative to that of the wild-type control) are shown at the bottom (Bases).

analyses can be performed with contigs of the genome available at the U.S. Department of Energy (DOE) Joint Genome Institute (JGI) website ([http://www.jgi.doe.gov/JGI\\_microbial/html/magnetococcus/magneto\\_homepage.html](http://www.jgi.doe.gov/JGI_microbial/html/magnetococcus/magneto_homepage.html)) and at NCBI ([http://www.ncbi.nlm.nih.gov/Microb\\_blast/unfinished\\_genome.html](http://www.ncbi.nlm.nih.gov/Microb_blast/unfinished_genome.html)). The *E. coli* LexA 202-amino-acid sequence was used to query the MC-1 genome with the TBLASTN program, revealing the presence of a 223-amino-acid protein showing significant homology. The 223-residue protein conserves the two main domains of LexA repressors: an N-terminal DNA binding domain harboring a helix-turn-helix (HTH) motif and a C-terminal domain containing serine protease activity. When compared to current databases, MC-1 LexA shows high homology to the LexA proteins of *Streptomyces clavuligerus*, *Clostridium perfringens*, and *Bacillus halodurans*. MC-1 LexA protein shares the highest identity with *Streptomyces clavuligerus* LexA (44%), while the repressor is only 37 and 34% identical to the *Rhodobacter sphaeroides* and *E. coli* LexA proteins, respectively (Fig. 2). As deduced from CLUSTAL W alignment of different LexA proteins (Mac Vector, version 6.5; Oxford Molecular), MC-1 LexA contains the four conserved key residues involved in repressor autocleavage (Ala<sup>104</sup>, Gly<sup>105</sup>, Ser<sup>138</sup>, and Lys<sup>175</sup>) (Fig. 2). However, dramatic structural differences can be observed when the MC-1 LexA three-dimensional structure is modeled and compared to that of *E. coli* (data not shown). For example, the second helix of the HTH motif is shorter than that of *E. coli*, and in the hinge region between the DNA binding and C-terminal domains

of MC-1 LexA, there are 14 additional residues forming an extra double beta-sheet structure not present in *E. coli* (24).

An open reading frame in the opposite transcriptional direction, encoding a putative protein of unknown function consisting of 406 amino acids, was found 198 bp upstream of the *lexA* gene (Fig. 3A). Downstream of *lexA*, and in the same transcriptional direction, there are three open reading frames (Fig. 3A). The first two encode putative proteins of unknown function, whereas the third shows significant identity (58%) with the product of the *rumA* gene of the *E. coli* plasmid R391, which is a homologue of the *umuD* gene (20), encoding a subunit of the DNA damage-inducible polymerase V (33). The fact that the distance between *lexA* and these three open reading frames is very short (Fig. 3A) suggested that all of these genes are cotranscribed. To evaluate this hypothesis, RT-PCR analysis of total MC-1 RNA was carried out with primers designed to amplify a fragment of 1,542 bp or of 1,578 bp if a polycistronic mRNA was produced. Recovered PCR products demonstrated that transcription of these four genes is linked (Fig. 3B). It is worth noting that in the R391 plasmid a second gene called *rumB*, downstream of *rumA*, is cotranscribed with *rumA* (20), whereas in strain MC-1, neither gene downstream of the *rumA*-like gene is cotranscribed with it.

**Identification of the MC-1 LexA-binding site.** All of the LexA proteins characterized to date share a common feature: they negatively control their own transcriptional expression after binding specifically to a DNA motif located at their pro-

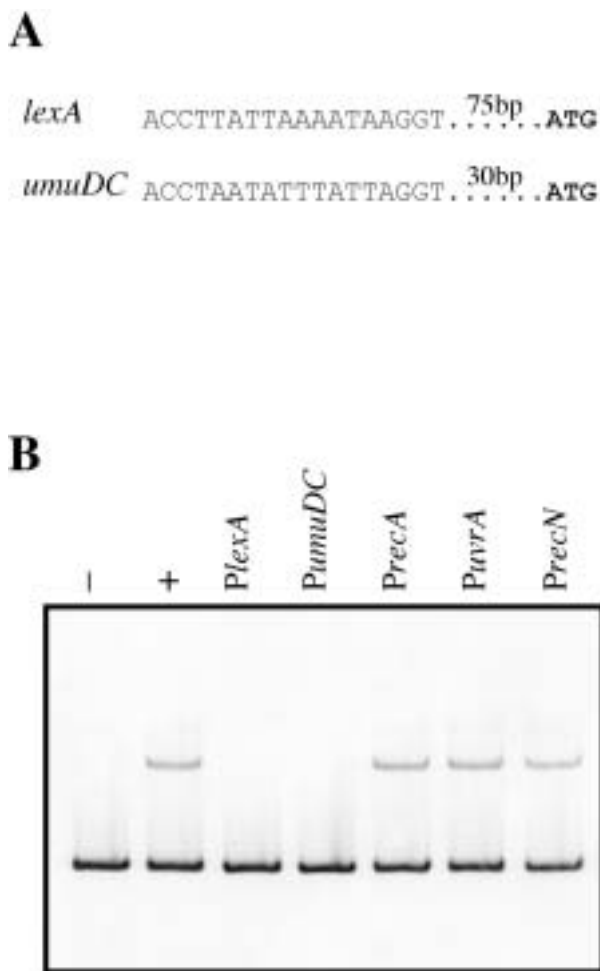


FIG. 6. (A) Upstream operator sequences of MC-1 *lexA* and *umuDC* genes presenting potential LexA binding sites. LexA binding sites are depicted, the putative translation initiation codons are bold-faced, and the distances between them are given. (B). EMSA experiments using the LexA3 fragment incubated with purified MC-1 LexA protein. At the same time, a DIG-labeled LexA3 fragment was incubated with MC-1 LexA protein and one of five unlabeled DNA competitors (100-fold molar excess) containing the promoter regions and potential LexA binding sites from the *lexA* (*PlexA*) and *umuDC* (*PumuDC*) genes. Similarly, promoter regions of previously characterized LexA-regulated genes in the alpha subclass of the *Proteobacteria* were used as competitors: *recA* (*PrecA*), *uvrA* (*PuvrA*), and *recN* (*PrecN*).

moter regions (3, 5, 10, 12, 38). To identify this sequence in the MC-1 strain, a 261-bp region upstream of the putative translational start codon of the *lexA* gene (LexA1) was PCR amplified using oligonucleotide primers JC7 and JC8DIG (Fig. 4A). This LexA1 fragment was then utilized as a probe in EMSA experiments with purified LexA protein of this microorganism. MC-1 LexA shifted the mobility of the LexA1 probe (Fig. 4B). Furthermore, the retarded band did not disappear when 3  $\mu$ g of nonspecific DNA was included in the same reactions (Fig. 4B). However, the shifted band was eliminated by use of a 100-fold molar excess (1  $\mu$ g) of the unlabeled LexA1 DNA fragment, indicating that binding of MC-1 LexA to its promoter region is indeed specific (Fig. 4B). To establish the

position of the MC-1 LexA binding site in its promoter, serial deletions of DIG-labeled fragments from this region were analyzed in EMSA experiments. Whereas probes LexA1, LexA2, and LexA3 showed a shifted band in the presence of pure LexA protein, LexA4 did not shift its electrophoretic mobility (Fig. 4C), suggesting that the LexA binding target, or at least part of it, is located between positions  $-100$  and  $-86$  relative to the putative translational start codon of the *lexA* gene (Fig. 4A). A visual study of this region revealed the presence of an inverted repeat, ACCTTATTAAAATAAGGT (ACCTTN<sub>8</sub>AGGT), 75 bp upstream of the ATG start codon. The ACCT sequence was changed to GGGG in the LexA3 probe by using the JC16 primer (Fig. 4A). The mobility was not changed in the presence of LexA (Fig. 4C), indicating that these four nucleotides of the inverted repeat are part of the specific recognition site for the MC-1 LexA protein. In other studies, the introduction of additional nucleotides in the linker region of the inverted repeat resulted in the abolishment of the binding ability of LexA for its target sequence (12). To determine whether this was also true for our motif, one to three adenines were introduced at position  $-84$  relative to the ATG start codon by using primers JC22 to JC24, thereby lengthening the N<sub>10</sub> region between the ACCT and AGGT motifs (Fig. 4A). When either one or two adenines were introduced into the LexA3 probe, some residual binding of LexA was observed. However, when three adenines were introduced into this linker region, the LexA binding activity disappeared completely, providing further corroboration that the ACCTN<sub>10</sub>AGGT motif is involved in the MC-1 LexA repressor binding.

To study the individual importance of each of the 18 bases of this inverted repeat in LexA binding activity, single-nucleotide changes were introduced by PCR using primers JC25 to JC43 in the LexA3 probe (Fig. 5). EMSA experiments with each single change revealed that bases belonging to the main motifs of the inverted repeat, CCT and AGG, are important in the recognition of LexA protein. More precisely, the most dramatic changes affecting LexA binding activity are located at positions  $-91$  (C),  $-79$  (A), and  $-78$  (G) relative to the ATG start codon, respectively (Fig. 5). When these nucleotides were changed, no DNA-LexA complex was detected, indicating their key importance in LexA binding. Quantification of the retarded band using ImageQuant (version 1.2) Macintosh software revealed that other nucleotides also play a significant role in the recognition of MC-1 LexA repressor. This is the case for A at positions  $-88$ ,  $-85$ ,  $-84$ , and  $-80$ , C at position  $-92$ , T at positions  $-90$  and  $-81$ , and G at position  $-77$  (relative to the putative translational start codon), since a single-nucleotide change resulted in a  $>50\%$  loss of LexA binding activity (Fig. 5). These results strongly indicate that CCTTATTAAAATAAGG is the specific binding site recognized by the MC-1 LexA protein in the *lexA* promoter.

