



Universitat Autònoma de Barcelona

Departament de Genètica i Microbiologia

**Evolució del motiu d'unió de la
proteïna LexA al Domini *Bacteria***

Gerard Mazón i Busquets

2004

UAB

Universitat Autònoma de Barcelona

Departament de Genètica i Microbiologia

**Evolució del motiu d'unió de la
proteïna LexA al Domini *Bacteria***

Memòria redactada per Gerard
Mazón i Busquets per optar al grau
de Doctor en Microbiologia per la
Universitat Autònoma de
Barcelona.

Vist i plau

El Director de la tesi

Dr.Jordi Barbé i García

BELLATERRA 2004

*“L’eternitat està enamorada
de les obres del temps”*

W. Blake

*“Si me dijeras, pide un deseo
te pediria un rabo de nube,
(...)
un barredor de tristezas,
un aguacero en venganza,
que cuando escampe parezca,
nuestra esperanza”*

S. Rodríguez

Als meus pares,
a tots aquells i aquelles
que estimo i he estimat,

Resum

El sistema SOS o sistema de reparació d'emergència és una xarxa multigènica que s'indueix en presència de lesions al DNA. Aquest sistema ha estat caracteritzat en diferents espècies bacterianes, grampositives i gramnegatives, determinant-se per a totes elles el motiu d'unió del seu repressor, la proteïna LexA.

En el present treball s'han caracteritzat els llocs d'unió de la proteïna LexA o caixes LexA dels microorganismes *Xylella fastidiosa*, *Anabaena* sp. i *Fibrobacter succinogenes*. Per aconseguir-ho, s'han clonat els seus respectius gens *lexA* i s'han purificat els seus productes mitjançant columnes d'afinitat a una cua d'histidines en els dos primers casos i a la Glutatió-S-Transferasa en el darrer. La utilització d'aquestes proteïnes LexA en assaigs de mobilitat electroforètica (EMSA) i de "footprinting" amb els promotors del gen *lexA* d'aquests bacteris ens ha permès identificar la seqüència a la qual aquesta proteïna s'uneix específicament en cadascun dels tres casos, TTAGN₆TACTA a *X. fastidiosa*, RGTACNNNDGTWCB a *Anabaena* i TGCNCN₄GTGCA a *F. succinogenes*. Aquests tres motius han servit per identificar gens del reguló LexA a l'ordre Xanthomonadales i a les divisions Cyanobacteria i Fibrobacter, i això ens ha permès demostrar una important variabilitat en la composició d'aquest reguló i, també, la presència de gens induïbles davant del dany al DNA de manera independent de LexA a *X. fastidiosa*.

Per determinar la possible evolució del motiu d'unió de LexA al Domini *Bacteria*, s'han dissenyat una sèrie de mutacions a les caixes LexA d'*Anabaena* i *F. succinogenes* per generar-hi derivats amb operadors LexA ja caracteritzats d'altres grups filogenètics i provar la capacitat d'unió de les seves proteïnes LexA. Aquest estudi, conjuntament amb l'anàlisi filogenètica comparativa entre la proteïna LexA i la proteïna RecA (àmpliament conservada al Domini *Bacteria*), ens ha permès reconstruir la història evolutiva de la caixa LexA i demostrar que el motiu present als α -Proteobacteria prové d'un altre grup filogenètic, probablement per una pèrdua del seu gen *lexA* transmès verticalment i una incorporació per transferència

genètica horitzontal d'una nova còpia provinent d'una espècie bacteriana relacionada amb els cianobacteris.

Els següents articles descriuen i discuteixen els resultats sobre els quals està basada aquesta tesi. Es presenten a la secció d'Annexos del treball i es troben al text identificats amb números romans:

- I. **Campoy, S., Mazón, G., Fernández de Henestrosa, A.R., Llagostera, M., Monteiro, P.B. i Barbé, J.** 2002. 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.
- II. **Mazón G., Lucena J.M., Campoy, S., Fernández de Henestrosa, A.R., Candau, P. i Barbé J.** 2004. LexA-binding sequence in Gram-positive and cyanobacteria are closely related. *Mol. Gen. Genomics* **271**: 40 – 49.
- III. **Mazón, G., Erill, I., Campoy, S., Cortés, P., Forano, E. i Barbé, J.** 2004. Reconstruction of the evolutionary history of the LexA binding sequence. *Microbiology* **150**: 3783-3795 (En premsa en el moment de redacció de la Tesi).

ÍNDIX

1. Introducció.....	1
1.1. El sistema SOS a <i>Escherichia coli</i>	3
1.1.1. Inducció del sistema SOS a <i>E. coli</i>	5
1.1.1.1. La proteïna LexA.....	6
i. Estructura de la proteïna LexA.....	9
ii. Funció de la proteïna LexA.....	10
iii. Autohidròlisi de la proteïna LexA.....	12
1.1.2. Funcions SOS	13
1.1.2.1. Reparació per escissió de nucleòtid	14
1.1.2.2. Reparació per recombinació	15
1.1.2.3. Mutagènesi SOS	16
1.1.2.4. Altres funcions del sistema SOS.....	18
1.2. El sistema SOS a altres bacteris	18
1.2.1. El sistema SOS als bacteris Gramnegatius	19
1.2.1.1. Grup dels γ -proteobacteris	19
1.2.1.2. Grup dels α -proteobacteris	20
1.2.1.3. Grup dels δ -proteobacteris.....	20
1.2.2. El sistema SOS als bacteris Grampositius.....	22
1.3. Estudis filogenètics moleculars en bacteris.....	25
1.3.1. Filogènia basada en la presència d' <i>indels</i> a les proteïnes.....	27
1.4. Objectius	31

2. Resultats i discussió.....	33
2.1. Identificació de la caixa d'unió de LexA a <i>Xylella fastidiosa</i>	33
2.1.1. Clonació del gen <i>lexA</i> de <i>X. fastidiosa</i> i sobreexpressió del seu producte	33
2.1.2. Anàlisi transcripcional de <i>lexA</i> a <i>X. fastidiosa</i>	34
2.1.3. Estudis de la mobilitat electroforètica del promotor de <i>lexA</i> de <i>X. fastidiosa</i>	35
2.1.4. Efecte a la unió de la proteïna LexA de mutacions puntuals a la zona operadora	37
2.1.5. Caracterització del reguló LexA a <i>X. fastidiosa</i>	38
2.1.6. Estudi de la inducció dels gens regulats per LexA a <i>X. fastidiosa</i>	40
2.2. Identificació de la caixa d'unió de LexA a <i>Anabaena sp.</i> PCC7120.....	41
2.2.1. Clonació del gen <i>lexA</i> d' <i>Anabaena</i> i sobreexpressió del seu producte	41
2.2.2. Estudis de la mobilitat electroforètica del promotor <i>lexA</i> d' <i>Anabaena</i>	42
2.2.3. Estudi de protecció (“ <i>footprinting</i> ”) de la regió promotora de <i>lexA</i> d' <i>Anabaena</i>	44
2.2.4. Efecte a la unió de la proteïna LexA de mutacions puntuals a la zona operadora	46
2.2.5. Caracterització del reguló LexA a <i>Anabaena</i>	47
2.2.6. Estudis de mutagènesi dirigida per analitzar la proximitat entre els motius d'unió de LexA d' <i>Anabaena</i> i <i>B. subtilis</i>	49

2.3.	Identificació de la caixa d'unió de LexA a <i>Fibrobacter succinogenes</i> S85 ..	51
2.3.1.	Clonació del gen <i>lexA</i> de <i>F. succinogenes</i> i sobreexpressió del seu producte.....	51
2.3.2.	Estudis de la mobilitat electroforètica i assaig de protecció (“ <i>footprinting</i> ”) de la regió promotora de <i>lexA</i> de <i>F. succinogenes</i>	53
2.3.3.	Efecte a la unió de LexA de mutacions puntuals a la zona operadora	54
2.3.4.	Caracterització del reguló LexA a <i>F. succinogenes</i>	56
2.4.	Anàlisi comparatiu de les caixes d'unió de LexA al Domini <i>Bacteria</i>	57
2.4.1.	Estudi comparatiu entre els arbres filogenètics generats amb les proteïnes RecA i LexA	58
2.4.2.	Estudis de mutagènesi dirigida per analitzar la possible evolució del motiu d'unió de LexA des de Grampositius fins a Proteobacteris.	60
3.	Conclusions	62
4.	Bibliografia.....	63
5.	Annexos	
6.	Agraïments	

1. INTRODUCCIÓ

1. Introducció

L'estudi del contingut del missatge hereditari dels éssers vius permet entreveure quins camins divergents ha seguit una hipotètica espècie primera per donar lloc a la diversitat actual. La similitud en les seqüències de DNA o d'aminoàcids entre dues espècies ens indiquen la distància relativa a la que es troben del punt on van començar a divergir d'un ancestre comú. Aquesta informació ha permès establir una filiació d'espècies, una filogènia, ja sigui pel conjunt del seu genoma o, a petita escala, per determinats gens o sistemes gènics. El coneixement de l'evolució d'un sistema genètic ve determinat pel coneixement del seu funcionament, de la seva composició, del seu control, i també per l'estudi comparatiu d'aquest amb sistemes similars d'altres espècies.

La conservació de la integritat del material genètic és una de les funcions més importants d'un organisme viu, ja sigui eucariota o procariota. Els canvis i modificacions a la seqüència del DNA d'una espècie permeten l'evolució, però, en un espai temporal més reduït, l'acumulació de canvis i modificacions al DNA és interpretada per la cèl·lula com un compromís per la seva viabilitat i supervivència. Paradoxalment doncs, podríem dir que l'organisme vetlla per la perpetuació de l'espècie garantint l'estabilitat del DNA i alhora permet un grau sostenible de mutació facilitant una millor adaptació a un medi en constant procés de canvi.

Les espècies bacterianes posseeixen diversos mecanismes que els permeten reparar els diferents tipus de lesions que s'introdueixen al DNA com a conseqüència de la interacció amb agents lesius, ja siguin endògens, producte del mateix metabolisme cel·lular, o exògens deguts a l'exposició a agents químics o físics externs. Segons el tipus de lesions i el grau en què es produeixen, la cèl·lula bacteriana disposa d'un ventall de respostes defensives per restaurar el DNA danyat o atenuar i minimitzar els efectes del dany.

En general, es pot parlar de tres grans tendències de mecanismes de reparació: reversió directa, escissió i recombinació. Es coneixen fins al moment dos mecanismes de reversió directa: la

fotoreactivació i la transferència d'alquils. Són sistemes encaminats a eliminar *in situ* el dany que s'ha produït gràcies a l'acció d'enzims (fotoliasas en el primer cas i alquil transferases en el segon) que reconeixen específicament la lesió i són capaços de catalitzar la seva reparació per revertir el canvi. A la reparació per escissió, la lesió és reconeguda i es duu a terme una acció de substitució de la seqüència afectada, ja sigui de manera exacta sobre la base lesionada (escissió de base) o sobre la cadena que conté aquesta base lesionada (escissió de nucleòtid). La recombinació es dona quan les lesions introduïdes a la molècula del DNA impliquen una complexitat major, amb enllaços entre les dues cadenes del DNA, talls mono o bicatenaris i efectes sobre la cadena motlle. Les vies són diverses i impliquen diferents proteïnes però totes elles utilitzen la proteïna RecA en el pas de l'aparellament de les cadenes o sinapsi.

A *Escherichia coli*, el microorganisme més estudiat dins dels procariotes, s'ha determinat recentment que s'expressen més de 1000 gens davant de la presència d'un agent lesiu com la mitomicina C (Khil i Camerini-Otero, 2002). Molts d'aquests 1000 gens formen part de xarxes multigèniques capaces d'activar-se de manera coordinada segons el tipus de dany o el nivell d'aquest. Les xarxes induïbles per lesions al DNA millor caracteritzades fins al moment són :

1. Resposta adaptativa als agents alquilants

La resposta als agents alquilants com N-metil-N'-nitro-N-nitrosoguanidina (MNNG), metil metanosulfonat (MMS) i etil metanosulfonat (EMS) està integrada en un reguló controlat pel producte del gen *ada*, i inclou els gens *alkA*, *alkB*, *aidB* i el propi gen *ada*. La proteïna Ada té una funció dual, en reparació i control, i la resta estarien implicades en el procés de reparació (Landini i Volkert, 2000).

2. Resposta a l'estrès oxidatiu

És la resposta a les lesions produïdes per radicals químics com el peròxid d'hidrogen (H_2O_2), radicals superòxid (O_2^-) o nitrils (NO) que es generen de manera endògena pel metabolisme aeròbic però també de manera exògena. Existeixen dos regulons caracteritzats, OxyR i SoxR, que impliquen més de 20 gens. OxyR està implicat en la regulació de catalases i reductases, que actuen en presència de peròxid d'hidrogen. Mitjançant el reguló SoxR, s'activa la via que

codifica la superòxid dismutasa, entre altres enzims que permeten fer front al dany oxidatiu (Demple, 1997; Michán *et al.*, 1999; Zheng *et al.*, 1999).

