

### **III.B. Toxicidad de la secuencia TRAPP**

Los resultados expuestos referentes a la expresión de formas delecionadas Ct y Nt del fragmento A5 así como a la fusión de estas formas, hacen referencia siempre a un mismo contexto y una estrategia de expresión recombinante. El contexto es siempre el de la secuencia de aminoácidos de la gE, y la estrategia es la expresión como fusión a una proteína acompañante o 'carrier', la MBP. El hecho que la secuencia TRAPP se haya expresado siempre junto con el resto de la secuencia de aminoácidos de la gE y fusionada a la MBP no permite concluir a todas luces que sea la secuencia TRAPP la máxima responsable de los efectos tóxicos observados. Para determinar la aportación que hace la secuencia TRAPP a la toxicidad observada se ha llevado a cabo dos tipos de experimentos. Por un lado, se ha expresado el dominio extracelular de la gE sin la secuencia TRAPP y no fusionado a la MBP, y por otro, se ha expresado la secuencia TRAPP como parte integrante de proteínas naturales y estables en *E. coli*.

#### **III.B.1. Expresión del dominio extracelular de la gE**

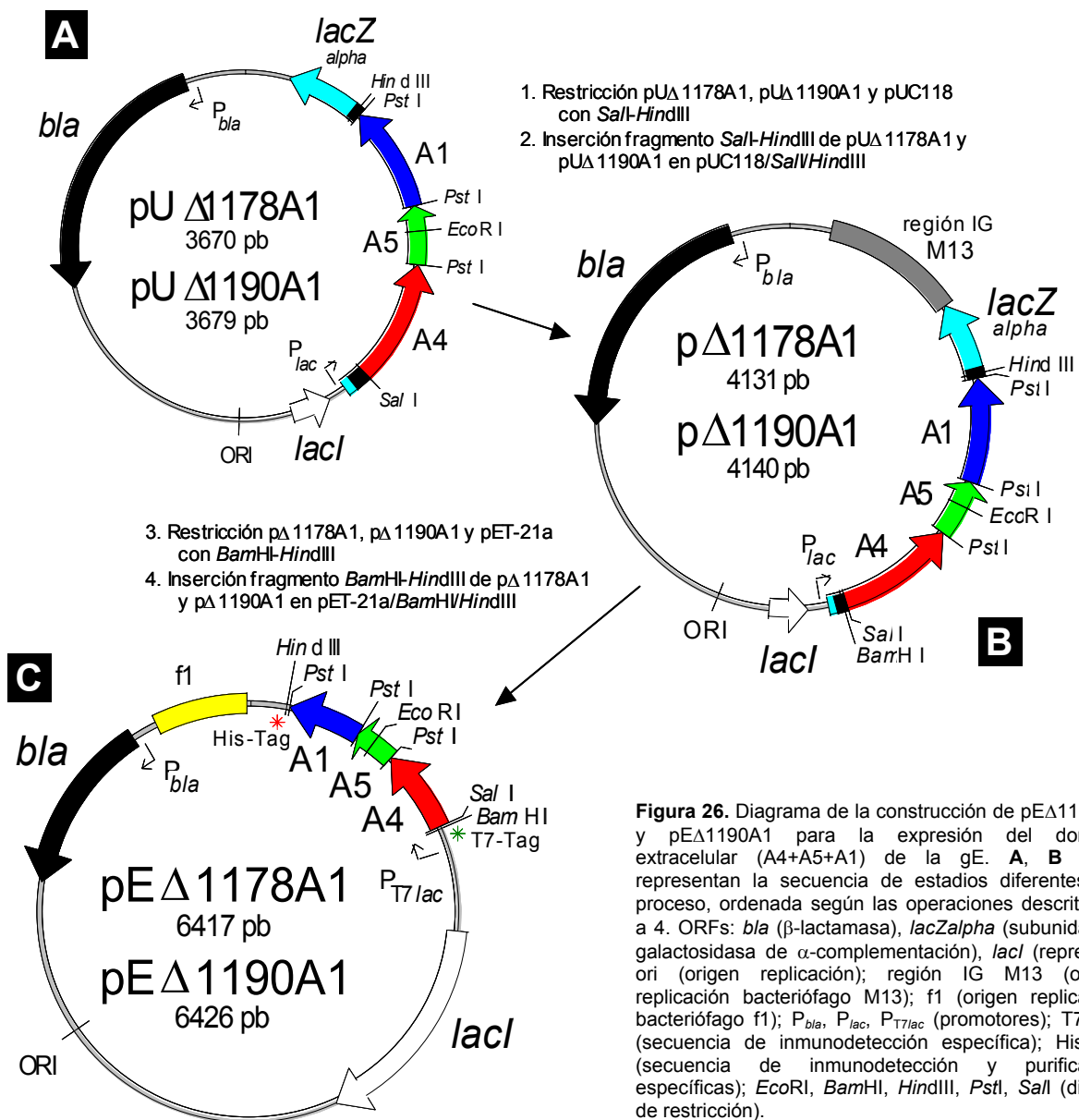
Los precedentes en la expresión de la gE del BoHV-1 llevados a cabo en nuestro laboratorio se iniciaron con la expresión del dominio extracelular íntegro de la gE. Inicialmente se ensayaron los sistemas de expresión pET-22b(+) y pMAL™-p2 que exportan la proteína recombinante al periplasma de *E. coli*, pero ambos se desestimaron porque la expresión de la proteína iba asociada a la inhibición del crecimiento del cultivo y a un elevado grado de proteólisis de la proteína recombinante. Ante la imposibilidad de expresar íntegramente el dominio extracelular de la gE se iniciaron entonces los ensayos de expresión de las formas truncadas A4, A5 y A1.