***umuD* and *umuC* are part of the LexA regulon in strain MC-1.** In *E. coli*, a minimum of 40 genes are directly regulated by the LexA repressor (6, 11, 17). This is, thus far, the bacterial species for which the largest number of LexA-regulated genes have been described. Hence, these genes were used as a query in a TBLASN search using the unfinished MC-1 genome sequence. Homologues of most of the *E. coli* SOS genes were identified in the MC-1 chromosome. However, other SOS gene homologues such as *dinI*, *dinG*, *polB*, *sbmC*, *yebG*, *ydjM*, *yjiW*,





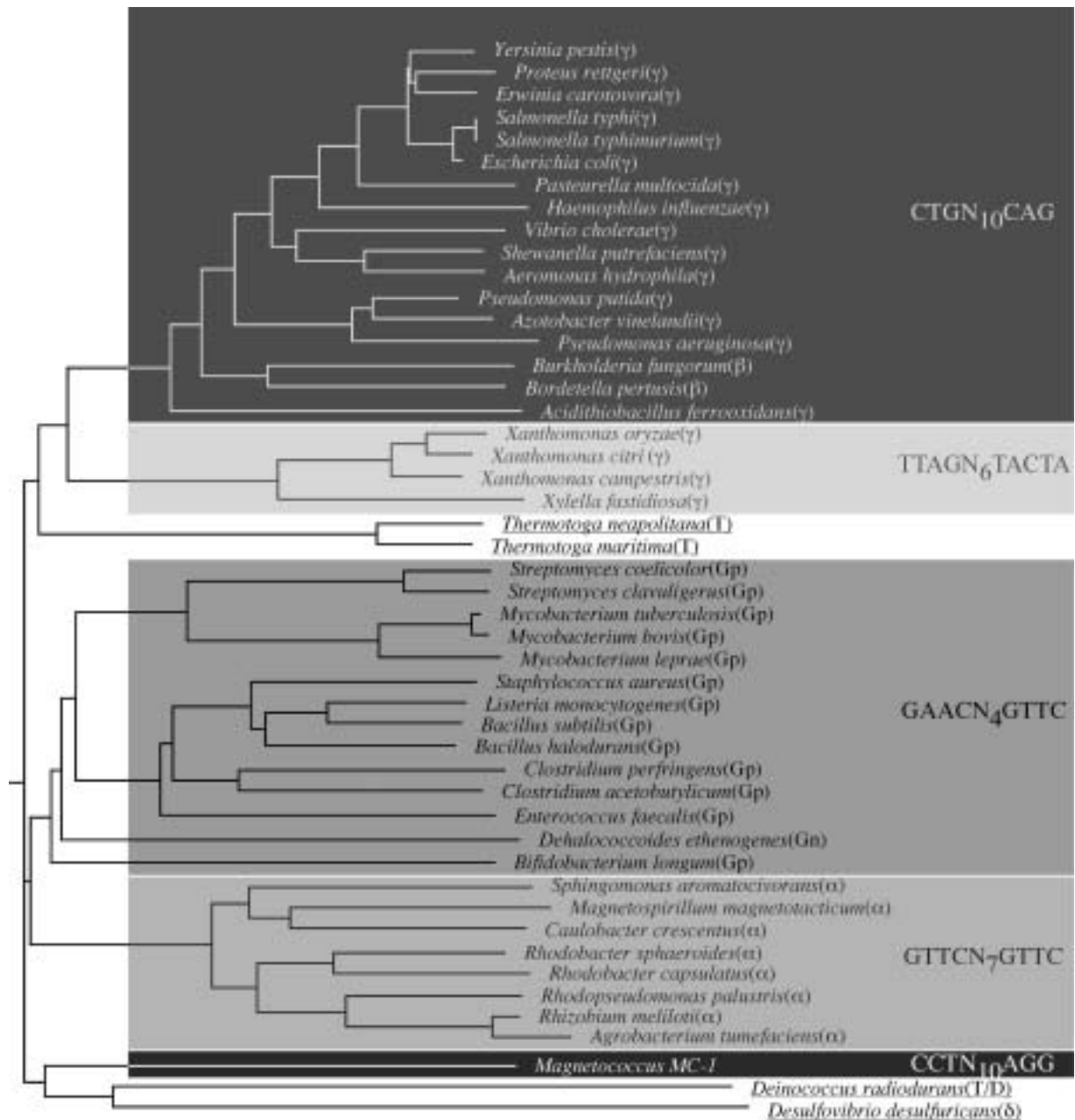


FIG. 8. Dendrogram constructed with CLUSTAL W software (Mac Vector, version 6.5) using 48 N-terminal LexA binding domains from different phylogenetic groups (from the first residue to the Ala-Gly cleavage site). The matrix employed was BLOSUM30. Clusters are identified by shades of gray grouping, in each case, the different bacterial species showing a determined LexA binding site found in their *lexA* promoter. Organisms that do not show any of the five LexA boxes characterized to date in their *lexA* promoters are underlined. Phylogenetic groups are shown in parentheses.  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ , the alpha, beta, gamma, and delta subclasses of the *Proteobacteria*, respectively; Gp, gram-positive; Gn, green nonsulfur bacteria; T, *Thermotogales*; T/D, *Thermus/Deinococcus* group.

to contain only a few. So far, by use of computational searches, only three LexA-regulated transcriptional units have been characterized in the entire genomes of these two microorganisms. This is also the case for *D. ethenogenes*, in whose entire genome only two LexA-regulated genes have been identified (10). In fact, most of the *E. coli* LexA-regulated genes found in the *X. fastidiosa*, *D. ethenogenes*, and MC-1 genomes are not under the control of their LexA repressor (5, 6, 10, 11, 17). A possible explanation for this wide variability in the number of

LexA-regulated genes in different species of bacteria might be related to the environment which these microorganisms inhabit. Bacterial species subjected to a permanent stress in their natural habitat should constitutively express many DNA repair genes and thus be not sensitive to the fine-tuning of LexA. It should be mentioned that the lack of a LexA binding site in the *recA* promoter of strain MC-1 indicates that this classical SOS gene is not directly regulated by LexA in this organism. This situation is not unique and has been proposed to occur in

*Deinococcus radiodurans* and *D. ethenogenes* also (10, 28). In the case of *D. radiodurans*, the *recA* gene is induced by DNA damage but in a LexA-independent manner (2, 28), although it has been reported that *D. radiodurans* contains a second copy of the *lexA* gene whose role in *recA* regulation has not been determined (28).

As discussed earlier, strain MC-1 is phylogenetically affiliated with the alpha subclass of the *Proteobacteria*, which contains several other magnetotactic species including *Magneto- spirillum magnetotacticum* (7, 32). In this phylogenetic group, *recA*, *uvrA*, and *ssb* have been described as LexA-regulated genes (12, 34). In this study, we demonstrated that the LexA repressor in strain MC-1 does not bind to the promoter regions of such genes. In the case of *ssb*, we could not completely confirm this, since *ssb* is present in a very short contig (contig 471) lacking its promoter region (the genome has not yet been annotated and completely mapped). We have shown that the MC-1 LexA binding site is different from the direct repeat typically found in the alpha subclass of the *Proteobacteria*. Thus, our findings do not correlate with the current phylogenetic affiliation of strain MC-1 based on the sequence of the 16S rRNA gene. A dendrogram performed from CLUSTAL W alignment of 48 N-terminal DNA binding domains of different bacterial LexA proteins belonging to diverse phylogenetic groups demonstrates this difference and shows that the MC-1 LexA protein clearly diverges very early from the main branch of the alpha subclass of the *Proteobacteria* (Fig. 8). Despite the fact that MC-1 LexA shares its highest level of identity with the LexA proteins of gram-positive bacteria, the MC-1 LexA protein appears to form an independent branch separate from the gram-positive bacteria that is relatively closer to gram-negative representatives of the alpha subclass of the *Proteobacteria* (Fig. 8). Another possibility is that the MC-1 *lexA* gene is of viral origin, because the amino acid residues involved in the autocleavage of both the LexA protein and the lytic repressors of bacteriophage as well as their hydrolysis mechanisms are the same (36). In fact, it has recently been hypothesized that the second *lexA* gene of *D. radiodurans* might be of viral origin, since this gene is at the left end of a defective temperate bacteriophage inserted in chromosome I of this organism (27). However, there are no experimental data that confirm this supposition. Moreover, it seems unlikely that this is the case for strain MC-1, since there appear to be no bacteriophage-related genes surrounding the *lexA* gene or the *umuDC* genes, which, as shown in this work, belong to its LexA network ([http://www.jgi.doe.gov/JGI\\_microbial/html/magnetococcus/magneto\\_homepage.html](http://www.jgi.doe.gov/JGI_microbial/html/magnetococcus/magneto_homepage.html)).