3. Sistema de reparació d'emergència del DNA o sistema SOS

En ser aquest sistema de reparació l'objecte del present estudi, a continuació es farà una descripció més detallada.

1.1 El sistema SOS a *Escherichia coli*

El sistema SOS és una xarxa que a *E. coli* podria implicar aproximadament 40 gens (Taula 1.1) i que s'activa davant la presència de lesions al DNA o una aturada de la replicació (Fernández d'Henestrosa *et al.*, 2000; Courcelle *et al.*, 2001). Aquest nivell de lesions és detectat pel sistema com a presència de DNA de cadena senzilla (ssDNA), generalment deguda a la inhibició de la replicació i al processament d'extrems derivats del trencament de DNA. El sistema, doncs, s'activa gràcies a la capacitat per detectar aquest senyal, capacitat que ve donada per la proteïna RecA, que podríem considerar l'element activador del sistema.

L'activació de RecA un cop reconeguda la presència de ssDNA porta a la desrepressió gradual dels gens de la xarxa, que en situació basal estan reprimits per l'altre element regulador, la proteïna LexA o repressor del sistema, que s'uneix a una seqüència palindròmica dels promotors dels gens SOS sovint anomenada caixa SOS o caixa d'unió de LexA. Els propis gens reguladors (*recA* i *lexA*) formen part del sistema SOS a *E. coli*.

La repressió dels gens SOS, no és total i permet una expressió basal suficient per tenir una concentració de LexA que pugui reprimir tots els gens implicats i una concentració basal de RecA que possibiliti les seves funcions en recombinació. De totes maneres, hi ha gens del sistema SOS que tenen efectes tòxics sobre la cèl·lula i que són reprimits de manera més forta. Els gens implicats en la resposta d'emergència, també anomenats *din* ("damage-inducible"), estan implicats en diferents funcions cel·lulars (reparació del DNA, divisió cel·lular, tolerància al dany i recombinació, entre d'altres).

Taula 1.1. Gens que formen part del reguló LexA a *E. coli*

Gens	Funció
<u>Cromosoma</u>	
<i>dinA</i> (<i>polB</i>)	DNA polimerasa II
<i>dinB</i> (<i>dinP</i>)	DNA polimerasa IV
<i>dinD</i> (<i>pscA</i>)	Desconeguda. Mutant de sensibilitat al fred
<i>dinF</i>	Desconeguda. Forma unitat transcripcional amb <i>lexA</i>
<i>dinG</i>	Desconeguda
<i>dinI</i>	Inhibició d' <i>umuD</i> , possible modulador del sistema
<i>dinL</i> (<i>sosC</i> , <i>yjiW</i>)	Desconeguda
<i>dinM</i> (<i>sosD</i> , <i>ydjQ</i>)	Possible homòleg d' <i>uvrC</i>
<i>dinO</i> (<i>sosF</i> , <i>molR</i>)	Possible regulador de la captació de molibdat
<i>dinQ</i>	Desconeguda
<i>dinS</i>	Possible transposasa
<i>ftsK</i> (<i>dinH</i>)	Desconeguda
<i>hokE</i> (<i>ybdY</i>)	Possible proteïna "assassina"
<i>lexA</i>	Repressor del sistema SOS o reguló LexA
<i>recA</i>	Recombinació i coproteasa
<i>recN</i>	Recombinació per la via RecF, reparació de trencaments de doble cadena i cadena senzilla
<i>recX</i>	Possible modulador del sistema, inhibició funcions RecA
<i>ruvAB</i>	Resolvasa. Recombinació
<i>sbmC</i>	Resistència a la Microcina B17
<i>Ssb</i>	Unió a ssDNA
<i>sulA</i> (<i>sfiA</i>)	Inhibidor de la divisió cel·lular
<i>umuDC</i>	Mutagènesis SOS. DNA polimerasa V
<i>uvrA</i> (<i>dinE</i>)	Reconeixement de dímers de pirimidina i reparació per escissió de nucleòtid
<i>uvrB</i>	Reconeixement de dímers de pirimidina i reparació per escissió de nucleòtid
<i>uvrD</i>	Helicasa II. Reparació per escissió de nucleòtid i reparació d'aparellaments incorrectes
<i>ybfE</i>	Desconeguda
<i>ydjM</i>	Desconeguda
<i>yebG</i>	Desconeguda
<i>ysdAB</i>	Desconeguda
<u>Extracromosòmics</u>	
<i>colA</i> (<i>caa</i>)	Colicina (ColA)
<i>colE</i> (<i>cea</i>)	Colicina (Col E1)
<i>impAB</i>	Homòleg d' <i>umuDC</i>
<i>mucAB</i>	Homòleg d' <i>umuDC</i>
<i>samAB</i>	Homòleg d' <i>umuDC</i>
Repressors fàgics*	Reguladors del cicle lític de profags

Modificat de Walker (1996), Koch i Woodgate (1998), Fernández de Henestrosa *et al.* (2000) i Courcelle *et al.* (2001).

*Regulats per un mecanisme anàleg al del repressor LexA amb un domini C-terminal molt similar.

1.1.1 Inducció i regulació del sistema SOS

L'aparició de regions de DNA de cadena senzilla (ssDNA) es produeix quan la cèl·lula intenta fer progressar la forca de replicació en el punt on s'ha produït una lesió i aquesta impedeix l'avanç de la DNA polimerasa III (Sassanfar i Roberts, 1990). La generació de regions ssDNA s'incrementa per l'activitat del complex RecBCD (holoenzim exonucleasa V) amb funció helicasa sobre el dsDNA (Churchill *et al.*, 1999). RecA reconeix aquesta senyal i adopta una conformació activa formant filaments helicoïdals al voltant del DNA en presència de trifosfat d'adenosina (ATP), filaments que consten de 6 monòmers de RecA per volta en una estructura similar a la de RuvB (Egelman, 1998). En el procés de nucleació intervindrien altres proteïnes com Ssb, que facilita l'accés al ssDNA, i el complex RecFOR, que millora l'eficiència del procés (Kuzminov *et al.*, 1999).

El complex nucleoprotèic de RecA (RecA*) actua com una coproteasa. S'uneix a LexA gràcies a un domini d'interacció de RecA on sembla que estarien implicats els residus 156-165 i 229-243 de RecA (Yu i Egelman, 1993; Karlin i Brocchieri, 1996; Mirshad i Kowalczykowski, 2003). Aquesta unió ajuda a la proteïna LexA a realitzar una autohidròlisi a l'enllaç dels residus Ala₈₄ i Gly₈₅ que inactiva el repressor (Little, 1991). Els fragments de LexA són llavors ràpidament destruïts a l'aparèixer senyals que detecta la proteasa ClpXP i que no eren accessibles quan la proteïna estava sencera (Neher *et al.*, 2003). D'aquesta manera, els promotors que estaven reprimits per la unió de LexA passen a incrementar la seva transcripció, que es mantindrà mentre no desaparegui la forma activada de RecA* i els monòmers de LexA deixin de ser hidrolitzats. Això succeeix en desaparèixer les lesions i recuperar-se la normal replicació del DNA. Com que ja no hi ha DNA de cadena senzilla, RecA no passa a la forma activa i la proteïna LexA que es sintetitza *de novo* torna a unir-se als promotors dels gens SOS, recuperant l'estat basal del sistema (Friedberg *et al.*, 1995; Walker *et al.*, 1996; Koch i Woodgate, 1998) (Fig. 1.1).

Recentment s'ha incidit en el paper de dos moduladors de l'activitat de RecA: les proteïnes DinI i RecX, codificades ambdues per gens del reguló LexA. DinI exerceix un paper dual: al principi de la resposta d'emergència es troba en baixes concentracions i estabilitza la unió de

RecA al ssDNA que dóna lloc a RecA*. Quan la resposta arriba a la seva fase final, la concentració elevada de DinI juga el paper contrari, desestabilitzant el nucleofilament RecA* (Lusetti *et al.*, 2004). RecX a elevades concentracions inhibeix la funció coproteasa de RecA* (Stohl *et al.*, 2003) i elimina la senyal inductora, la elevada concentració de RecX a la fase final de la resposta SOS ajudaria a limitar els nivells d'inducció i a recuperar l'estat basal (Stohl *et al.*, 2003). S'ha trobat també un gen cromosòmic, anomenat *isfA*, que en ser inactivat provoca una major sensibilitat a la radiació UV i una inhibició de funcions SOS (Felczak *et al.*, 1999). El producte d'aquest gen podria, per tant, estar actuant com un activador addicional del sistema.

1.1.1.1 La proteïna LexA

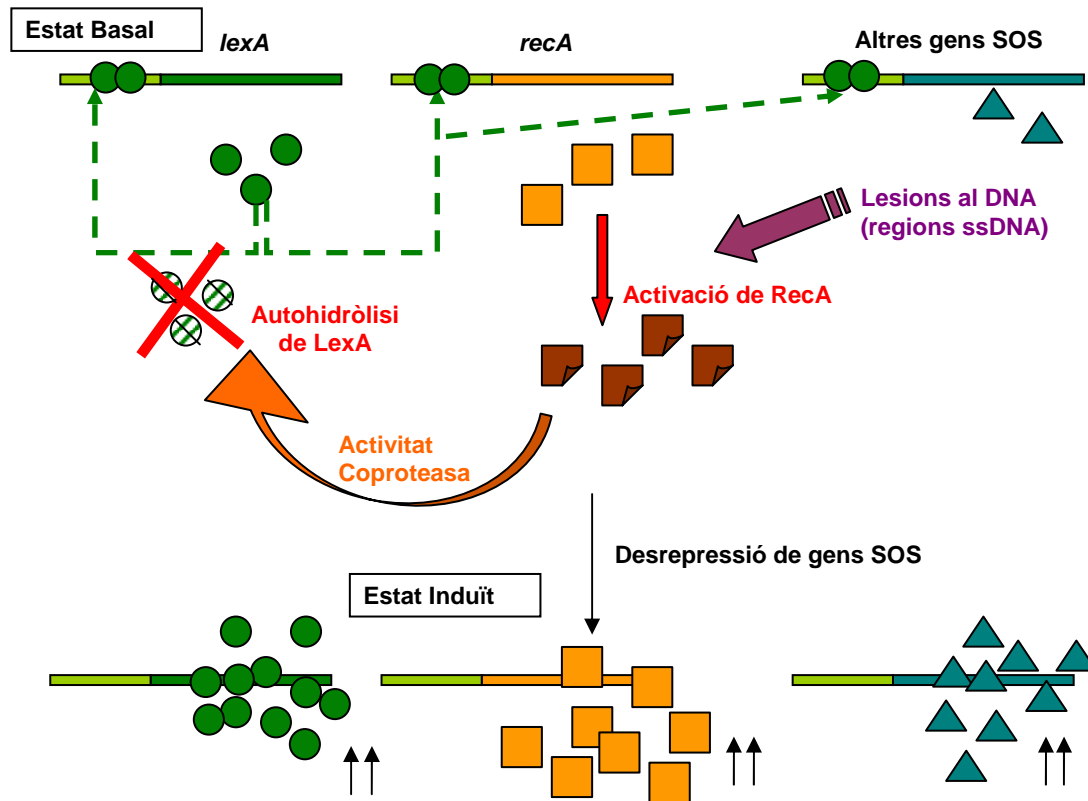
La proteïna LexA d'*E. coli* és un pèptid de 202 aminoàcids de longitud i un pes molecular de 22,7 KDa (Horii *et al.*, 1981). Està codificada pel gen *lexA*, que a *E. coli* es troba formant unitat transcripcional amb *dinF*. Aquesta proteïna és el repressor transcripcional del sistema SOS i, per tant, un dels principals reguladors de la resposta d'emergència.

Aquesta acció es duu a terme gràcies a la capacitat d'unió al DNA de la proteïna LexA, que reconeix una seqüència palindròmica, denominada caixa SOS o caixa LexA, present als promotors dels gens del sistema. Aquesta caixa està ampliamment caracteritzada a *E. coli* i la seva seqüència consens és CTGT-(AT)₄-ACAG (Walker *et al.*, 1984; Wertman i Mount, 1985). La unió de LexA es produeix en forma de dímer, reconeixent cadascun dels monòmers de LexA un dels motius inversament simètrics del palíndrom. Una altra capacitat essencial de LexA és la d'autohidrolitzar-se. En condicions fisiològiques, això succeeix gràcies a l'acció catalitzadora de RecA*.

Generalment, cadascun dels gens del sistema SOS té una única caixa SOS al seu promotor, però no és tampoc extrany trobar gens amb dues o més regions operadores (*lexA* i *recN*). Cadascuna de les bases de la seqüència consens de la caixa té una importància diferent en l'establiment específic de la unió, essent les més conservades CTGT i ACAG (Taula 1.2) que per altra banda, s'ha vist que són les que tenen les interaccions principals amb els residus de la

proteïna encarregats del seu reconeixement (el "DNA binding domain" o DBD) (Knegtel *et al.*, 1995; Thliveris *et al.*, 1992).

A) Inducció del sistema



B) Retorn al nivell basal un cop reparades les lesions i desapareguda RecA*

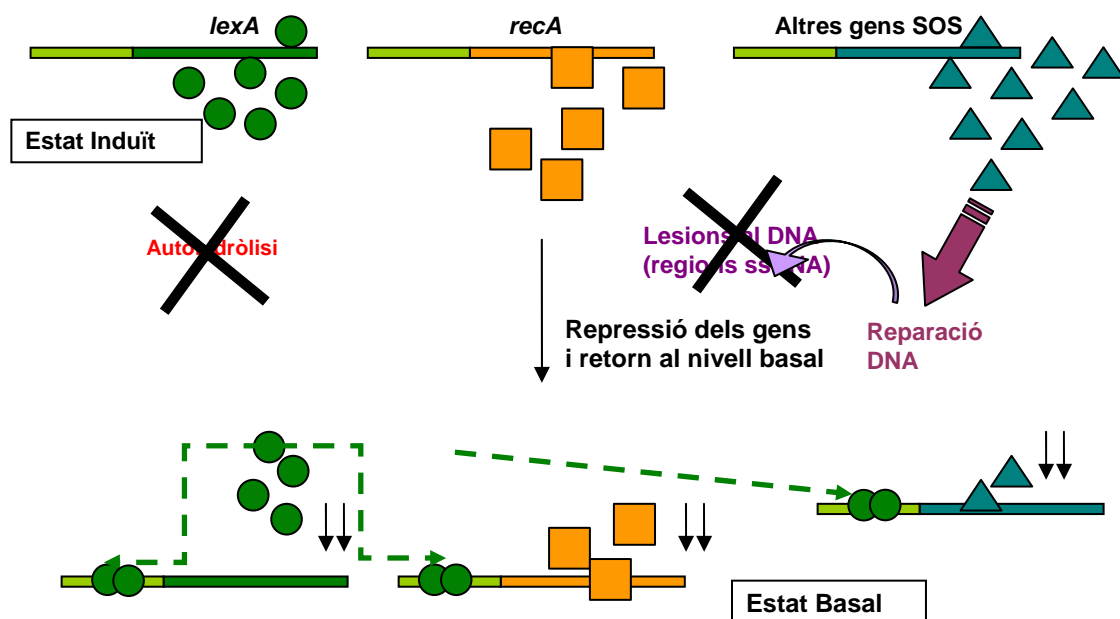


Fig. 1.1. Model esquemàtic del procés d'activació del sistema SOS.

El repressor s'uneix a la regió promotora dels gens SOS i, per tant, podríem dir que actua inhibint l'inici de la transcripció a l'evitar la unió de la RNA polimerasa (RNAP) a la seva caixa -10 (Brent i Ptashne, 1981; Weisemann i Weinstock, 1991). El grau de repressió que exerceix la proteïna és variable en funció de l'especificitat del reconeixement de la caixa en els diferents gens, que vindrà determinada per la proximitat d'aquesta a la seqüència consens (Schnarr *et al.*, 1991) i per la localització de la caixa respecte les regions d'unió de la RNAP. De la mateixa manera, és obvi que la presència de regions operadores addicionals incrementarà l'efecte de la repressió.

Taula 1.2. Gens cromosòmics d'*E. coli* regulats per LexA i seqüències de les seves respectives caixes SOS.

Gens	Seqüència
<i>dinA</i> (<i>polB</i>)	GACTGTATAAAAACCACAGCC
<i>dinB</i> (<i>dinP</i>)	CACTGTATACTTTACCAGTG
<i>dinD</i> (<i>pscA</i>)	AACTGTATATAAATACAGTT
<i>dinG</i>	TAT TGG CTGTTTATACAGTA
<i>dinI</i>	ACCTGTATAAATAACCAGTA
<i>dinL</i> (<i>sosC</i> , <i>yjiW</i>)	TACTG A TGATATATACAGGT
<i>dinM</i> (<i>sosD</i> , <i>ydjQ</i>)	CACTG G ATAGATAAACAGCA
<i>dinO</i> (<i>sosF</i> , <i>molR</i>)	A A CTG G ATAAAATTACAGGG
<i>dinQ</i>	TACTGTATGATTAT C CAGTT
<i>dinS</i>	AGCTGTATTTGTCT C CAGCA
<i>ftsK</i> (<i>dinH</i>)	TCCTGTTAATCCATACAGCA
<i>hokE</i> (<i>ybdY</i>)	CACTGTATAAATAAACAGCT
<i>lexA/dinf</i> (1)	TGCTGTATATACTCACAGCA
<i>lexA/dinf</i> (2)	TGCTGTATATACACCAGGG
<i>recA</i>	TACTGTATGCTCATAACAGTA
<i>recN</i> (1)	TACTGTATATAAAAACCAGTT
<i>recN</i> (2)	TACTGTACACAATAACAGTA
<i>recN</i> (3)	T A A TGG TTTTTTCATACAGGA
<i>ruvAB</i>	CGCTG G ATGTCTAT C CAGCA
<i>sbmC</i>	TACTGTATATAAAAAACAGTA
<i>Ssb</i>	ACCTG A ATGAATATACAGTA
<i>sulA</i> (<i>sfiA</i>)	TACTGTACATCCATACAGTA
<i>umuDC</i>	TACTGTATATAAAAAACAGTA
<i>uvrA</i> (<i>dinE</i>)	TACTGTATATTCAT C CAGGT
<i>uvrB</i>	A A CTGTTTTTTTAT C CAGTA
<i>uvrD</i>	ATCTGTATATATACCAGCT
<i>ybfE</i>	A A CTG A TTAAAAACCAGCG
<i>ydjM</i> (1)	TACTGTACGTATCGACAGTT
<i>ydjM</i> (2)	CACTGTATAAAAAT C C T A T A
<i>yebG</i>	TACTGTATAAAATCACAGTT
<i>ysdAB</i>	CACTGTTTATTTATACAGTA
Consens	TACTGTATATATATACAGTA

Modificat de Fernández de Henestrosa *et al.* (2000).

Les bases ombrejades assenyalen els motius més conservats.

Les bases en negreta assenyalen desviacions dels motius conservats CTGT i ACAG.

i. Estructura de la proteïna LexA

Podem diferenciar tres regions estructurals i funcionals a la proteïna LexA (Fig. 1.2):

1. Regió Amino-terminal. Comprèn els aminoàcids que van des de la posició 1 a la 72. És el domini encarregat del reconeixement de la caixa SOS i la unió al DNA. La seva estructura és similar a la d'una hèlix-gir-hèlix (HTH), que sovint es troba en altres repressors transcripcionals (Harrison i Aggarwal, 1990; Schnarr *et al.*, 1991). En concret, l'estructura tridimensional d'aquest domini de LexA ha estat caracteritzada per ressonància magnètica nuclear (RMN) (Fogh *et al.*, 1994) i està constituïda per tres hèlixs α (residus 6-21, 28-35 i 40-52) connectades per regions "turn" i amb la presència de dues regions β antiparal·leles (entre els residus 50-58 i 66-68). Les tres regions hèlix α sovint s'han anomenat $\alpha 1$, $\alpha 2$ i $\alpha 3$. Les dues últimes serien, per la seva disposició, les més pròximes a una estructura "HTH-like". A la tercera hèlix trobaríem els residus implicats de manera principal en el reconeixement de la caixa (Fig. 1.4 A). Aquesta regió serà variable en altres espècies en funció de la caixa SOS que reconegui LexA.
 2. Regió connectora ("hinge region"). Fragment central de la proteïna (residus 73-94) on es troba l'enllaç Ala₈₄-Gly₈₅ susceptible de ser trencat a la reacció d'autohidròlisi. Aquesta regió, però, també tindria una certa estructura de plegament, necessària per a una interacció amb el domini carboxi-terminal que permeti la realització de l'autohidròlisi.
 3. Regió carboxi-terminal. Comprèn els aminoàcids 95-202 de la proteïna, implicats a la reacció d'autohidròlisi, la dimerització i la interacció amb RecA*. És la regió més conservada de LexA, amb una similitud elevada amb la proteïna UmuD i amb el repressor del cicle lític de diferents bacteriòfags (Battista *et al.*, 1990), que tot i això realitzen funcions d'autohidròlisi sensiblement diferents (Mustard i Little, 2000). La seva estructura secundària és majoritàriament de fulles β (Hurstel *et al.*, 1986).
-

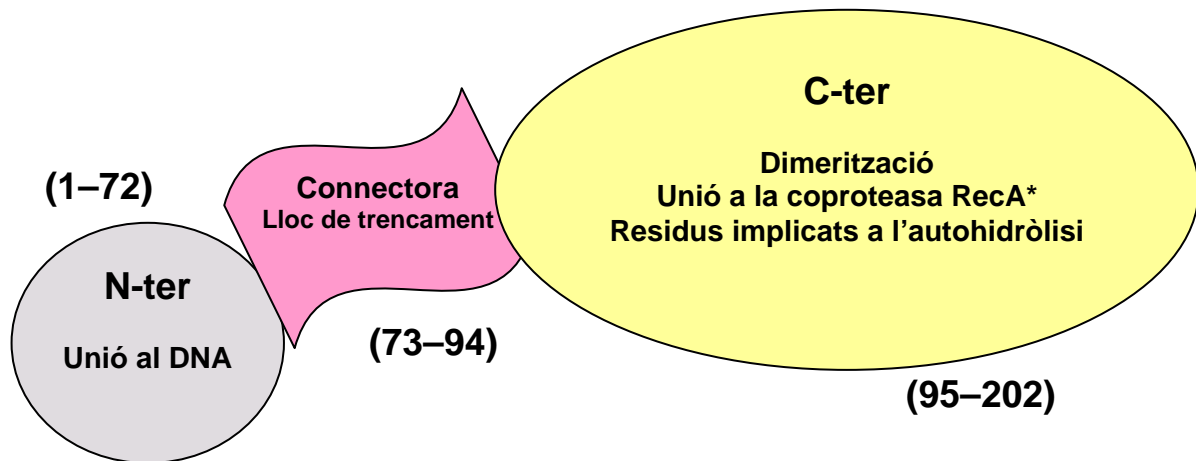


Fig. 1.2. Esquema dels dominis estructurals de la proteïna LexA. Entre parèntesi les posicions dels aminoàcids que pertanyen al domini a la proteïna LexA d'*E. coli*.

ii. Funció de la proteïna LexA

L'estructura de LexA li permet realitzar la seva funció reguladora gràcies a les característiques del seu domini amino-terminal (reconeixement d'una seqüència reguladora) i del seu domini carboxi-terminal (dimerització i sensibilitat a la senyal inductora de RecA*).

La unió de LexA al DNA és un procés dinàmic i reversible que es produeix en forma de dímer i on cadascun dels monòmers de LexA reconeix un dels motius de la caixa. Experimentalment s'ha demostrat que aquest procés es produeix de manera seqüencial. Primer, s'uneix un monòmer al motiu CTGT i, posteriorment, un segon monòmer al motiu CTGT de l'altra cadena, gràcies a la interacció proteïna-proteïna amb la regió carboxi-terminal del primer (Kim i Little, 1992; Mohana-Borges *et al.*, 2000) (Fig. 1.3).

Com hem detallat abans, la regió amino-terminal conté tres hèlixs α , la tercera de les quals comprèn els residus 40-52. Fa temps que es postula que els residus d'aquesta hèlix són els directament implicats en el reconeixement de la seqüència CTGT-N₈-ACAG, ja que mutants amb delecions o mutacions en aquesta hèlix veuen alterada la seva capacitat d'unió (Oertel Buchheit *et al.*, 1990; Oertel Buchheit *et al.*, 1992). Un estudi més acurat (Thliveris *et al.*, 1992) va senyalar els residus 42-47 com els essencials per a la unió directa al DNA. Les mutacions a les posicions A₄₂, A₄₃, E₄₄, E₄₅ i R₂₈ de LexA donen lloc a mutants negatius dominants de la

proteïna (sense capacitat d'unió al DNA però amb capacitat de dimeritzar amb monòmers salvatges de LexA).

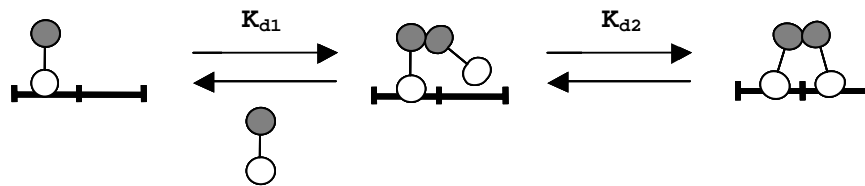


Fig. 1.3. Model de l'equilibri químic a la reacció d'unió i dimerització de LexA a l'operador SOS. El domini amino terminal és representat per un cercle blanc i el domini carboxi terminal per un de gris. Modificat de Kim i Little (1992).