The CLUSTAL W alignment is a useful tool for assigning a new LexA protein to a determined group on the basis of the target sequence that this protein recognizes. As seen in Fig. 8, with CLUSTAL W, different bacterial species can be clustered by using the LexA binding sites. Species belonging to different phylogenetic groups, such as *Burkholderia fungorum* and *Bordetella pertussis* (beta subclass of the *Proteobacteria*), are grouped together with the members of the gamma subclass that recognize the same LexA box, CTGN<sub>10</sub>CAG (Fig. 8). The LexA protein of the green nonsulfur bacterium *D. ethenogenes* recognizes the DinR box, and as a consequence, this organism is clustered with the gram-positive bacteria (Fig. 8). Species of the genera *Xylella* and *Xanthomonas*, despite being affiliated

with the gamma subclass of the *Proteobacteria*, group independently of the gamma subclass (Fig. 8), since their LexA repressor recognizes a different target, TTAN<sub>6</sub>TACTA, as has been recently demonstrated (5).

This characterization of the MC-1 LexA binding site raises the total number of such motifs in the domain *Bacteria* to five. However, this number will likely increase, since other bacterial species including *Desulfovibrio desulfuricans* (delta subclass of the *Proteobacteria*) and *Thermotoga maritima* (*Thermotogales*) do not contain any of these five motifs but have *lexA*-like genes in their genomes (<http://igweb.integratedgenomics.com/GOLD/index.cgi?want=Prokaryotic+Ongoing+Genomes>), suggesting that they utilize other, undescribed sequences as LexA binding sites. Identification of new targets for LexA in these species will provide valuable information about the evolution of the LexA repressor and the gene network regulated by this damage-inducible protein.

#### ACKNOWLEDGMENTS

This work was funded by grants BMC2001-2065 from the Ministerio de Ciencia y Tecnología (MCyT) de España and 2001SGR-206 from the Departament d'Universitats, Recerca i Societat de la Informació (DURSI) de la Generalitat de Catalunya. A. R. Fernández de Henestrosa is the recipient of a postdoctoral reincorporation contract, and Gerard Mazón and Jordi Cuñé are recipients of a predoctoral fellowship from the MCyT. D. A. Bazylinski is supported by grant NAG 9-1115 from the U.S. National Aeronautics and Space Administration (NASA) Johnson Space Center.

We acknowledge Pilar Cortés and Joan Ruiz for excellent technical assistance.

A. R. Fernández de Henestrosa and Jordi Cuñé contributed equally to this work.

#### REFERENCES

- Bazylinski, D. A., A. J. Garratt-Reed, and R. B. Frankel. 1993. Electron microscopic studies of magnetosomes in magnetotactic bacteria. *Microsc. Res. Tech.* **27**:389–401.
- Bonacossa De Almeida, C., G. Coste, S. Sommer, and A. Bailone. 2002. Quantification of RecA protein in *Deinococcus radiodurans* reveals involvement of RecA, but not LexA, in its regulation. *Mol. Genet. Genomics* **268**:28–41.
- Brent, R., and M. Ptashne. 1981. Mechanism of action of the *lexA* gene product. *Proc. Natl. Acad. Sci. USA* **78**:4204–4208.
- Brooks, P. C., F. Movahedzadeh, and E. O. Davis. 2001. Identification of some DNA damage-inducible genes of *Mycobacterium tuberculosis*: apparent lack of correlation with LexA binding. *J. Bacteriol.* **183**:4459–4467.
- Campoy, S., G. Mazon, A. R. Fernández de Henestrosa, M. Llagostera, R. B. Monteiro, and J. Barbé. 2002. A new regulatory DNA motif of the gamma subclass *Proteobacteria*: identification of the LexA binding site of the plant pathogen *Xylella fastidiosa*. *Microbiology* **148**:3583–3597.
- Courcelle, J., A. Khodursky, B. Peter, P. O. Brown, and P. C. Hanawalt. 2001. Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics* **158**:41–64.
- DeLong, E. F., R. B. Frankel, and D. A. Bazylinski. 1993. Multiple evolutionary origins of magnetotaxis in bacteria. *Science* **259**:803–806.
- del Rey, A., J. Diestra, A. R. Fernández de Henestrosa, and J. Barbé. 1999. Determination of the *Paracoccus denitrificans* SOS box. *Microbiology* **145**:577–584.
- Eisen, J. A., and P. C. Hanawalt. 1999. A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat. Res.* **435**:171–213.
- Fernández de Henestrosa, A. R., J. Cune, I. Erill, J. K. Magnuson, and J. Barbé. 2002. A green nonsulfur bacterium, *Dehalococcoides ethenogenes*, with the LexA binding sequence found in gram-positive organisms. *J. Bacteriol.* **184**:6073–6080.
- Fernández de Henestrosa, A. R., T. Ogi, S. Aoyagi, D. Chafin, J. J. Hayes, H. Ohmori, and R. Woodgate. 2000. Identification of additional genes belonging to the LexA regulon in *Escherichia coli*. *Mol. Microbiol.* **35**:1560–1572.
- Fernández de Henestrosa, A. R., E. Rivera, A. Tapias, and J. Barbé. 1998. Identification of the *Rhodobacter sphaeroides* SOS box. *Mol. Microbiol.* **28**:991–1003.
- Fogh, R. H., G. Ottleben, H. Ruterjans, M. Schnarr, R. Boelens, and R. Kaptein. 1994. Solution structure of the LexA repressor DNA binding domain determined by <sup>1</sup>H NMR spectroscopy. *EMBO J.* **13**:3936–3944.

14. Frankel, R. B., D. A. Bazylinski, M. S. Johnson, and B. L. Taylor. 1997. Magneto-aerotaxis in marine coccoid bacteria. *Biophys. J.* **73**:994–1000.
15. Friedberg, E. C., G. C. Walker, and W. Siede. 1995. DNA repair and mutagenesis. American Society for Microbiology, Washington, D.C.
16. Garriga, X., S. Calero, and J. Barbé. 1992. Nucleotide sequence analysis and comparison of the *lexA* genes from *Salmonella typhimurium*, *Erwinia carotovora*, *Pseudomonas aeruginosa* and *Pseudomonas putida*. *Mol. Gen. Genet.* **236**:125–134.
17. Khil, P. P., and R. D. Camerini-Otero. 2002. Over 1000 genes are involved in the DNA damage response of *Escherichia coli*. *Mol. Microbiol.* **44**:89–105.
18. Kimble, L. K., L. Mandelco, C. R. Woese, and M. T. Madigan. 1995. *Hellobacterium modesticaldum*, sp. nov., a thermophilic heliobacterium of hot springs and volcanic soils. *Arch. Microbiol.* **163**:259–267.
19. Koch, W. H., and R. Woodgate. 1998. The SOS response, p. 107–134. In J. A. Nickoloff and M. F. Hoekstra (ed.), DNA damage and repair: DNA repair in prokaryotes and lower eukaryotes, 1st ed. Humana Press, Totowa, N.J.
20. Kulaeva, O. L., J. C. Wootton, A. S. Levine, and R. Woodgate. 1995. Characterization of the *umu*-complementing operon from R391. *J. Bacteriol.* **177**:2737–2743.
21. Little, J. W. 1993. LexA cleavage and other self-processing reactions. *J. Bacteriol.* **175**:4943–4950.
22. Little, J. W., S. H. Edmiston, L. Z. Pacelli, and D. W. Mount. 1980. Cleavage of the *Escherichia coli* LexA protein by the RecA protease. *Proc. Natl. Acad. Sci. USA* **77**:3225–3229.
23. Little, J. W., D. W. Mount, and C. R. Yanisch-Perron. 1981. Purified LexA protein is a repressor of the *recA* and *lexA* genes. *Proc. Natl. Acad. Sci. USA* **78**:4199–4203.
24. Luo, Y., R. A. Pfuetzner, S. Mosimann, M. Paetzel, E. A. Frey, M. Cherney, B. Kim, J. W. Little, and N. C. Strynadka. 2001. Crystal structure of LexA: a conformational switch for regulation of self-cleavage. *Cell* **106**:585–594.
25. Marmur, J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. *J. Mol. Biol.* **3**:208–218.
26. Meldrum, F. C., S. Mann, B. R. Heywood, R. B. Frankel, and D. A. Bazylinski. 1993. Electron microscopy study of magnetosomes in a cultured coccoid magnetotactic bacterium. *Proc. R. Soc. Lond. B* **251**:231–236.
27. Morgan, G. J., G. F. Hatfull, S. Casjens, and R. W. Hendrix. 2002. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. *J. Mol. Biol.* **317**:337–359.
28. Narumi, I., K. Satoh, M. Kikuchi, T. Funayama, T. Yanagisawa, Y. Kobayashi, H. Watanabe, and K. Yamamoto. 2001. The LexA protein from *Deinococcus radiodurans* is not involved in RecA induction following gamma irradiation. *J. Bacteriol.* **183**:6951–6956.
29. Radman, M. 1974. Phenomenology of an inducible mutagenic DNA repair pathway in *Escherichia coli*: SOS repair hypothesis, p. 128–142. In L. Prakash, F. Sherman, M. Miller, C. W. Lawrence, and H. W. Tabor (ed.), Molecular and environmental aspects of mutagenesis. Charles C. Thomas, Springfield, Ill.
30. Riera, J., and J. Barbe. 1995. Cloning, sequence and regulation of expression of the *lexA* gene of *Aeromonas hydrophila*. *Gene* **154**:71–75.
31. Sambrook, J., and D. W. Russell. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
32. Spring, S., and D. A. Bazylinski. 1 December 2000, posting date. Magneto-tactic bacteria. In M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (ed.), The prokaryotes: an evolving electronic resource for the microbiological community, release 3.4. [Online.] Springer-Verlag New York, Inc., New York, N.Y. <http://141.150.157.117:8080/prokPUB/index.htm>.
33. Tang, M., X. Shen, E. G. Frank, M. O'Donnell, R. Woodgate, and M. F. Goodman. 1999. UmuD'2C is an error-prone DNA polymerase, *Escherichia coli* pol V. *Proc. Natl. Acad. Sci. USA* **96**:8919–8924.
34. Tapias, A., and J. Barbé. 1999. Regulation of divergent transcription from the *uvrA-ssb* promoters in *Sinorhizobium meliloti*. *Mol. Gen. Genet.* **262**:121–130.
35. Voloshin, O. N., B. E. Ramirez, A. Bax, and R. D. Camerini-Otero. 2001. A model for the abrogation of the SOS response by an SOS protein: a negatively charged helix in DinI mimics DNA in its interaction with RecA. *Genes Dev.* **15**:415–427.
36. Walker, G. C. 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol. Rev.* **48**:60–93.
37. Wertman, K. F., and D. W. Mount. 1985. Nucleotide sequence binding specificity of the LexA repressor of *Escherichia coli* K-12. *J. Bacteriol.* **163**:376–384.
38. Winterling, K. W., A. S. Levine, R. E. Yasbin, and R. Woodgate. 1997. Characterization of DinR, the *Bacillus subtilis* SOS repressor. *J. Bacteriol.* **179**:1698–1703.