En alguns casos, es van trobar mutants ($E_{45}K$, $E_{45}V$, $N_{41}H$ i $E_{44}Q$ o els dobles mutants $N_{41}S A_{42}S$) que presentaven un canvi d'especificitat en reconèixer caixes SOS diferents de la seqüència consens en alguna de les posicions. Estudis de RMN (Knegtel *et al.*, 1995) han aportat informació sobre la naturalesa bioquímica del reconeixement de la caixa SOS per part de LexA mitjançant l'establiment de ponts d'hidrògen entre els residus N_{41} , E_{44} , E_{45} i les bases CTGT de la caixa i interaccions hidrofòbiques entre la primera timina de la caixa i els residus S_{39} , A_{42} i N_{41} . Però més enllà d'aquests residus, hi hauria un ampli ventall d'aminoàcids que contactarien amb l'esquelet fosfat del DNA i una gamma d'interaccions i equilibris hidrofòbics i polars que garantirien l'estabilitat de la unió (Fig. 1.4).

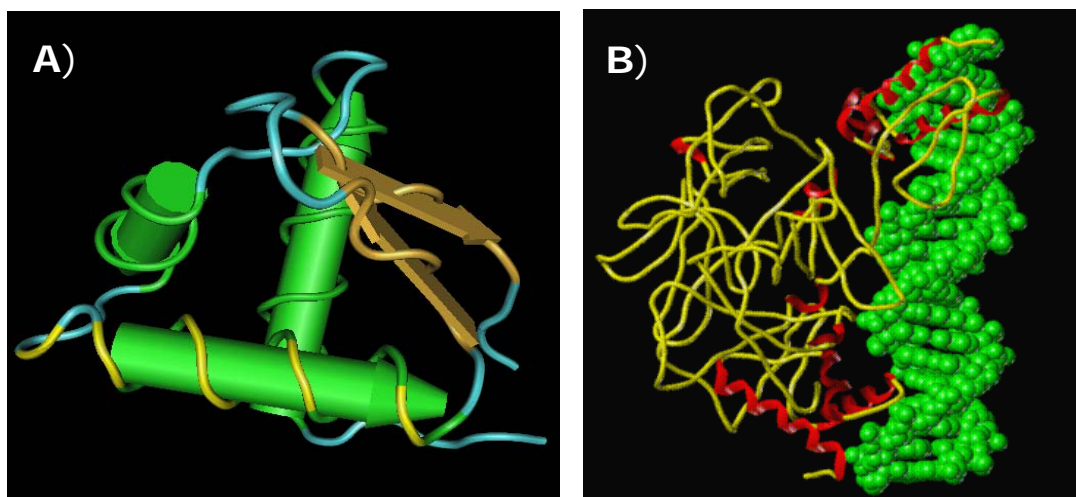


Fig. 1.4. A) Model estructural del domini amino terminal de la proteïna LexA d'*E. coli* (generat amb el programa Cn3D 4.1). En primer terme hèlix $\alpha 3$ amb els principals residus que reconeixen la caixa LexA (en groc) B) Model de la interacció de LexA amb la caixa SOS (http://www.rtc.riken.go.jp/jouhou/image/dna-protein/all/small_N1qaa.gif).

iii. Autohidròlisi de la proteïna LexA

Al trencament de la proteïna LexA, tant el substrat de la reacció (lloc de trencament) com el centre catalític (residus que intervenen en el procés químic que porta al trencament de l'enllaç peptídic) es troben dins la molècula de LexA i per això es considera un procés d'autohidròlisi (Fig. 1.5). En condicions de pH neutre, la reacció no es dona de manera espontània sinó que és necessària la catàlisi de RecA*, però en condicions de pH bàsic s'ha aconseguit *in vitro* un trencament espontani de LexA sense la intervenció de cap altre element (Little, 1984; Little *et al.*, 1994).

El lloc de trencament, com ja ha estat esmentat, es troba a la regió connectora de la proteïna, concretament a l'enllaç peptídic de l'Ala₈₄ i la Gly₈₅. El centre catalític estaria format pels aminoàcids Ser₁₁₉ i Lys₁₅₆ (Slilaty i Little, 1987; Roland *et al.*, 1992). La distància dels residus Ala₈₄-Gly₈₅ i els residus Ser₁₁₉ i Lys₁₅₆ del centre catalític es manté constant en les proteïnes LexA caracteritzades fins ara amb petites variacions d'1 o 2 aminoàcids i serveix per identificar aquesta proteïna com a tal.

El procés químic que porta al trencament de LexA consisteix en una desprotonització del grup amino de la Lys₁₅₆ (amb un pK de 9,5) deguda a una pujada del pH o, en condicions *in vivo*, a l'efecte catalític de RecA*. La sostracció d'un protó al grup hidroxil de l'aminoàcid Ser₁₁₉ per part de la Lys₁₅₆ desencadenaria un atac nucleofilic sobre l'enllaç Ala₈₄-Gly₈₅ que provocaria el trencament del mateix i, per tant, de la proteïna (Fig. 1.5).

Aquest mecanisme d'autohidròlisi havia estat descrit anteriorment per les serin-proteases, essent el domini carboxi-terminal de LexA similar al que podríem trobar en aquest grup de proteases (Roland i Little, 1990). Hi ha dades, però, que apunten a un mecanisme de trencament més similar al de les beta-lactamases (Little, 1993). LexA tindria dues conformacions, una compatible amb l'autohidròlisi i una altra amb els centres catalítics allunyats del punt de trencament. La interacció entre el lloc catalític i l'enllaç Ala₈₄-Gly₈₅ podria donar-se per un canvi d'una conformació a l'altra gràcies a la flexibilitat de la zona connectora (Oertel-Buchheit

et al., 1998). RecA* activaria l'autohidròlisi establint la conformació compatible amb aquesta reacció (Luo *et al.*, 2001).

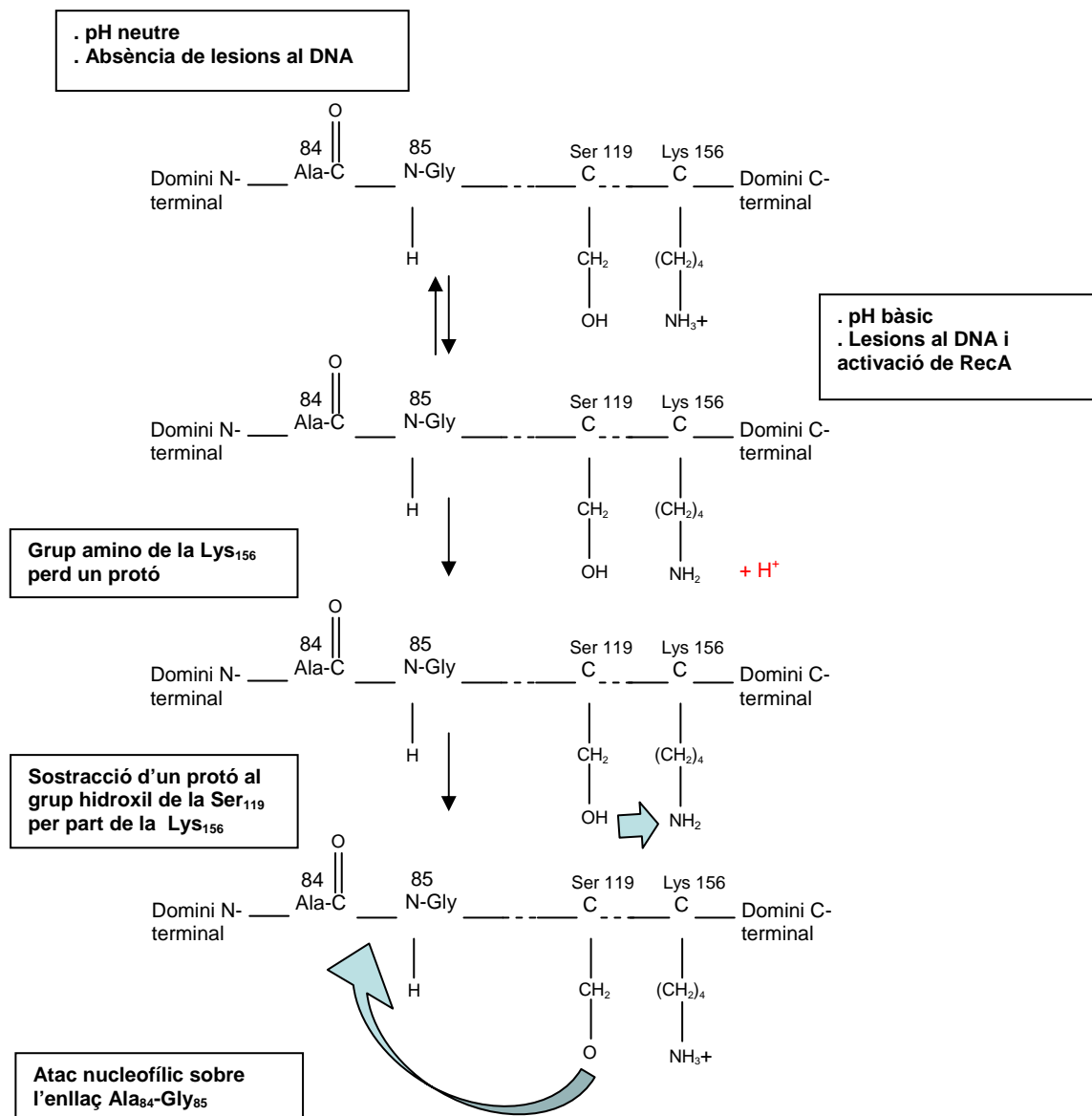


Fig. 1.5. Representació esquemàtica del procés d'autohidròlisi de la proteïna LexA. L'estructura dels residus de LexA no està a escala, es mostren tan sols aquells elements químics implicats directament a la reacció.

1.1.2 Funcions SOS

El reguló SOS duu a terme múltiples funcions, majoritàriament relacionades amb els processos de reparació de les lesions del DNA per garantir la supervivència de la cèl·lula. Aquestes

funcions poden ser englobades en tres grans grups: reparació, mutagènesi SOS i altres funcions no directament relacionades amb la reparació.

1.1.2.1 Reparació per escissió de nucleòtid

Aquest mecanisme de reparació detecta un mal aparellament de nucleòtids i, utilitzant com a motlle la cadena de DNA original, és capaç d'eliminar el nucleòtid danyat i substituir-lo pel corresponent. La radiació UV i alguns tipus de productes químics són els agents que causen el tipus de lesions que poden ser afrontades per aquest sistema (Friedberg *et al.*, 1995).

Els elements que participen a la reparació per escissió de nucleòtid són els productes dels gens *uvr* (*uvrA*, *uvrB*, *uvrC* i *uvrD*) que, tanmateix, no estan formant unitat transcripcional en un únic operó ni tenen el mateix tipus de control. El gen *uvrC*, per exemple, no forma part del reguló LexA a diferència dels altres tres. També participen en el procés de reparació la DNA polimerasa I i la DNA lligasa.

UvrA és la proteïna que reconeix el dany de manera preliminar com a part del complex UvrA₂B. Quan és reconeguda una lesió o una estructura anormal del DNA, la proteïna UvrB verifica la presència d'una lesió gràcies a la seva estructura de “ β -hairpin” (Moolenaar *et al.*, 2001) i actua com a helicasa, amb el consum d'ATP, unint-se de manera estable al lloc danyat i formant el complex de preincissió (Moolenaar *et al.*, 1994). En aquest procés s'allibera UvrA. A continuació, la proteïna UvrC s'uneix al complex de preincissió i realitza els talls als enllaços fosfodièster als dos costats de la lesió, en direcció 5' i 3' (Verhoeven *et al.*, 2001). S'ha proposat la participació en el reconeixement del dany d'un segon monòmer d'UvrB que asseguraria la detecció de la lesió a les dues cadenes del dsDNA i que s'alliberaria amb l'entrada d'UvrC (Verhoeven *et al.*, 2002). Recentment s'ha descrit també la presència d'un homòleg d'UvrC, la proteïna Cho, que formaria amb UvrB un complex actiu que participa a la reparació d'alguns tipus de lesions conjuntament amb UvrC. A diferència d'UvrC, Cho, que està codificada pel gen *ydjQ*, forma part del reguló LexA (Moolenaar *et al.*, 2002). UvrD és també una proteïna amb funció helicasa (Helicasa II), i participaria juntament amb la DNA

polimerasa I a la síntesi de la nova cadena que ocuparà el lloc de la deletada però restablint el nucleòtid correcte (Moolenaar *et al.*, 2000). Per realitzar aquesta acció és necessària la presència de (desoxi)nucleòsid trifosfats (dNTPs) i l'acció última de la DNA lligasa.