## **ARTICLE III**



## The *Leptospira interrogans* *lexA* Gene Is Not Autoregulated

Jordi Cuñé,<sup>1</sup> Paul Cullen,<sup>2</sup> Gerard Mazon,<sup>3</sup> Susana Campoy,<sup>3</sup> Ben Adler,<sup>2</sup>  
 and Jordi Barbe<sup>1,3\*</sup>

Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain<sup>1</sup>; Australian Bacterial Pathogenesis Program, Department of Microbiology, Monash University, Victoria 3800, Australia<sup>2</sup>; and Centre de Recerca en Sanitat Animal, Bellaterra, 08193 Barcelona, Spain<sup>3</sup>

Received 17 March 2005/Accepted 30 May 2005

**Footprinting and mutagenesis experiments demonstrated that *Leptospira interrogans* LexA binds the palindrome TTTGN<sub>5</sub>CAA found in the *recA* promoter but not in the *lexA* promoter. In silico analysis revealed that none of the other canonical SOS genes is under direct control of LexA, making the leptospiral *lexA* gene the first described which is not autoregulated.**

Genome integrity in bacterial cells depends on their DNA repair ability. In free-living bacterial species, many environmental factors, such as UV radiation, chemical compounds, and some antibiotics, can produce chromosomal lesions. DNA from pathogenic bacteria may be also affected by several host factors (e.g., oxygen and iron concentration). Likewise, meta-

bolic endogenous compounds can also cause damage to bacterial DNA. For these reasons, bacterial cells possess several pathways involved in DNA repair. Despite the fact that most of them target specific kinds of DNA lesions (e.g., oxidative damage or presence of alkyl radicals in DNA), a global DNA damage response, known as the SOS system, is present in many

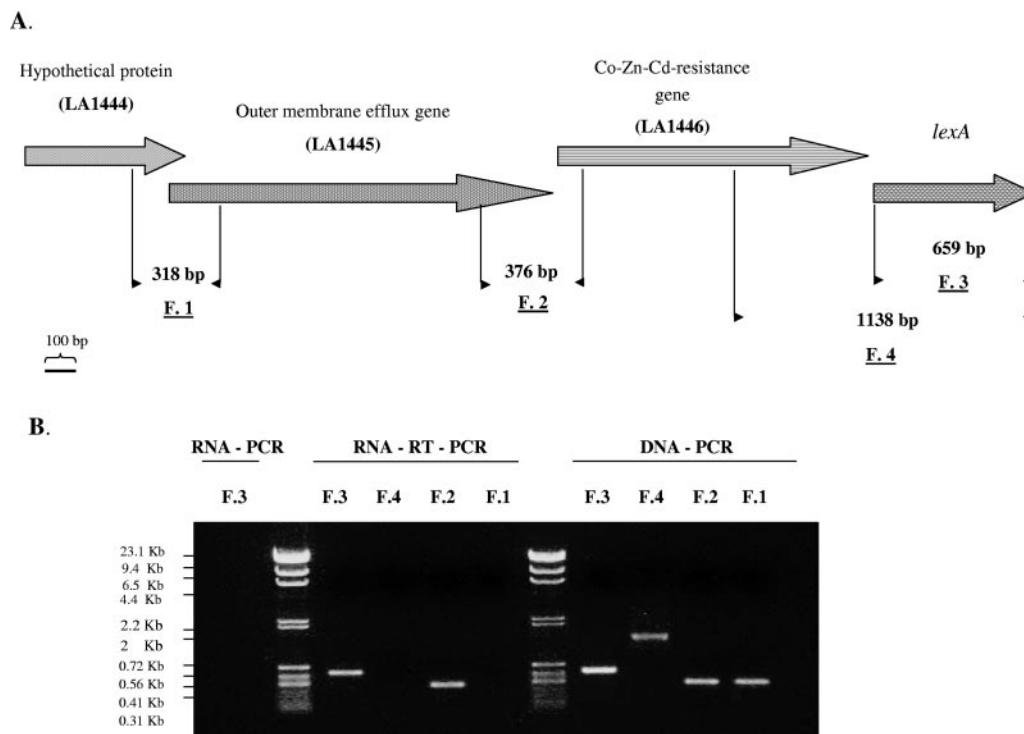


FIG. 1. (A) Genetic organization of the *L. interrogans* *lexA* gene region. The sizes of the fragments (F.1 to F.4) obtained with several primers used are indicated. (B) RT-PCR analysis of putative transcripts covering regions F.1 to F.4 in *L. interrogans* (RNA-RT-PCR). As controls, PCR experiments were carried out with the same primers, but without reverse transcriptase, with either RNA (RNA-PCR) or DNA (DNA-PCR) as a template. The positions of standard DNA size markers are shown on the left.

\* Corresponding author. Mailing address: Department of Genetics and Microbiology, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain. Phone: 34-93-581 1837. Fax: 34-93-581 2387. E-mail: jordi.barbe@uab.es.

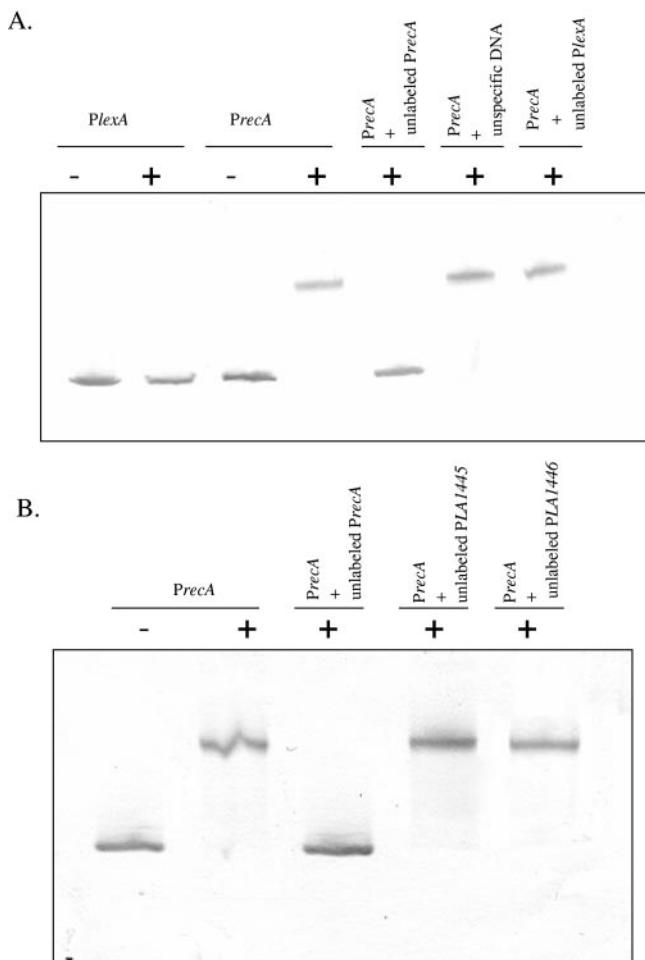


FIG. 2. (A) Electrophoretic mobilities of the *lexA* and *recA* promoter regions in the absence (–) or presence (+) of 40 ng of purified *L. interrogans* LexA protein. The effect of a 300-fold molar excess of unlabeled *recA* (*PrecA*) or *lexA* (*PlexA*) promoter regions as well as of nonspecific DNA on the electrophoretic mobility of the *recA* promoter region is also shown. (B) Effect of a 300-fold molar excess of unlabeled LA1445 or LA1446 promoters on the electrophoretic mobility of the *recA* promoter in the presence (+) of 40 ng of purified *L. interrogans* LexA protein. As controls, the mobilities of the *recA* promoter in the absence (–) of LexA protein or in its presence with and without the unlabeled *recA* promoter are also shown.

bacterial species. This system facilitates cell survival when massive DNA damage occurs and the normal DNA replication of the bacterial cell is disturbed (20). Genes belonging to the SOS system are directly repressed by the LexA protein, which has two clearly differentiated domains: the N-terminal domain, which specifically recognizes the operator (known as the SOS box) at which it binds, and the C-terminal one containing the amino acids Ser and Lys involved in the autocatalytic cleavage which takes place between Ala-Gly residues (12). In the absence of DNA damage, LexA binds the SOS box present in the promoter region of the SOS genes, thereby blocking their transcription. Different sequences for the SOS box, which seems to be conserved for each bacterial evolutionary division, have been described (4, 5, 21). Conversely, in the event of DNA damage, the product of the bacterial *recA* gene acquires

an active conformation after binding to single-stranded DNA fragments generated either by DNA damage-mediated replication inhibition or by enzymatic processing of broken DNA ends (18). Upon activation, the RecA protein promotes autocatalytic cleavage of the LexA repressor through a mechanism similar to that observed for serine proteases (11), thus initiating a global induction of the SOS response. After DNA lesions have been repaired, activated RecA protein concentration declines and noncleaved LexA returns to its usual levels, inhibiting again the expression of SOS genes. In all bacterial species studied so far, the product of the *lexA* gene also regulates its own transcription. On the other hand, the *recA* gene is not always under the negative control of the LexA protein (8).

Spirochetes are helically coiled, gram-negative bacteria present in many microhabitats and may be free living, commensal, or parasitic. This bacterial order is important because of its many pathogenic species but also because of the role that the serial endosymbiotic theory has assigned to these organisms in the origin of the eukaryotic cell (13). The genome sequences of several spirochetes have been reported: *Borrelia burgdorferi*, *Treponema pallidum*, *Treponema denticola*, and *Leptospira interrogans* (6, 7, 15, 17, 19). Among them, only *L. interrogans* has a gene whose deduced product shows both DNA-binding and serine-protease domains compatible with those of a LexA-like protein structure. Importantly, other regulatory proteins, such as prophage and integrating-conjugative-element repressors, may also display a serine protease-like domain (1, 16). Then, when bacterial genome sequences are analyzed in silico, some of these other regulatory genes may be inappropriately annotated as *lexA*. In the present work, the putative *L. interrogans* *lexA* gene was cloned and its protein product purified to determine if it is actually a LexA repressor, as well as to further understand the evolution of the LexA regulon in spirochetes.

**Characterization of the transcriptional organization of the *L. interrogans* *lexA* gene.** To determine the component genes of a bacterial regulon, the sequence to which any regulatory proteins bind must first be identified. All LexA proteins from bacterial species analyzed so far are able to regulate directly their own transcription, even in polycistronic transcriptional units. Analysis of the *L. interrogans* serovar Lai genome region containing the putative *lexA* gene revealed two open reading frames (LA1445 and LA1446) upstream of the *lexA* gene (Fig. 1A).

In order to establish the genomic region in which the *L. interrogans* LexA box might be located, it was necessary to test if these three genes were cotranscribed. To perform this, reverse transcriptase (RT) PCR analysis of total RNA from *L. interrogans* serovar Lai cells, grown in EMJH medium, was carried out as described previously (3) with a set of appropriate oligonucleotide primers based on the *L. interrogans* sequence (17) designed to amplify fragments of suitable size in the event that *lexA*-LA1446 or *lexA*-LA1446-LA1445 was transcriptionally linked (Fig. 1A). Data showed that transcription of the *L. interrogans* *lexA* gene is not linked with that of LA1445 or LA1446 (Fig. 1B). On the other hand, LA1445 and LA1446 open reading frames constitute a single transcriptional unit (Fig. 1B).

Coding strand

Non-coding strand



+LexA



AATTTTGGCTATACAAATTAC

GTATTTGTATAGCAAAATT

FIG. 3. DNase I footprinting assay with coding and noncoding Q5-labeled strands of the DNA fragment containing the *L. interrogans recA* promoter in the absence (-LexA) or in the presence (+LexA) of purified LexA. The translational start codon is shown in italics and underlined. The palindrome contained in the protected sequence is indicated in bold and underlined.

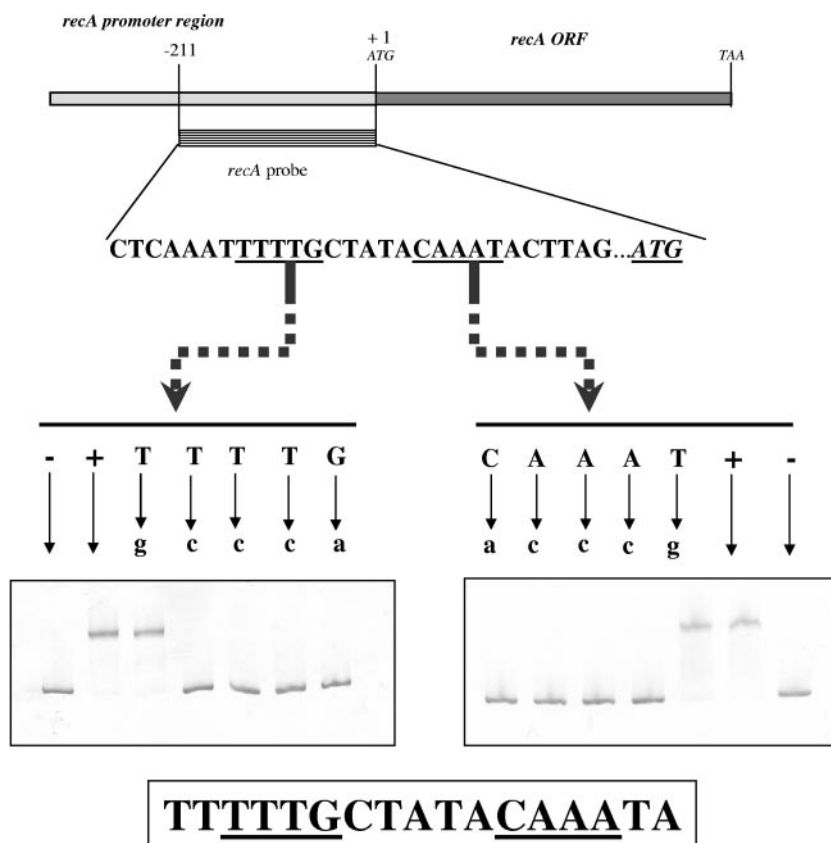


FIG. 4. Single-nucleotide substitutions in the TTTTGCTATACAAA palindrome and their effect on the electrophoretic mobility of the *L. interrogans* *recA* promoter region in the presence of purified LexA. The mobility of the wild-type *L. interrogans* *recA* promoter region in the absence (-) or presence (+) of LexA is also shown.

#### Identification of the *L. interrogans* LexA binding sequence.

Results described above allowed us to conclude that a putative *L. interrogans* LexA box might be located between the end of LA1446 and the beginning of *lexA*. To confirm this, the *L. interrogans* *lexA* gene was amplified by PCR using a forward primer (NdeI-*lexA*) containing an NdeI restriction site which incorporated the *lexA* ATG start codon. The reverse primer (BamHI-*lexA*) started 30 bp downstream of the *lexA* stop codon. The 659-bp PCR fragment containing the *L. interrogans* *lexA* gene was cloned into pGEM-T and then into the pET15b expression vector. The pET15b derivative containing the *lexA* gene was transformed into the *Escherichia coli* BL21(DE3) CodonPlus strain for overexpression of the LexA protein, which was subsequently purified using the Talon metal affinity resin kit (Clontech) as described previously (14). The purity of the LexA protein was >95% as determined with Coomassie blue staining of sodium dodecyl sulfate-polyacrylamide gel electrophoresis (15%) gels (data not shown) following standard methodology (10).

Likewise, and using LpLexAup (5'-CGAAACGGAAAAC TAAAACCG-3') and LpLexAdwn (5'-CTCAATTGCTTTG A GATGATC-3') primers, a 490-bp fragment containing the region upstream of *lexA* and 202 nucleotides of the 3' end of LA1446 was obtained. When this fragment was used in electrophoretic mobility shift assay experiments with purified *L.*

*interrogans* LexA protein, no specific shift in electrophoretic mobility was detected (Fig. 2A).

Since transcription of many bacterial *recA* genes is negatively regulated by LexA, we analyzed whether the *L. interrogans* *recA* region binds leptospiral LexA. Figure 2B indicates that the *L. interrogans* LexA protein binds to a 225-bp fragment containing the upstream region of *recA*. This DNA-protein interaction is specific because it was abolished by an excess of unlabeled *L. interrogans* *recA* promoter but not when an excess of nonspecific DNA was added (Fig. 2A). Likewise, and in accordance with data presented above, the presence of unlabeled *L. interrogans* *lexA* promoter region did not eliminate this DNA-LexA complex (Fig. 2A).

A possible explanation for the negative results obtained in the RT-PCR analysis (Fig. 1A) concerning the transcriptional organization of the leptospiral *lexA* gene could be a consequence of a dramatically low basal expression of this gene. If this were the case, a leptospiral LexA box could still be present upstream of either LA1445 or LA1446. To test this possibility, two unlabeled DNA fragments, containing the region upstream of either LA1445 or LA1446, were used as competitors against the LexA-*recA* promoter region mixture. Results showed that neither region inhibited the binding of LexA to the *recA* promoter, consistent with the absence of a leptospiral LexA box upstream of these two open reading frames (Fig.

2B). All these data supported the conclusion that the *L. interrogans* *lexA* gene is not autoregulated but that it does encode a LexA repressor because, as described for many other LexA proteins, its product is able to bind the *recA* gene promoter region.