1.1.2.2 Reparació per recombinació

Quan les lesions que s'han produït impliquen, per exemple, enllaços entre les dues cadenes del DNA, trencaments o s'ha danyat la cadena motlle, no es pot dur a terme una reparació per escissió. El tipus de lesions més habituals en aquest grup són els trencaments de doble cadena i els derivats del bloqueig de replicació. Aquest últim cas el trobem quan la DNA polimerasa està intentant replicar i es troba davant de lesions que no pot superar. Hi haurà majoritàriament un salt en aquest punt, deixant un espai d'un miler de nucleòtids sense replicar. Per tant, la cadena presentarà discontinuïtats replicatives en forma de forats a les cadenes filles.

Quan es parla del mecanisme de reparació per recombinació no es fa referència en sentit estricte a un mecanisme que repari les lesions, sinó a un mecanisme cel·lular encaminat a generar un substracte susceptible de ser reparat posteriorment per altres mecanismes. La cèl·lula aconsegueix que un dany present en les dues cadenes passi a estar representat tan sols en una cadena i pugui ser, per tant, solucionat per altres mecanismes (Kuzminov *et al.*, 1999; Dronkert *et al.*, 2001; Dudás *et al.*, 2004) (Fig. 1.6).

Hi ha dos grans vies de reparació per recombinació. La via *recBCD* és la principal i estaria implicada en la reparació de la majoria dels trencaments de doble cadena mitjançant la formació d'un holoenzim amb funció mixta exonucleasa-helicasa (Anderson i Kowalczykowski, 1998). La via *recF* s'activa quan *recBCD* no és funcional (Ivancic-Bace *et al.*, 2002) i també estaria implicada en els casos de bloqueig de replicació. Aquesta segona via seria la via controlada per LexA i per tant formaria part del seu reguló LexA. De tots els gens de la via *recF* (*recF*, *recJ*, *recO*, *recR*, *recQ* i *recN*) tan sols *recN* formaria part del sistema SOS a *E. coli* (Kuzminov *et al.*, 1999). En ambdues vies seria imprescindible el paper de les proteïnes RecA i Ssb (Roca i Cox, 1990; Cox *et al.*, 2000).

Hi ha altres vies molt menys importants que participen en casos especials. RecE, per exemple, actua en relació a la presència d'un fag críptic i mutacions en el gen *sbcA*. És important recalcar que en les fases de migració de cadena i postsinàptica hi ha implicades altres proteïnes, com ara RecG amb funció helicasa i els productes dels gens *ruvABC*. Entre tots aquests gens, tan sols *ruvA* i *ruvB* estan regulats directament per LexA (Kogoma *et al.*, 1997 i Kuzminov *et al.*, 1999).

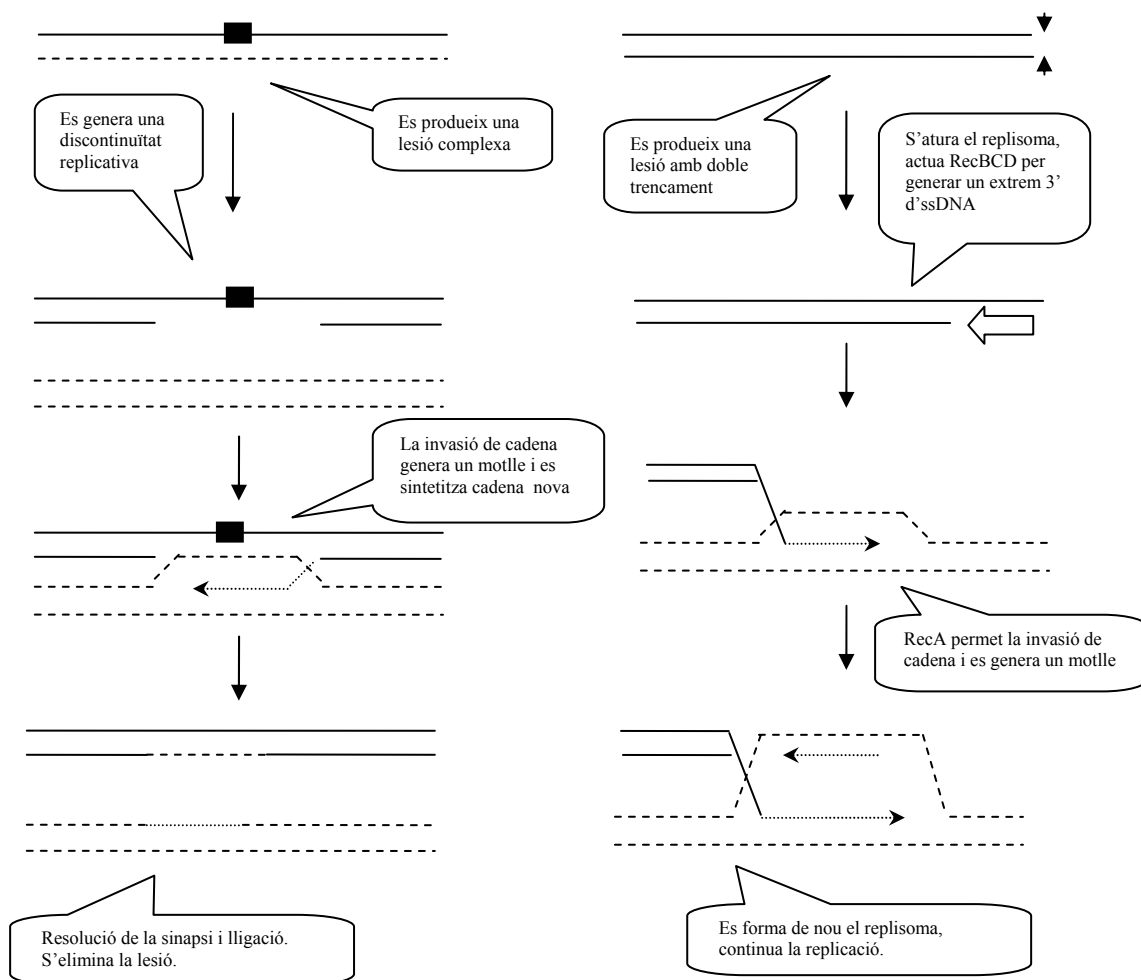


Fig. 1.6. Model esquemàtic del mecanisme de reparació per recombinació davant una discontinuïtat replicativa i un doble trencament de cadena.

1.1.2.3 Mutagènesi SOS

La mutagènesi SOS s'ha anomenat també síntesi de translesió o reparació tendent a l'error. És un mecanisme de "by pass" replicatiu que utilitza la cèl·lula quan el nivell de lesions és molt elevat per poder continuar replicant el cromosoma. És un fenomen estretament regulat i que tan

sols s'indueix quan el nivell d'inducció del sistema SOS és molt elevat. El mecanisme té dues lògiques superposades. Per una banda li permet a la cèl·lula continuar amb la replicació i seguir endavant amb el cicle cel·lular esperant poder solucionar aquestes lesions en etapes posteriors. D'altra banda, també és un mecanisme que aconsegueix incrementar la taxa de mutació en el genoma, afavorint l'aparició de mutacions que podrien ser necessàries per escapar de la pressió selectiva a la qual està sotmesa la cèl·lula i per la qual està rebent aquesta dosis tan elevada de dany al DNA.

Els dos elements centrals en aquest tipus de resposta són els productes dels gens *umuDC*, que a *E. coli* es troben formant una única unitat transcripcional sota el control de LexA. La proteïna UmuD presenta unes característiques similars a les del domini carboxi terminal de LexA, amb un punt de trencament format per un enllaç Cys₂₄-Gly₂₅ (Sutton *et al.*, 2000; Sutton *et al.*, 2001). La presència de dany al DNA induïx la transcripció d'*umuD* i *umuC*, que fins llavors s'expressaven en molt baixes quantitats. El complex nucleoprotèic RecA*, al igual que fa amb LexA, és capaç de catalitzar l'autohidròlisi d'UmuD, que passa a una conformació activa UmuD' (Nohmi *et al.* 1988). Dues subunitats de UmuD' formen un complex amb UmuC. El complex Umu(D')₂UmuC, conegut com DNA polimerasa V (Tang *et al.*, 1999), s'uneix llavors al DNA amb l'ajuda de RecA* (Ekaterina *et al.*, 1993) i és el responsable de la síntesi de translesió (González i Woodgate, 2002).

Novament, la proteïna Ssb jugaria un paper important en la intervenció de RecA* i, de fet, s'ha apuntat la possibilitat que RecA i Ssb siguin claus en la realització del "by pass" sobre la lesió (Reuven *et al.*, 2001). El control del procés de la síntesi de translesió es realitza a múltiples nivells, donada la seva perillositat per a la integritat de la informació genètica. Per una banda, l'operador d'*umuDC* és dels que presenta major afinitat per LexA. Les proteïnes són làbils i la modificació d'UmuD a UmuD' és catalitzada per RecA* amb menor eficiència que l'autohidròlisi de LexA. Per altra banda, la formació d'homodímers UmuD'₂ està desfavorida estequiòmicament davant la formació d'heterodímers UmuD-UmuD' (Battista *et al.*, 1990). Un altre element modulador de la mutagènesi SOS seria el producte del gen *dinI*. DinI limitaria

la síntesi de translesió gràcies a la seva afinitat per RecA*, que és major que l'afinitat de RecA* per UmuD. Així la unió de DinI a RecA* limitaria la possibilitat d'activació d'UmuD a un cas de màxima inducció SOS (Yasuda *et al.*, 2001).

1.1.2.4. Altres funcions del sistema SOS

Dins del grup de funcions no reparatives dels gens SOS trobem la inhibició de la divisió cel·lular, per l'acció del producte del gen *sulA* (Huisman *et al.*, 1981). Aquest és un inhibidor de la septació que actua sobre FtsZ i la seva inducció depenent de LexA produeix un fenotip conegut com filamentació (Bi *et al.*, 1993). Aquesta filamentació és el resultat de la manca de divisió cel·lular, que es retarda d'aquesta manera fins que les lesions que s'estan reparant puguin ser reparades. SulA però, és una proteïna molt inestable i per tant s'observarà tan sols filamentació en casos d'inducció molt elevada (Walker, 1996). Dins els efectes de la inducció del sistema SOS també es troba la inducció de profags (Sauer *et al.*, 1982), gràcies a la analogia entre el funcionament dels seus repressors de cicle lític i el repressor LexA (Mustard i Little, 2000). Tots dos tipus de repressors són autohidrolitzats davant de la presència de dany al DNA i l'aparició de RecA* amb funció coproteasa.

1.2. El sistema SOS a altres bacteris

Fins aquí s'han exposat les característiques bàsiques del sistema SOS al microorganisme millor estudiat, *Escherichia coli*, amb una abundant bibliografia al respecte. Hom podria creure que el sistema SOS és universal i no presenta canvis remarcables entre diferents espècies i grups filogenètics, especialment si tenim en compte la suposada centralitat de les seves funcions en el manteniment de la informació genètica i la supervivència de l'espècie. És cert que el sistema SOS és un sistema extès en el món bacterià trobant-se homòlegs de LexA a la majoria de microorganismes coneguts (Eisen i Hanawalt, 1999). Així mateix, al domini *Archaea* s'ha suggerit la presència d'un sistema similar per estudis de genòmica comparativa (Makarova *et al.*, 2002).

De totes maneres, a diferència de la proteïna RecA, que es troba de manera ubiqua en tot el domini *Bacteria* amb un origen ancestral i una funció vital que en garanteix la seva conservació a través de l'evolució, hi ha divisions bacterianes que no presenten el repressor LexA. Possiblement, aquesta absència de *lexA* és deguda a una pèrdua en el transcurs de l'evolució associada a una reordenació de material genètic, ja que la majoria d'aquests microorganismes han patit una marcada reducció del tamany del seu genoma (*Aquifex aerolicus*, *Borrelia burgdorferi*, *Chlamydia pneumoniae*, *Mycoplasma pneumoniae*, *Campylobacter jejuni*, *Helicobacter pylori*, *Bacteroides fragilis* o *Porphyromonas gingivalis*). De la mateixa manera, no hi ha una homogeneïtat del sistema SOS a nivell de funcionament. Aquest, en termes generals, és homologable al d'*E. coli* en tots els microorganismes on s'ha estudiat, però existeix una certa variabilitat en la seva regulació i, sobretot, en la composició del reguló LexA.