The sequence to which the *L. interrogans* LexA protein binds was localized precisely through footprinting experiments with the *recA* promoter region using the ALF sequencer (Pharmacia Biotech) as described previously (2). The results showed that a core region of 20 nucleotides (AATTTTTGCTATACA AATAC), centered at the -206 position with respect to the hypothetical translational start point of *recA*, was protected when both coding and noncoding strands of this gene were analyzed (Fig. 3). Inspection of this sequence revealed the presence of a perfect palindrome whose left and right halves are TTTG and CAAA, respectively.

To confirm the footprinting results and to determine if this palindrome is the sequence recognized by the leptospiral LexA, point mutations were introduced into each of the nucleotides, as well as in immediate flanking regions. Directed mutagenesis of the *L. interrogans* *recA* promoter was carried out by PCR as described previously (2), and the presence of DNA changes was confirmed by sequencing on an ALF sequencer (Pharmacia Biotech). The results showed that leptospiral LexA binding requires both halves of the palindrome, since mutagenesis of nucleotides in either half diminished the formation of the DNA-LexA complex (Fig. 4). As expected, no copy of this palindrome was found upstream of the *L. interrogans* *lexA* gene when a search using the RCGScanner informatics program (4) was carried out. Furthermore, none of the other *L. interrogans* canonical genes belonging to the bacterial SOS system (*uvrAB*, *ruvAB*, *ssb*, and *recN*) (4, 5) have a copy of the TTTGN<sub>5</sub>CAAA palindrome in their promoter regions, indicating that the only gene directly regulated by the LexA repressor in this organism is *recA*. Moreover, analysis of the recently sequenced genome of *L. interrogans* Copenhageni Fiocruz L1-130 (15), a strain different from that studied in this work, also shows the presence of the TTTGN<sub>5</sub>CAAA palindrome upstream of *recA* but not upstream of *lexA*.

The fact that only the *recA* gene is under the direct control of LexA in *L. interrogans* could indicate an intermediary stage in the process of genomic reduction which seems to occur in spirochetes. Thus, the presence of a *lexA* gene in *L. interrogans* appears to be an evolutionary step prior to the loss of this gene in pathogenic spirochetes such as *T. pallidum*, *T. denticola*, and *B. burgdorferi*, whose genomes are significantly smaller than that of *L. interrogans*. Furthermore, it must be noted that the absence of a *lexA* gene in these three spirochetes may increase their fitness against the continuous DNA damaging host defense factors because they exhibit constitutive expression of genes belonging to the recombinational DNA repair pathway (*recA*, *ruvAB*, and *recN*), which is the principal mechanism for DNA damage repair in chromosomes of pathogenic bacteria (9).

This work was funded by grants BFU2004-02768/BMC from the Ministerio de Educación y Ciencia de España and 2001SGR-206 from the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya and by a program grant from the National Health and Medical Research Council, Canberra, Australia. J. Cuñé was recipient of a predoctoral fellowship from the Ministerio de Educación y Cultura, and S. Campoy is recipient of a postdoctoral contract from INIA-IRTA.

We are deeply indebted to Joan Ruiz and Pilar Cortés for their excellent technical assistance.

#### REFERENCES

1. **Beaber, J. W., B. Hochhut, and M. K. Waldor.** 2004. SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* **427**:72–74.
2. **Campoy, S., M. Fontes, S. Padmanabhan, P. Cortes, M. Llagostera, and J. Barbe.** 2003. LexA-independent DNA damage-mediated induction of gene expression in *Myxococcus xanthus*. *Mol. Microbiol.* **49**:769–781.
3. **Cullen, P. A., S. J. Cordwell, D. M. Bulach, D. A. Haake, and B. Adler.** 2002. Global analysis of outer membrane proteins from *Leptospira interrogans* serovar Lai. *Infect. Immun.* **70**:2311–2318.
4. **Erill, I., M. Escribano, S. Campoy, and J. Barbé.** 2003. *In silico* analysis reveals substantial variability in the gene contents of the gamma Proteobacteria LexA-regulon. *Bioinformatics* **19**:2225–2236.
5. **Erill, I., M. Jara, N. Salvador, M. Escribano, S. Campoy, and J. Barbé.** 2004. Differences in LexA regulon structure among Proteobacteria through *in vivo* assisted comparative genomics. *Nucleic Acids Res.* **32**:6617–6626.
6. **Fraser, C. M., S. Casjens, W. M. Huang, G. G. Sutton, R. Clayton, et al.** 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* **390**:580–586.
7. **Fraser, C. M., S. J. Norris, G. M. Weinstock, O. White, G. G. Sutton, et al.** 1998. Complete sequence of *Treponema pallidum*, the syphilis spirochete. *Science* **281**:375–388.
8. **Jara, M., C. Nuñez, S. Campoy, A. R. Fernández de Henestrosa, D. R. Lovley, and J. Barbé.** 2003. *Geobacter sulfurreducens* has two autoregulated *lexA* genes whose products do not bind the *recA* promoter: differing responses of *lexA* and *recA* to DNA damage. *J. Bacteriol.* **185**:2493–2502.
9. **Kline, K. A., E. V. Sechman, E. P. Skaar, and H. S. Seifert.** 2003. Recombination, repair and replication in the pathogenic *Neisseria*: the 3 R's of molecular genetics of two human-specific bacterial pathogens. *Mol. Microbiol.* **50**:3–13.
10. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685.
11. **Little, J. W.** 1991. Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease. *Biochimie* **73**:411–421.
12. **Luo, Y., R. A. Pfuetzner, S. Mosimann, M. Paetzel, E. A. Frey, M. Cherney, B. Kim, J. W. Little, and N. C. Strynadka.** 2001. Crystal structure of LexA: a conformational switch for regulation of self-cleavage. *Cell* **106**:585–594.
13. **Margulis, L.** 1996. Archaeal-eubacterial mergers in the origin of Eukarya: phylogenetic classification of life. *Proc. Natl. Acad. Sci. USA* **93**:1071–1076.
14. **Mazon, G., J. M. Lucena, S. Campoy, A. R. Fernández de Henestrosa, P. Candau, and J. Barbé.** 2003. LexA-binding sequences in gram-positive and cyanobacteria are closely related. *Mol. Gen. Genomics* **271**:40–49.
15. **Nascimento, A. L., A. I. Ko, E. A. Martins, C. B. Monteriro-Vitorello, et al.** 2004. Comparative genomics of two *Leptospira interrogans* serovars reveals novel insights into physiology and pathogenesis. *J. Bacteriol.* **186**:2164–2172.
16. **Pabo, C. O., R. T. Sauer, J. M. Sturtevant, and M. Ptashne.** 1979. The  $\lambda$  repressor contains two domains. *Proc. Natl. Acad. Sci. USA* **76**:1608–1612.
17. **Ren, S. X., G. Fu, X. G. Jiang, R. Zeng, Y. G. Miao, H. Xu, Y. X. Zhang, et al.** 2003. Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. *Nature* **422**:888–893.
18. **Sassanfar, M., and J. W. Roberts.** 1990. Nature of SOS-inducing signal in *Escherichia coli*. The involvement of DNA replication. *J. Mol. Biol.* **212**:79–96.
19. **Seshadri, R., G. S. Myers, H. Tettelin, J. A. Eisen, J. F. Heidelberg, et al.** 2004. Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. *Proc. Natl. Acad. Sci. USA* **101**:5646–5651.
20. **Walker, G. C.** 1984. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol. Rev.* **48**:60–93.
21. **Winterling, K. W., D. Chafin, J. J. Hayes, J. Sun, A. S. Levine, R. E. Yashin, and R. Woodgate.** 1998. The *Bacillus subtilis* DinR binding site: redefinition of the consensus sequence. *J. Bacteriol.* **180**:2201–2211.



## **AGRAÏMENTS**

## AGRAÏMENTS

La realització i presentació d'aquesta tesi, aparentment sembla ser, el final de la meua etapa formativa, amb categorització com a tal, ja que pel que diuen, reciclar-se o morir.

Com a tal, també acaba una etapa vital, en la que hi han participat i col·laborat gran quantitat de persones, no sols a nivell acadèmic, sinó evidentment personal. M'agradaria agrair a totes i cada una de elles el paper que han jugat durant tots aquests anys, fent així una mica de justícia a la seva aportació. Així caldria començar pels meus pares i la meua germana, per tot el que m'han donat, durant tota la vida, començant per una família i acabant per tots els mitjans per a poder créixer com a persona, tant en el camp acadèmic com en qualssevol altra vessant per la que hagi mostrat interès. Aquest agraïment es fa extens a tots els integrants de la meua família: avis, tiets, cosins...

També, agrair de manera especial a la Núria Busquets. Agrair-li tota la seva paciència i amor durant els moments difícils, tant del doctorat, que són molts, com de qualssevol altre tipus. Gràcies a tots.

També voldria agrair a tots els professors que he trobat en tots aquests anys d'estudis, començant pels del desaparegut EGB a Guissona, el desaparegut BUP i COU a Cervera (la meua tieta Blanca també, tot i les cigarretes que vaig haver de llençar), o en la llicenciatura de Biologia. També citar la Lourdes *for his perfect english technical assistance* quan ha estat necessari.

Pel que fa a la realització d'aquest doctorat, agrair al Dr. Jordi Barbé la concessió de la possibilitat de realitzar aquesta tesi doctoral, junt amb els coneixements, fruit de l'experiència i estudi, transmesos durant tot aquest temps. També mostrar el meu agraïment a la Dra. Montserrat Llagostera, i per extensió, a tot el Departament de Genètica i Microbiologia. En tant els companys de lab, començar pel Toni, que va ser la persona amb qui em vaig iniciar en els meus coneixements tècnics en la biologia mol·lecular. Gràcies. Al Gerard Mazón, actualment envoltat de sabó de gran qualitat. També em va donar un cop de mà quan va ser necessari, a part de pinzellades de *Trivial*. La Mònica, per les seves visions serenes i la seva gran energia encomanadissa. Al Marc, noi extraordinàriament tranquil de Rubí, i amb qui vaig prendre contacte amb el *lab* simultàniament, ja fa uns quants anys. Informàtic oficial que ens ha tret de més d'un *apuru*. La Noelia, directament importada de *La Comarca*, junt amb el Marc, van venir de *pack*. Una bona noia, sempre disposada a donar un cop de mà. El Joan, també es pot dir *Juà*. Gràcies pels seus consells i ànims, i per les bones estones compartides, bromes incloses (com

quan trucaven de Consergeria...). La Susanna Campoy, amb qui varem coincidir en una festa romana de Guissona sense conèixer-nos encara, coses de la vida... Moltes gràcies per tenir sempre una resposta i per amenitzar els àpats i les estones de *poiata* amb rocamboliques històries. L'Alfonso, amb qui vaig coincidir poc, un començava l'altre acabava, però de qui m'enduc un bon record, igual que el José Antonio. L'Anna Hervàs, per la seva alegria, les seves ganes d'ajudar i per tenir sempre un somriure. La Sònia, per la seva rialla encomanadissa i la seva serenitat, tranquil·litat i saber escoltar. La Laura, última incorporació al Lab 1, noia molt agradable a la que espero que li vagi molt bé aquesta etapa doctoral.

Pel que fa al *lab 2*, tot i que ja fa un temps que varen deixar el vaixell, fer referència a la Montse Bosch, la Lorena, l'altra Montse (la Rebollo) i la Mirle. Les quatre em van ajudar en els meus inicis, tant en la integració en el grup com en el camp de treball. A totes gràcies. L'Anna Bigas, noia extrovertida, alegre i que fa *sentir la seva veu*, que tot i que li recordem, sap que s'agraeix. La Vanessa, junt amb l'Anna les *Fred Aster* del Lab, amb qui he rigut i compartit opinions, potser per proximitat geogràfica, en ser *del tros de Catalunya que tenim per Aragó...* La Maria Elena, amb un riure que fa més alegre tant el treball com els àpats, sense oblidar el personatge Zuñiga. El Jesús, deixa'l anar, sembla que no però... Un bon noi que li desitjo el millor. El Gerard Àlvarez, potser encara més marcat que el Jesús, sembla que no però... També el millor en aquesta etapa. El doctor José González, que després de temps d'enyorança i treball té el seu merescut premi. La Pilar Cortés, per la seva experiència, força i per les tertúlies mantingudes (tot i que sovint amb visions bastant diferents). La Susana Escribano, *Escribanyo!! Sempre amb presses!! Calma dona!!* També gràcies per totes les estones compartides i cops de mà donats. La Isabel, per la seva vitalitat, opinió, criteris, i els cafès compartits amb les dues!! A la Núria Tomàs, l'Olga, la Núria Viguès, la Mayra, el Jordi Mas i la Núria Gajú (*lab 3 i 4*) i per extensió a totes les persones que m'he trobat aquests anys en el Departament. Gràcies a tots.

Com que la meua vida aquests anys com en la vida en general no tot ha estat feina i treball, voldria fer esment als membres de la Colla, els amics de Guissona, déu ni do quin personal!!! Segurament, estan entre les persones amb qui he passat més hores de la meua vida: ja fos al *Centro*, o als locals que hem tingut... i alguns arrasat (La Farinera, Cal Manyà, La Granja, el Local de Tapioles) al Gatzara o al Xaplin, agafant el cotxe per anar a Cervera, Tàrrega, a *Molle* més cap allà o a mil llocs més!! Bé, espero no deixar-me ningú. Així per exemple, i per estricte ordre del llistat que tinc al davant, podríem citar el Sr. Vidal: bastant, bastant *girat*, però un bon noi. Escriptor amb un llibre publicat i preparat per a fer el sal a la fama. Crec que té totes les qualitats per a fer-ho. Així en honor a les seves equacions humilment proposo la següent: Sergi Vidal = Joaquín Sabina + Quim Monzó + Ed Gein (*xit*, no estem parlant d'aparença física....quina quimera seria!!).



El Jordi Vilaplana, *Fu*. Un *tio* alegre, extravertit, molt treballador i una mica eixelebrat! Amb un do, que el permet dormir 2 hores i apuntar-se a tot com el qui més. Ànima i animador de moltes de les nostres trobades massives. Un *bon partit*, que s'ha adjudicat la que és la seva xicoteta de Cervera, la Pilar... I ja té cotxe!!

El Joan Condal, *Xipu* (nom artístic complert: *Xipulin d'Ubrique*). Un ilergeta, amb arrels a Iesso. Mestre del Soccer (fins que ens varem viciar el que s'havia viciat ell), persona amb qui es passen bones estones i amb certs problemes al canell. El que és diu un bon noi, que ja no vèiem tant (bàsicament a les Festes, diga-li tonto). No cal ni dir-ho, Rei dels *mutis*.

L'Eva Guasch, una gran persona i una amiga de tota la vida, que ara abandona un veïnatge de 15 anys, tot i que a Guissona, tampoc pots anar molt lluny. *Bueno*, cada cop més. Persona amb molta creativitat, que empra com a aptitud de treball en el seu estudi de disseny. I també al seu xicot, el Jordi, un xicot de Sabadell que esperem veure més sovint ara que ja no depèn tant de les magnífiques comunicacions lleidatanes (no només hi ha centralisme a Madrid...).

*Lo Pipe*. Sí, crec que és diu Jordi Ribera o així....Un *tio* amb seny i barman en el seu pub que tenim bastant *apubmat*!! Bon noi, però és de l'Espanyol (igual que el Vidal) i això vulguis que no, treu punts. Pocs, però en treu, més o menys els mateixos a que es queda del descens cada any l'*Apanyó*. Però bé, després d'haver compartit equip tant en futbol sala com en futbol i tot el viscut junts, se li perdona.

El Xavier Bureu, *Jony* (*Jhonny*, si es vol més rockavil.li). Que en diríem del Xavier? El podríem fer Màrtir de la diversió i la desobediència. Encara el puc veure aguantant estoicament, bé, potser amb un punt de rebel·lia, les estocades, de la mai prou agraïts Antonia: mestressa de la *Granja*, i que presentava al Xavier, el llistat de queixes i despropòsits que fèiem tots. No oblidarem mai *Esos maravillosos años* per la *Granja*, fins i tot quan encara no era *Granja*. Bé, aquests anys i tots els altres, és clar.

El Gerard. *En fin*. El que es diu un boig en el sentit sa de la paraula , perquè quan toca *currar* no hi ha qui el calmi, però quan no toca treballar, pot ser encara es més perillós. Com d'altres dels d'aquí citats, potser ens coneixem d'abans que tinguem records, i amb els quals hem compartit totes les etapes de la nostra vida, començant pel Col·legi de Guissona, l'institut de Cervera (en la nostra època no n'hi havia a Guissona...i millor!), i d'un pèl que no anem el primer any junts a la Vila. Crec que en el fons varem tenir sort. El nombre d'experiències junts crec que és incomptable, i ens ho hem passat extraordinàriament bé.

El David Beneit, el *Beny*. El que és diu un bon xicot!! Una ànima de blues a ritme d'havanera. La quantitat de *bolos* que ha fet arreu del nostre país...Ara que no es estrany que hagi tingut temps d'anar tant amunt i avall, diguéssim que és un noi que optimitza els trajectes modulant la variable velocitat, però controlant, *eh que sí?* Tot i que ara el tenim de nou a Guissona, es deixa veure menys que abans quan estudiava a Tarragona, que és més o menys per on l'han *caçat*....la

Gemma, *fitxatge* del grup bastant recent, però que per la seva extraversió, es com si fes molt temps.

El Joan Morros, el *Nyua*. Si féssim un pessebre seria St. Josep: persona bondadosa, amb paciència, tranquil i fuster! No sé si li colaria el tema de l'Esperit Sant...Com molts dels d'aquí citats, amics de tota la vida: literalment tota la vida. La quantitat d'hores compartides jugant a bàsquet, futbol, futbol sala, tennis, escoltant música o anant a alguna festa... Alguna. Espero no trenca mai la relació, almenys fins que tingui els mobles!!!

El 300. Altrament dit, el *Manyà*. El trobaríem acompanyant al Xavier i al *Fulli* (veure Alexandre Pujol) en el retaule, dels qui patiren en nom de tots, del nostre temple de joventut. També, va *pillar*, les queixes (amb perspectiva, justificades) dels seus progenitors com a propietaris de l'immoble que allotjava també la nostra disbauxa! – *Ja són macos ja, però una mica esvalotats*. A part d'això, un *tio* ben parit, tranquil i total, absoluta i completament imprevisible.