1.2.1 El sistema SOS als bacteris Gramnegatius

1.2.1.1 Grup dels γ -proteobacteris

En aquesta divisió trobem dos grans grups de microorganismes. Un grup de patògens de vertebrats, sovint entèrics, i un altre grup de bacteris que habiten el sòl. Tots ells són molt pròxims a nivell filogenètic i coherentment amb això presenten un sistema SOS molt similar al d'*E. coli*, màxim representant d'aquest grup. Comparteixen el mateix motiu regulador CTGT-N₈-ACAG i una regulació similar. Dins d'aquests microorganismes trobem *Pseudomonas sp.*, *Salmonella sp.*, *Erwinia sp.*, *Providencia rettgeri*, *Aeromonas sp.* (Calero *et al.*, 1991; Garriga *et al.*, 1992; Riera i Barbé, 1993; Riera i Barbé, 1995) i molts altres dels quals es coneix la presència d'un LexA pràcticament idèntic al d'*E. coli* per estudis comparatius amb seqüències dels bancs de dades. No obstant això, ha estat demostrada una substancial heterogeneïtat a nivell de composició gènica del reguló en aquest grup en estudis *in silico* utilitzant la caixa SOS d'*E. coli* per trobar gens regulats per LexA en diferents espècies d'aquesta subdivisió (Erill *et al.*, 2003).

Tot i formar part del grup dels β -proteobacteris, *Ralstonia metallidurans*, segons estudis recents del nostre grup de recerca, presenta una caixa SOS com la dels γ -proteobacteris, amb un

sistema homòleg de regulació. Seria l'excepció en un grup caracteritzat per l'absència de *lexA* a la majoria dels seus representants, com ara la família *Neisseriaceae* (Black *et al.*, 1998; Fyfe i Davies, 1990) o el gènere *Chromobacterium* (Duarte *et al.*, 2004).

1.2.1.2 Grup dels α -proteobacteris

En aquest grup s'inclouen les famílies *Rhizobiaceae* i *Rhodobacteriaceae*, formades generalment per bacteris de vida lliure que habiten el sòl i presenten una gran varietat metabòlica. Al nostre laboratori es va dur a terme un estudi exhaustiu de la regulació del sistema SOS en aquest grup (Fernández de Henestrosa *et al.*, 1998; Tapias *et al.*, 1997; 1999; Labazi *et al.*, 1999; del Rey *et al.*, 1999). L'estudi de microorganismes com *Rhodobacter sphaeroides*, *R.capsulatus*, *Sinorhizobium meliloti* i *Paracoccus denitrificans* va revelar la presència d'un motiu de regulació diferent del d'*E. coli* (GTTC-N₇-GTTC o GAAC-N₇-GAAC), no tan sols en la seqüència, sinó també perquè presenta una simetria de repetició directa (Taula 1.3). Contràriament al que succeeix a *E. coli*, un mutant *lexA* (Def) es viable en aquest grup de proteobacteris (Tapias *et al.* 2000) sense necessitat de mutacions compensatòries que evitin l'aturada de la divisió cel·lular deguda a SulA. Aquest fet ens permetria explicar la variabilitat que s'observa en les caixes SOS a través de l'evolució, ja que si en totes les espècies hi hagués un mecanisme anàleg al de SulA controlat per LexA, seria poc probable que es donés un canvi al domini d'unió al DNA de LexA sense un efecte letal. Una última diferència respecte el funcionament del LexA d'*E. coli* és que a *R. sphaeroides* LexA tindria una funció dual, actuant de repressor o d'activador del gen *recA* en funció del nombre de caixes ocupades al promotor (Tapias *et al.*, 2002).

1.2.1.3 Grup dels δ -proteobacteris

El sistema SOS del grup delta dels proteobacteris ha estat estudiat recentment pel nostre laboratori a *Myxococcus xanthus*, una espècie bacteriana que presenta un cicle vital peculiar, amb un estadi de cossos fructífers.

A *M. xanthus* tan sols es troba una còpia de *lexA* i s'ha caracteritzat la seqüència consens de la seva caixa com a CTRHAMRYBYGTTTCAGS (Campoy *et al.*, 2003).

Tabla 1.3. Caixes SOS del grup α dels Proteobacteris

Microorganisme	Gen	Seqüència
<i>Rhodobacter sphaeroides</i>	<i>recA</i> (1)	GTTTCGCCTTATG A TC
	<i>recA</i> (2)*	GTTTCGCCTCAAGTTC
	<i>lexA</i> (1)*	GTTCTGCCCCGCGTTC
	<i>lexA</i> (2)*	GTTCACCGCCTGTTC
	<i>uvrA</i>	GTTCATACTATGTTC
	<i>uvrB</i> (1)	G C TCCGCCCTTGTTC
	<i>uvrB</i> (2)*	G A TCCGTTTTTGTTC
<i>Rhodobacter capsulatus</i>	<i>recA</i> (1)	GTTCCGAAATGTTC
	<i>recA</i> (2)*	GTTCTGCTTTCGTTC
	<i>uvrA</i>	GTTCCCTGTTCCGTTC
<i>Rhodopseudomonas viridis</i>	<i>recA</i> (1)	GTTCTCTTCTTGTTC
	<i>recA</i> (2)*	GT A CACGATTTGTTC
<i>Rhodopseudomonas palustris</i>	<i>recA</i> (1)*	GTTTCGCAATATGTTC
	<i>recA</i> (2)*	GT A CCCTATTTGTTC
<i>Agrobacterium tumefaciens</i>	<i>recA</i> (1)*	GTTCTGCTTTCGTTC
	<i>recA</i> (2)*	GT A CTCTATTTGTTC
	<i>recA</i> (1)*	GTTCTATATTTGTTC
<i>Rhizobium etli</i>		
<i>Sinorhizobium meliloti</i>	<i>recA</i> (2)*	GT A CCCTATTTGTTC
	<i>recA</i> (1)*	GTTTCGATTCTTGTTC
	<i>recA</i> (2)*	GT A CATGTTTTGTTC
<i>Mesorhizobium loti</i>	<i>uvrA</i>	GTTCTTTTTTTGTTC
	<i>recA</i> (1)*	GTTCTTTTTTTCGT A C
<i>Brucella abortus</i>	<i>recA</i> (2)*	GT A CCTTTTTTGTTC
	<i>recA</i> (1)	GTTTCGTGGATAGTTC
	<i>recA</i> (2)*	GTTCCATTCTTGTTC
<i>Paracoccus denitrificans</i>	<i>uvrA</i>	GTTTCGATATTTGTTC
	<i>Ssb</i> *	GTTCCCTGTTTTGTTC
	<i>recA</i> *	GTTTCACGGGTGTTC
<i>Aquaspirillum magnetotacticum</i>	<i>uvrA</i>	GTTCCCTGTGATGTTC
<i>Acidiphilium facilis</i>	<i>recA</i>	GTTCTCCTCTCGTTC
<i>Zynomonas mobilis</i>	<i>recA</i> *	GTT T TGTCAACGTTC
	<i>recA</i>	GTTCACCTTATGTTC
	<i>uvrA</i> (1)	A TTCCCCCTTTGTTC
<i>Sphingomonas aromaticivorans</i>	<i>uvrA</i> (2)	A TTCTGCTACCGTTC
	<i>recA</i> (1)	GTTCCCCCTTGTTC
	<i>recA</i> (2)*	GT A CTCGTTGTGTTC
<i>Caulobacter crescentus</i>	<i>recA</i> *	GT A CACTCTTTGTTC
	<i>lexA</i> (1)*	GTTCTCCTGGTGTTC
	<i>lexA</i> (2)*	GTT T GCGGTTTGTTC
Consens		GTTTCYYYTTTTGTTC

Modificat de Labazi *et al.*(1999) i Tapias i Barbé (1999).

Les bases ombrejades assenyalen els motius més conservats. Les bases en negreta assenyalen desviacions dels motius conservats GTTC. (*) *Seqüències reverses complementàries de la caixa original per facilitar la comparació*

Aquest motiu presenta una certa similitud amb el que es troba a *E. coli*, encara que no és reconegut per la proteïna LexA d'aquesta darrera. De totes maneres, *M. xanthus* és un cas interessant per la presència de dos gens *recA* (Norioka *et al.*, 1995). Un està regulat per LexA i és induïble (*recA2*), mentre que l'altre (*recA1*) presenta al seu promotor una reminiscència de la caixa SOS que no és funcional, si bé pot ser regenerada amb mutagènesi dirigida canviant tan sols alguns nucleòtids (Campoy *et al.*, 2003). Estudis amb un mutant *lexA* (Def) de *M. xanthus* han permès demostrar que gens com ara *ssb* i *recN*, que a *E. coli* estan regulats per LexA, estan regulats per un altre mecanisme independent i són induïbles davant de lesions al DNA (Campoy *et al.*, 2003).

1.2.2 El sistema SOS als bacteris Grampositius

Als bacteris grampositius, el sistema SOS ha estat extensament estudiat, tant a *Bacillus subtilis*, model per aquest grup de microorganismes, com més recentment a *Mycobacterium tuberculosis* i altres micobacteris (Davis *et al.* 2002).

B. subtilis és un bacteri formador d'espores i aeròbic. El seu hàbitat és el sòl i gaudeix d' un sistema de regulació SOS similar al d'*E. coli*. El seu gen *recA* presenta un 50% de similitud al d'*E. coli* i el seu anàleg de *lexA*, un 34%. Aquest anàleg de *lexA*, ha estat anomenat habitualment *dinR*, però últimament s'accepta la seva nomenclatura com a *lexA* i codifica el repressor del sistema SOS en aquest bacteri (Miller *et al.*, 1996; Winterling *et al.*, 1997). El motiu d'unió del repressor LexA a *B. subtilis* és conegut com la caixa "Cheo" (Cheo *et al.*, 1991; Yasbin *et al.*, 1992) i la seva seqüència consens, revisada i ampliada recentment (Miller *et al.*, 1996; Winterling *et al.*, 1998), seria el palíndrom CGAACRNRYGTTTCG.

La proteïna RecA és activada, als grampositius, de manera similar a com ho fa a *E. coli* i forma un filament nucleoprotèic que catalitza l'autohidròlisis de LexA. En aquest cas, el lloc de trencament es troba entre els aminoàcids Ala₉₁-Gly₉₂. La desaparició de LexA provoca la desrepressió del sistema i la inducció de les funcions SOS, on s'inclouen mecanismes de reparació, síntesi de translesió, filamentació cel·lular i inducció de profags (Yasbin *et al.*, 1992). Només una desena de gens *din* regulats per LexA han estat identificats fins al moment a

B. subtilis, encara que es suposa que el reguló SOS en aquest bacteri està compost per molts més gens. El sistema SOS a *B. subtilis* s'indueix també quan les cèl·lules presenten competència natural (Love *et al.*, 1985), però aquesta activació es dona de manera independent de RecA gràcies al producte del gen *comK*, que codifica la proteïna CTF (Haijema *et al.*, 1996; Hamoen *et al.*, 2001). A nivell de funcions SOS, el mecanisme de filamentació no és degut a l'acció de la proteïna Sula, sinó al producte del gen *yneA* (Kawai *et al.*, 2003), que actua sobre la proteïna FtsZ inhibint la septació. Aquest gen forma unitat transcripcional amb altres gens i està controlat pel mateix promotor que controla el gen *lexA* (de manera divergent).

Comparant els dominis estructurals i funcionals de la proteïna LexA de *B. subtilis* amb la d'altres bacteris, es pot veure com la regió carboxi-terminal és la més conservada. Aquesta regió és la responsable de la dimerització i de l'autohidròlisi, amb els residus Ser₁₁₉ i Lys₁₅₆. La regió amino-terminal presenta una elevada similitud estructural amb la resta de LexAs coneguts, amb les tres hèlixs característiques del domini d'unió al DNA, però la seva seqüència primària és molt diferent a la dels altres LexA caracteritzats fins al moment, amb canvis a les posicions dels residus que a *E. coli* contacten amb la caixa SOS (Knegtel *et al.*, 1995). Aquests canvis explicarien la diferent especificitat al reconèixer una caixa o l'altra.

No és estrany trobar múltiples operadors als promotors de gens SOS de *Bacillus subtilis*. De totes maneres, la unió de LexA a aquests operadors es produeix en forma de dímers de manera similar a com ho fa LexA a *E. coli* (Winterling *et al.*, 1998). Altres bacteris grampositius com *Clostridium perfringens*, *Corynebacterium glutamicum*, *Mycobacterium leprae*, *M. smegmatis*, *M. tuberculosis*, *Staphylococcus aureus*, *Streptomyces lividans* y *S. coelicolor* presenten un sistema SOS homòleg al de *B. subtilis* amb un motiu regulador tipus caixa "Cheo" (Taula 1.4). De totes maneres, i com s'ha comentat abans, hi ha estudis recents amb *M. tuberculosis* (Davis *et al.*, 2002) on s'ha proposat la caixa consens TCGAAC-N₄-GTTCGA per als micobacteris, establint una petita diferència entre bacteris grampositius d'alt contingut de GC i de baix contingut de GC.

Tabla 1.4. Possibles caixes SOS a bacteris grampositius i relacionats.

Microorganisme	Gen	Seqüència	
<i>Bacillus subtilis</i>	<i>recA</i>	CGAATATG CC TTTCG	
	<i>dinA</i>	CGAACTTTAGTTTCG	
	<i>dinB</i>	AGA ACTCATGTTTCG	
	<i>dinC</i> (1)	CGAACGTATGTT TG	
	<i>dinC</i> (2)	AGA ACAAGTGTTCG	
	<i>dinR</i> (1)	CGAACCTCAGTT TG	
	<i>dinR</i> (2)	CGAACAAACGTT TC	
	<i>dinR</i> (3)	GGA ATGTTTGTTCG	
	<i>uvrB</i>	AAA ACAAACGTTTCG	
	<i>dnaX</i>	CGAACCAAGGTT CA	
	<i>Bifidobacterium breve</i>	<i>recA</i> (1)	CGTT CACCC GCAT C
		<i>recA</i> (2)	CGAACAAATGTTTCG
<i>Clostridium perfringens</i>	<i>recA</i>	AGA ACTTATGTTTCG	
<i>Corynebacterium glutamicum</i>	<i>recA</i>	CGTAG GAATTTTCG	
<i>Corynebacterium pseudotuberculosis</i>	<i>recA</i>	AGA ATGGTCGTT AG	
<i>Lactococcus lactis</i>	<i>recA</i>	TGAT AGAAAGTT CC	
<i>Mycobacterium leprae</i>	<i>recA</i> (1)	CGAACAGATGTTTCG	
	<i>recA</i> (2)	CGTACTGCG AT TCG	
	<i>lexA</i> (1)	CGAACACATGTT TG	
	<i>lexA</i> (2)	CGAACATTCG AT TCG	
	<i>recA</i> (1)	CGAACAGGTGTTTCG	
<i>Mycobacterium smegmatis</i>	<i>recA</i> (1)	CGAACAGGTGTTTCG	
	<i>recA</i> (2)	GGA ACACCGGT CAG	
<i>Mycobacterium tuberculosis</i>	<i>recA</i>	CGAACAGGTGTTTCG	
	<i>lexA</i>	CGAACACATGTT TG	
<i>Spiroplasma melliferum</i>	<i>recA</i>	XGAT CACGAG AA CG	
<i>Spiroplasma citri</i>	<i>recA</i>	TGAT CACGAG AA CA	
<i>Staphylococcus aureus</i>	<i>recA</i>	CGAACAAAT AT TCG	
<i>Streptococcus mutans</i>	<i>recA</i>	CGAACATGCC CT TG	
<i>Streptococcus pneumoniae</i>	<i>recA</i>	GGAT CATTAG AA TG	
	<i>dinF</i>	TGA ACTTG AA ATCG	
<i>Streptococcus pyogenes</i>	<i>recA</i>	CG ATT AGGAG AA CG	
<i>Streptomyces lividans</i>	<i>recA</i>	CGAACATCC AT TCT	
<i>Streptomyces coelicolor</i>	<i>recA</i>	AGA ATGGATGTTTCG	
	<i>lexA</i>	CAA ACACACGTTTCG	
<i>Streptomyces clavuligenes</i>	<i>lexA</i>	CG TT CGAGT GAAA A	
<i>Streptomyces rimosus</i>	<i>recA</i>	CGAACGTCT AT TCA	
<i>Streptomyces ambofaciens</i>	<i>recA</i>	CGAACATCC AT TCA	
Consens		CGAACRNR Y GTTTCG	

Modificat de Yasbin *et al.* (1992), Durbach *et al.* (1997), Movahedzadeh *et al.* (1997a) i Winterling *et al.* (1998).

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.

Utilitzant aquesta seqüència consens, s'han trobat nous gens regulats per LexA a *M. tuberculosis*. Addicionalment, s'han trobat diferències remarcables amb *E. coli*, ja que gens regulats en aquest microorganisme per LexA no ho estarien a *M. tuberculosis*, tot i ser induïbles

en presència de dany al DNA. Aquest fet indica l'existència d'un mecanisme de regulació alternatiu. De fet, la majoria de gens induïbles pel dany al DNA ho farien de manera independent de LexA (Rand *et al.*, 2003). De la mateixa manera, alguns gens que a *E. coli* no formen part del reguló SOS sí hi estarien integrats als micobacteris. És interessant destacar que a *M. tuberculosis* no existeixen gens homòlegs d'*umuDC* que codifiquin per polimerases "error-prone". L'absència d'aquest tipus de polimerases suggereix una reparació més acurada a *M. tuberculosis* per poder sobreviure al dany al DNA. (Movahedzadeh *et al.*, 1997b; Brooks *et al.*, 2001).

Estudis recents realitzats al nostre laboratori han determinat la caixa LexA a *Dehalococcoides ethenogenes*, una bactèria que s'ha classificat dins del grup dels Bacteris verds no del sofre (Hugenholtz *et al.*, 1998). La seqüència AGAAC-N₄-GTTCT, motiu d'unió de LexA en aquest microorganisme, és pràcticament idèntica a la caixa de *B. subtilis*, el LexA del qual és capaç d'unir-s'hi específicament. A *D. ethenogenes*, però, no es troba més que un altre gen (*uvrA*) sota el control de LexA. Altres gens com *ruvA*, *ruvB*, *recN* i *recA* presentarien algun tipus diferent de regulació (Fernández de Henestrosa *et al.*, 2002). *D. ethenogenes* és el primer bacteri gramnegatiu on s'ha descrit una caixa SOS típica de bacteris grampositius.

1.3 Estudis filogenètics moleculars en bacteris

Generalment, la taxonomia molecular dels procariotes s'ha basat en la comparació de les seqüències de 16S rRNA. Aquesta filogènia diferencia dos grans grups de procariotes, els arquees i els bacteris, que van evolucionar de manera independent d'una arrel o ancestre comú (Woese, 1987). Aquesta classificació atorga dotze divisions o *phyla* al Domini *Bacteria* (Actinobacteris, Firmicutes, Cianobacteris, Clamídies, Proteobacteris, Espiroquetes, Grup CFB, *Deinococcus/Thermus*, Bacteris Verds no del Sofre, Planctomicets, Bacteris Verds del Sofre i Thermotogals). Es consideren divisions els grups que tenen una diferència d'entre el 20-25% entre les seves seqüències de 16S rRNA, utilitzant més d'un mètode d'anàlisi o diferents fragments del rRNA (Gupta *et al.*, 1998; Gupta *et al.*, 2000). Dues espècies s'accepten com a representants de la mateixa divisió quan la diferència entre les seves seqüències és, de manera

reproducible, inferior a aquest percentatge. Tanmateix, segueix encara vigent la discussió per estandarditzar els mètodes i els criteris amb què es construeixen els arbres filogenètics i els alineaments, ja que petites variacions en aquests mètodes poden fer variar substancialment els resultats.

Les dotze divisions de la classificació clàssica de l'arbre filogenètic del Domini *Bacteria* han estat posades en dubte per estudis recents i per la constant aparició de nous genomes als bancs de dades. Algunes revisions parlen d'aproximadament 40 divisions al Domini *Bacteria* (Hugenholtz *et al.*, 1998), que inclourien molts bacteris que no estan cultivats o divisions que tenen poques espècies cultivades. Moltes divisions es postulen tan sols a partir de la seqüenciació d'rRNA amplificat per PCR des de mostres ambientals. Aquestes divisions s'han anomenat "divisions candidates" per indicar la feblesa de les dades que donen suport a la seva existència (Fig. 1.7). Curiosament, algunes d'aquestes divisions candidates i algunes altres amb molts pocs membres seqüenciats fins al moment serien les responsables de processos ecològics molt importants i podrien trobar-se de manera ubíqua en gran varietat d'hàbitats. Per tant, encara ara estem subestimant bona part de la diversitat bacteriana en les nostres classificacions filogenètiques.

La utilització de proteïnes com RecA, presents tant en els bacteris com en els arqueobacteris, per establir arbres filogenètics (Eisen, 1995; Karlin *et al.*, 1995) obté resultats concordants quasi totalment amb als estudis amb rRNA. En l'actualitat però, l'intercanvi horitzontal d'informació genètica (LTG) s'ha començat a considerar un dels factors més importants de generació de diversitat bacteriana (Woese, 2000; Gupta i Griffiths, 2002). Les actuals divisions serien veritables calaixos de sastre on tenim espècies amb genomes construïts mitjançant combinacions d'adquisicions verticals i horitzontals on les últimes serien les principals responsables de l'evolució més recent de les espècies. Per intentar afinar en aquestes classificacions varis grups d'investigadors han posat èmfasi en la comparació de proteïnes "house-keeping", o altament conservades, al domini *Bacteria*, utilitzant la suma de la

informació d'insercions i delecions (*indels*) en regions de totes elles per determinar un arbre filogenètic amb punts d'embranchament més ben definits (Gupta, 1998; Gupta, 2000).

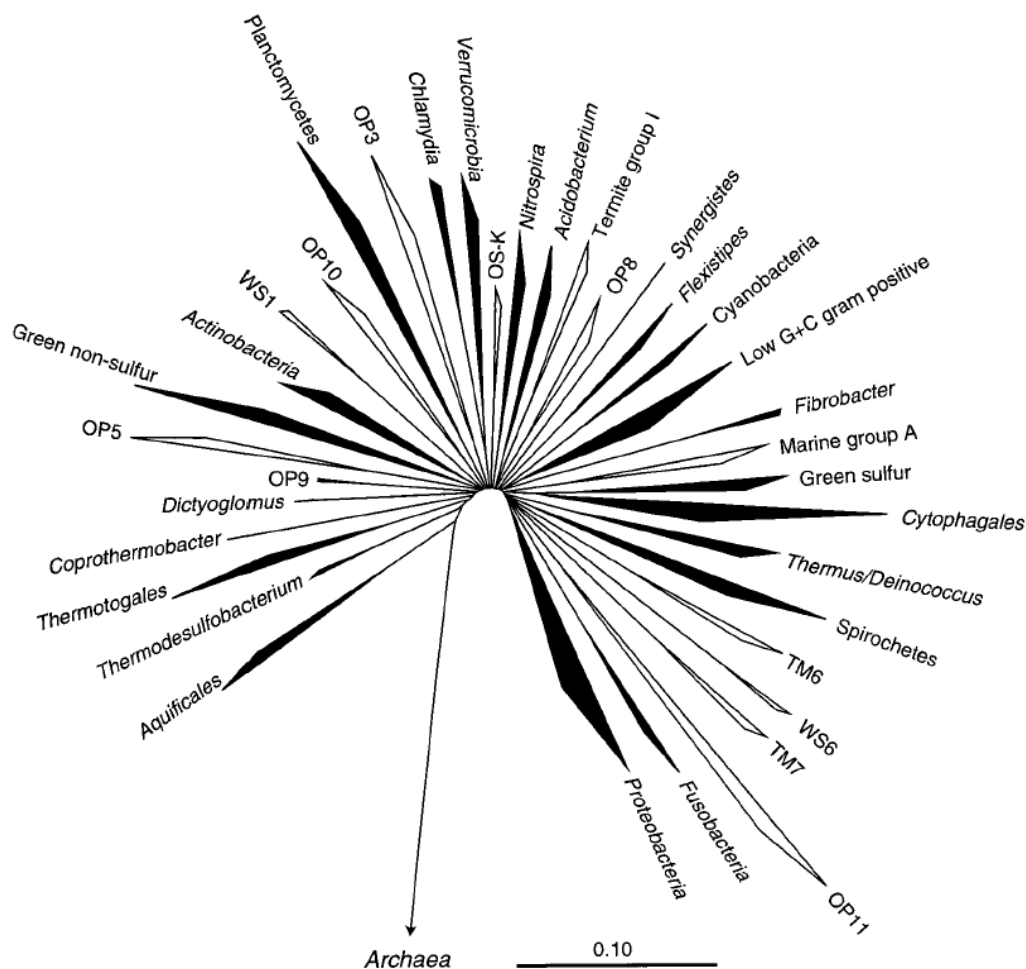


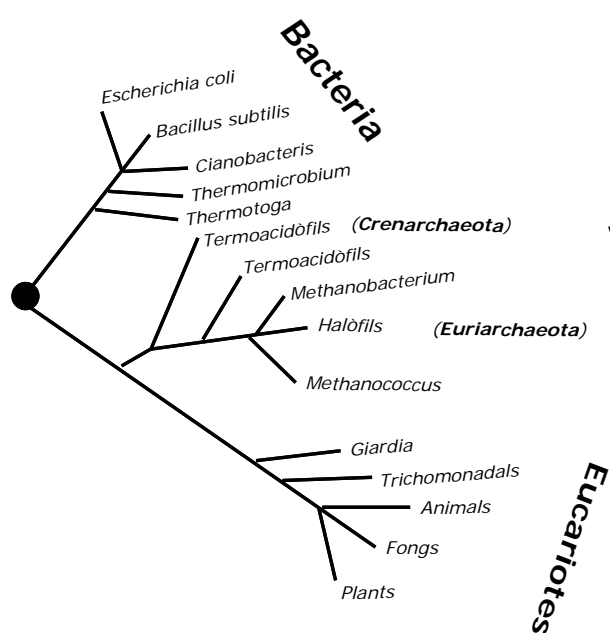
Fig. 1.7. Arbre filogenètic amb totes les divisions amb representants cultivats (en negre) i les divisions “candidates” basades en la seqüenciació directa de mostres de DNA obtingudes del medi (en blanc). Alguns termes apareixen en anglès. Modificat de Hugenholtz *et al.* (1998).