L'Albert Vendrell, el *Bruce*. Banda sonora oficial (junt amb els desapareguts grups locals Farm's juice, Sex-Garra, Atònik, Avernus i algun altre...), amb registres que han anat evolucionant. DJ oficial de les revetlles assenyalades (o no) i espontani en moltes altres ocasions. Dels primers en tenir cotxe, fet que varem explotar per anar amunt i avall... Una persona amb fusta d'actor i amb una gran oïda per la música, que es deu haver trinxat a cop de decibel. També a la seva xicoteta, la Laura de Cervera, que ja farà, igual que la Pilar, uns dos anys que la coneixem, tot i que *de vista* de l'Institut, ja fa més....*I reina com passa el temps!*

La Mireia Ortiz, la *Mire*. Quantes hores compartides als transports públics: el bus de l'Institut, el *Borriquero*, l'Alsina Graells, la Hispano Igualadina...però que varem aprofitar amb gratificants converses i estones divertides. Què faràs si no amb 2 o 3 llargues hores pel davant! Aquesta relació també es va veure incrementada (tot i els molts anys d'amistat), pel fet de que també era habitant de la Vila Universitària (déu ni do, amb 2.000 persones entre 18-25 anys, no es pot esperar res de bo. Res). Ara la tenim per Madrid, *masteritzant-se*, tot i que esperem que torni aviat, i no es quedi als *madriles* definitivament. Gràcies també, per haver fet possible viure aquest últim 0-3 a Madrid. Experiència molt recomanable!

L'Alexandre Pujol, *Fulli* (de fullero). Tercer element del retaule esmentat anteriorment, per haver-nos cedit, junt amb els seus pares, un dels pisos de l'antiga farinera de Guissona. Quina felicitat aquells anys!! Pocs, pocs mals de cap...fer el *workbook* i poca cosa més. Ara torna a ser Guissona després dels seus periples per Girona. Persona espavilada, amb sentit de l'humor i amb un gran talent per dibuixar i treure punta a les coses. Un bon noi. L'Anna Rosell, Anna o potser *rassel*. La ubico al cantó de l'Alex, ja que així ha estat des de temps immemorials, tants potser com anys fa que la conec. Persona alegre, simpàtica, viva i una mica *fullera* també, ja es sap: Déu els fa i ells s'ajunten...També molt ordenada! ja que mai t'agraïrem prou (bé, jo ho faig

aquí una mica) les hores dedicades a que el local *fes una mica més de goig*, després de cada nit, tot i que últimament el *Vidalet* pedala eh...

Continuant amb el *clan Pujis*, el Ricard Pujol, el *Ritxi*. Baixista destacat del grup Atònik, persona amb inquietuds on les hi hagi i cul del Jaumet oficial! *Perdó, aquest noi no es pot desendollar? // Quins nois més atrafegats, no paren! No senyora, és un però molt nerviós.* Sempre que varem voler ens va obrir les portes, amablement com ell és, del seu pis de Pamplona, on va cursar els seus estudis, fet que no varem desaproveitar. Actualment es tot un farmacèutic a Sanauja, però no perd els lligams amb el *Nord*, gràcies a la seva xicota, la Sara: noia molt agradable que tot i la distància vèiem força, i que ja fa força temps que tenim la sort de tractar.

Com a últim *Puji*, que no el menys important, el Lluís Pujol, *Lluizao*. La persona més tranquil·la i pacífica que conec, tot i ser un erudit en conflictes bèl·lics i armament (també cal dir que torna bastant vermell del *paint ball*... i la pintura és blanca). Després de les seves aventures a Barcelona i Lleida, ha fet de fill pròdig, i també de nou el tenim a Guissona. A part de la seva amistat, no oblidaré mai i romandran a la meva memòria (i crec que ho puc fer extensible a tothom) les *performances* amb que ens ha obsequiat..... simplement impagables!!

La Judit Farran, *Yudí* (si es vol fer-la empipar). És la *mare de la colla* (això també es per fer-la empipar) i amb la qual tenim tracte (tot i que com a colla fou més tard) des de molt petits: tant per la seva germana, amb qui anàvem a classe, com per estudiar tant al mateix Col·legi com el mateix Institut. Una noia esportista, disciplinada i eficient, que ja fa tants anys que festeja amb el Ramón Muntada (*Món*), com el *Vasqueta* i la Ramona! A veure si fem el cop de cap ja!!

El *Món*, un xicot intel·ligent incomprès per l'obsolet sistema acadèmic, però que actualment, ja és tot un enginyer industrial. També fa molts anys que el coneixem, des de ben petits, i amb el qual hem compartit moltes hores, especialment d'estiu, tant a la piscina com als *locals*, amb horaris diversos, així com un munt d'hores el meu primer any per aquestes terres barcelonines. Un bon *tio*, tranquil, amb opinió pròpia i principis.

El David Puig: *Patx* o *Patxi*, a escollir. Un vampir, entre setmana i el cap de setmana no perdia el ritme. Persona a qui no li fan por els kilòmetres, si el fi justifica els mitjans (ja ens entenem), siguin amb cotxe o *volant*. Un xicot de veure-les vindré, potser per jugar de porter des de petit, i amb una gran capacitat d'optimitzar recursos, especialment energia (pròpia), tot i que quan toca *currar* tocar *currar*.....Ja que és veritat el que diuen, que els cansats fan la feina. Un bon noi, amb qui també he compartit un munt de coses, i per tant, i això es característic seu, he rigut molt.

L'Anna Palau, o *Font de la Palau*. Vindria ser una mica la versió del Ricard en femení. Tampoc sap està quieta. Per exemple, tant la podies veure treballant al sector terciari d'oci nocturn urgellenc d'Agramunt, com encara no 6 hores després trobar-te-la amb la missió de socórrer qualssevol incident a la piscina. Incompatible? Jo diria que sí. Una noia com cal, que

últimament no es deixa caure massa per Guissona, degut al seu idil·li amb un xicot, l'Albert, i la seva ciutat, Reus. Un noi agradable i tranquil. Tots els ànims i sort ara que està passant una situació familiar difícil.

El Ferran García, *Ferritxes*. Esperit lliure i pròxim cap del Polisario. Començar els estudis de Filosofia, però tot i que acadèmicament no és filòsof, diria que potser és dels alumnes que millor varen captar l'essència...i amb poques classes que ho va fer!! Persona culta, espavilada, amb sentit del humor, i amb alguna anada d'olla...perquè caram no has vingut els últims 10 anys a la Festa de Cap d'any...? Ja sé, ja sé els teus arguments, però no ho acabo de veure clar...

L'Eva Suárez, a part dels anys de colla, hem acabat compartint durant un temps universitat (tot i que tothom va tant liat...), i ara també es troba en el Vallès, com un servidor, en la dura vida de l'*emigrant*, amb una feina que demostra que ha de tenir més paciència que un Sant. Tot i així, també sap que com a casa enlloc, i així anem coincidint, i espero continuar fent-ho per Guissona.

L'Albert Manaut, *Manau*. Tot i que els llibres d'història segurament no recolliran les seves cites (o potser sí, mai es sap), seria necessari: ens ha donat asseveracions dignes d'un savi, i que passarem de generació en generació (o potser no...). Noi amb ganes de divertir-se, sempre amb una bona cara per tothom i membre actiu del que es coneix des de fa temps *com el jovent de Tarroja*. El tenim per la Segarra mes septentrional, Torà, amb la seva parella, la Montse, amb qui ja fa un bon grapat d'anys que comparteix la vida: noia molt agradable i amb bones paraules sempre també.

Ubicats prop del final d'aquesta enumeració, com la z de *Zárágózá*, trobem el Diego (*Diegues* o Dídac si s'adherís a la normalització. Ai nen amb l'Estatut!) i l'Alicia. Amb el primer d'ells fa molt temps que ens coneixem, molt que vol dir uns 26 anys per haver estat veïns i més tard company de colla. Company d'hores i hores de riure tant a l'Institut, el *bus* (un record pel *Janma*) com en els moments de Colla. Persona amb gran sentit de l'humor, amb un gran ironia i amb qui es passen estones molt agradables. Els últims anys no l'hem vist tant degut a trobar-se a Saragossa, amb els estudis de Filologia i també per què no dir-ho, per l'Alicia, la seva xicota i que a diferència de la Judith, si que és la mare de la colla! No ens enfadem, que no ho sembla! Ara la trobem per Lleida amb la seva carrera docent que esperem que vagi molt bé, assolint la plaça que desitgi. Destacar la seva alegria, extraversió i *bibatxeria* (ho cito aquí però val per tots, que no citi una qualitat a algú no vol dir que no la presenti. Per aplacar possibles susceptibilitats ;-))

El Xavier i la Mercè Mercader, que tot i que mica en mica, no han aparegut tant per Guissona, els recordem i no oblidem les bones estones que varem passar junts.

La llista podria ser molt més extensa, però la cosa passaria de la *taca d'oli* que ja és cap a vessament, tanmateix m'agradaria citar tots els bons amics que vaig fer a l'Institut de Cervera: el *Cisco*, Xavi, *Radomir*, el Caus, Massana, Pané, Bertrán, la Gasol, la Bonell, l'Estruch,

l'Aldomà, la Carcassona... i per sort molts més!! Així com amics de la Universitat: el Santi, Ramón, Pere, *Pillo*, Ariel, Joan, Lluís, Sandra, Ariadna, Neus, Laia, Elisa, Abraham...Així com el Lluís Capdevila, company de pis durant molts anys, i una persona feliç i *pantxa-contenta* on les hi hagi, o el Marc Vidal de la Fuliola.....déu ni do quin personatge.... Aquells anys de vila! O l'Albert Ribera, consol, ja que també fa més hores que un rellotge...I un llarg etcètera que em sap greu no citar.

A tots moltes, moltes gràcies.