1.3.1 Filogènia basada en la presència d'*indels* a les proteïnes

Dins d'aquests estudis complementaris als alineaments de rRNA, els que han aconseguit més rellevància són els estudis de seqüències “signatura” (Gupta, 1998). Una seqüència “signatura” pot ser definida com una regió d'un alineament de proteïnes on es pot observar un canvi específic a la seqüència primària de la proteïna en tots els membres d'una divisió o de diverses, però no en membres d'altres divisions. Quan els canvis que observem són insercions i delecions d'aminoàcids, parlem d'*indels* (insertion-deletion). Aquestes seqüències han d'estar flanquejades per regions altament conservades que permeten descartar una naturalesa artificial de l'*indel* deguda a l'algoritme utilitzat per alinear les proteïnes. La característica que fa als

indels especialment atractius per als estudis filogenètics és el fet que no estan tan relacionats amb mutacions puntuals independents i poden ser considerats com un tret identificatiu d'un grup d'espècies. Quan un *indel* de longitud definida es troba de manera conservada en tots els membres d'un o diferents grups bacterians, la explicació més obvia i senzilla és que s'ha introduït un sol cop en un ancestre comú i ha passat als seus descendents. Per tant, les espècies que emergeixen més tard en l'evolució han de dur el mateix *indel* que aquelles que van fer-ho anteriorment. L'estudi d'*indels* de proteïnes com Hsp70, GroEL (Hsp60), GSI (glutamat-1-semialdehid 2,1 aminomutasa), proteïnes d'operons anabòlics i proteïnes ribosomals, totes elles conservades entre els dominis *Archaea* i *Bacteria*, demostra clarament la proximitat entre arquees i bacteris grampositius, que comparteixen *indels* característics. Com que ambdós grups comparteixen una mateixa estructuració dels embolcalls cel·lulars, més simple que els de la resta de bacteris, s'ha proposat un nou model per classificar els procariotes com a monoderms (Arqueobacteris i Grampositius) o diderms (Gramnegatius), on els monoderms serien ancestres dels diderms (Gupta, 1998) (Fig. 1.8). En aquest model els eucariotes no derivarien dels arquees sinó d'una quimera d'ambdós grups.

Model Woese



Model Gupta

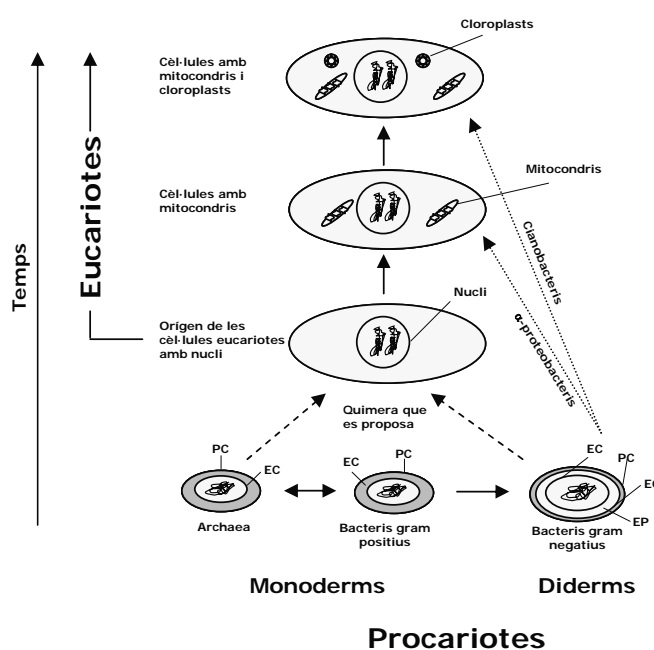


Fig. 1.8. Comparació dels models filogenètics de Woese (1987) i Gupta (1998). Modificat de Gupta *et al.* (1998). En el Model Gupta: EC (Embolcall cel·lular), PC (Paret cel·lular) i EP (Espai periplasmàtic).

De la mateixa manera, aquests autors han establert una línia evolutiva amb diferents punts de divergència o embrancament, on cada divisió actual seria la representant evolucionada d'un ancestre de la següent divisió (Fig. 1.9).

La possibilitat d'establir filogènies per diferents proteïnes i xarxes multigèniques, així com per les seqüències reguladores d'aquestes (Eisen i Hanawalt, 1999; Tan *et al.*, 2001; Makarova *et al.*, 2001, Panina *et al.*, 2001; Rodionov *et al.*, 2001 i Roy *et al.*, 2002), completa un mosaic que, poc a poc, ens aclareix el camí seguit per l'evolució en la gènesis de l'actual diversitat i ens dóna informació sobre la transferència horitzontal d'informació genètica (LTG) que ha passat de ser considerada un fenomen secundari en l'evolució a ser valorada com un dels principals mecanismes per generar diversitat. La caracterització dels motius reguladors i el funcionament del sistema SOS dels diferents grups bacterians pot dur-nos en aquest context a esbrinar la seva història evolutiva.

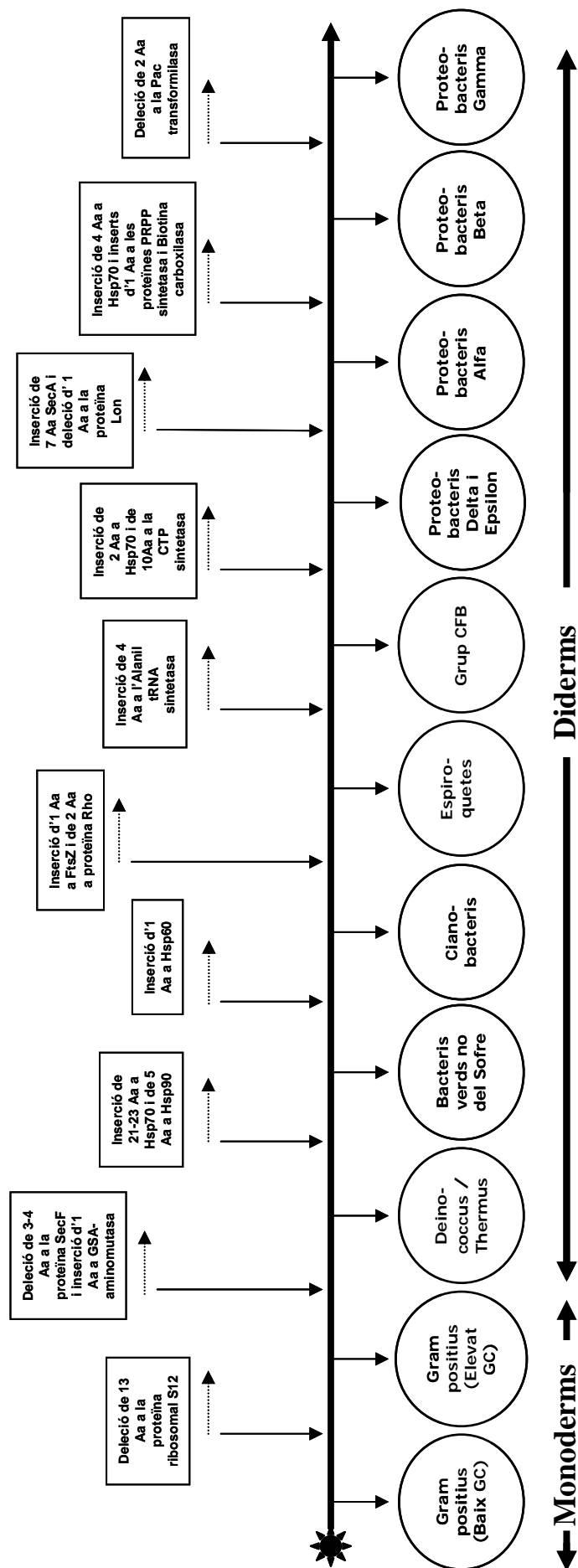


Fig. 1.9. Diferents punts de divergència entre les divisions del domini Bacteria basat en l'observació d'*indels* en les seqüències de proteïnes conservades. Modificat de Gupta *et al.* (2000).

1.4 Objectius

Tenint en compte tot el que s'ha exposat abans, ja podem definir les caixes d'unió de la proteïna LexA per a moltes de les divisions bacterianes conegudes, descartant aquelles en què s'ha descrit l'absència de *lexA* o on el reguló SOS presenta diferències molt notables en composició i regulació (Fig. 1.10). Això ens dóna una primera visió global de l'evolució de la caixa atenent a la disposició filogenètica clàssica, però no hi ha encara una relació directa evident entre la caixa LexA descrita als Grampositius i Bacteris verds no del sofre i els tres principals motius reguladors descrits als grups de Proteobacteris.

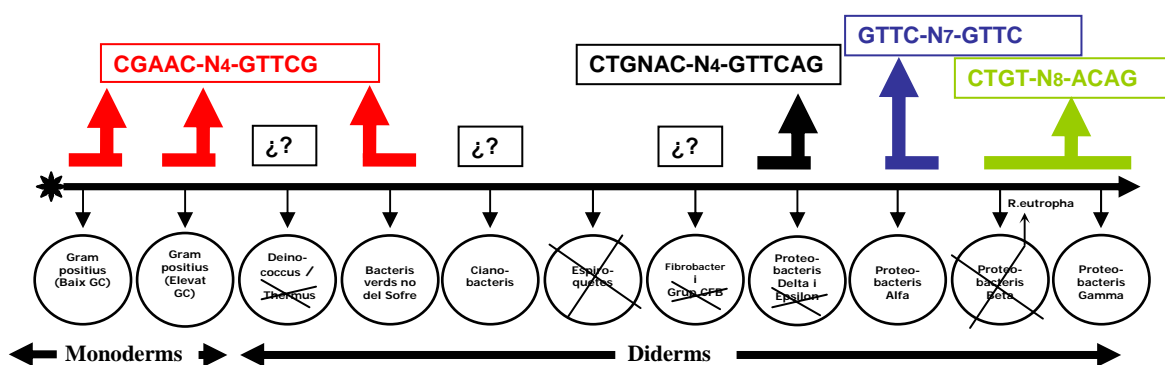


Fig.1.10. Principals motius reguladors descrits pel sistema SOS al domini *Bacteria* fins l'inici d'aquest treball disposats en ordre segons el model filogenètic de Gupta *et al.* (2000).

L'objectiu principal d'aquest treball ha estat la caracterització dels motius reguladors de dues divisions bacterianes (Cianobacteris i *Fibrobacter*) necessàries per completar una visió evolutiva del motiu d'unió de la proteïna LexA que ens permeti establir la relació entre els diferents motius reguladors descrits. Així mateix, s'ha dut a terme l'estudi del motiu d'unió de LexA a l'ordre Xanthomonadales del grup dels γ -proteobacteris per determinar la homogeneïtat a nivell de la caixa LexA en aquest *Phylum* bacterià.

Per a la divisió Cianobacteris es va escollir el microorganisme *Anabaena sp.* PCC7120, la seqüència del qual està disponible a la base de dades <http://www.kazusa.jp/cyanobase>. En el cas de la divisió *Fibrobacter*, es va utilitzar *Fibrobacter succinogenes* S85 per dur a terme els

estudis, obtenint la seqüència del seu gen *lexA* al banc de dades <http://www.tigr.org>. Per últim, es va escollir el patògen de plantes *Xylella fastidiosa* CVC 9a5c, on prèviament havia estat descrita la presència d'un gen *lexA* (Simpson *et al.*, 2000), per caracteritzar de manera precisa la seva caixa i determinar si aquesta es conserva a l'ordre *Xanthomonadales*.

En els tres casos s'ha seguit la mateixa aproximació estratègica:

1. Identificació als bancs de dades dels gens *lexA* mitjançant TBLASTN
 2. Amplificació per PCR i clonació del gen
 3. Sobreexpressió a *E. coli* BL21(λ DE) del producte del gen *lexA* identificat
 4. Purificació de la proteïna mitjançant el sistema d'afinitat de columna adequat
 5. Identificació de la regió d'unió de LexA mitjançant EMSA (mobilitat electroforètica), assaigs de protecció ("*footprinting*") i/o mutagènesi dirigida
 6. Experiments d'expressió *in vivo* i/o anàlisi *in silico* per definir la composició del reguló LexA
 7. Anàlisi comparatiu de les seqüències identificades i definició de la seva història evolutiva
-