Los resultados obtenidos en la expresión de los clones p $\Delta$ 1178A1, p $\Delta$ 1190A1 y p $\Delta$ 4198A1 constituyen el primer logro de expresión en *E. coli* del dominio extracelular de la gE en su secuencia de aminoácidos casi íntegra. Sin embargo, esta expresión se da como proteína de fusión a la MBP y ello no permite hacer una valoración precisa de lo que supone eliminar la secuencia TRAPP de la gE. Para expresar únicamente el dominio extracelular sin la secuencia TRAPP se ha utilizado el sistema de expresión citoplasmático pET.

##### *III.B.1.A.1.a. Construcción*

La restricción de los plásmidos pU $\Delta$ 1178A1 y pU $\Delta$ 1190A1 con *Sall* y *HindIII* libera un fragmento de DNA de 987 pb y 996 pb, respectivamente (Fig. 26A). Estos fragmentos codifican para la secuencia de aminoácidos V<sub>40</sub>-Q<sub>366</sub> de la gE sin la secuencia TRAPP. Una vez purificados, estos fragmentos se han insertado en el plásmido pUC118 linearizado con *Sall* y *HindIII*. Los plásmidos obtenidos se han llamado p $\Delta$ 1178A1 y p $\Delta$ 1190A1 (Fig. 26B). La digestión de p $\Delta$ 1178A1 y p $\Delta$ 1190A1 con *BamHI* y *HindIII* libera, respectivamente, un fragmento de 999 pb y de 1008 pb. Éstos, una vez purificados, se han insertado en el plásmido pET-21a linearizado con *BamHI* y *HindIII*. Los plásmidos obtenidos se han denominado pE $\Delta$ 1178A1

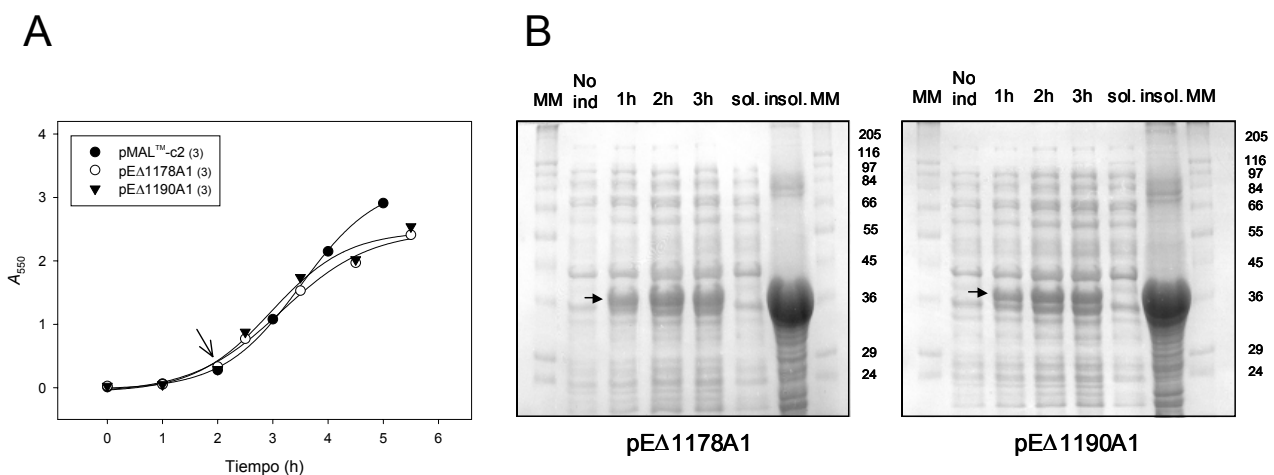
y pE $\Delta$ 1190A1 (Fig. 26C), y con ellos se han transformado células competentes de la cepa BL21(DE3) de *E. coli*.



La construcción de los plásmidos pET se podría haber realizado directamente mediante la inserción en el plásmido pET-21b del fragmento de DNA *SalI-HindIII* de pM $\Delta$ 1178A1 o pM $\Delta$ 1190A1, en la fase de lectura del gen *gE*. Sin embargo, la linearización *SalI-HindIII* de los plásmidos de la serie pET-21 puede tener un rendimiento bajo debido a la proximidad entre sí de estas dianas de restricción en la MCS del plásmido. Esta proximidad no afecta la eficacia de linearización del plásmido por la acción de una de las dos enzimas, pero sí condiciona la eficacia de corte de la segunda enzima pues su secuencia diana se halla entonces muy próxima a un extremo de la molécula lineal de DNA. Es por esto que se ha optado por incluir el paso intermedio de clonación en pUC118 donde las dianas *SalI* y *HindIII* están más distantes, y posteriormente utilizar las dianas *BamHI* y *HindIII* en la MCS del plásmido pET-21a.

## III.B.1.A.1.b. Expresión

La cinética de crecimiento de los cultivos de expresión de los clones BL21(DE3) pE $\Delta$ 1178A1 y pE $\Delta$ 1190A1 muestra que hay cierta inhibición del crecimiento tras la inducción de expresión de la proteína recombinante (Fig. 27A). Esta inhibición, sin embargo, no refleja la toxicidad asociada a la expresión de las formas tóxicas de la gE —los productos de delección Ct pM $\Delta$ 49, pM $\Delta$ 412 y pM $\Delta$ 41, los productos de delección Nt pM $\Delta$ 515A1 y pM $\Delta$ 715A1, o la fusión pM $\Delta$ 4859A1— sino que se asocia a la expresión del fragmento A4 de la gE. En este sentido cabe señalar que los cultivos de expresión de los productos de delección Ct no tóxicos pM $\Delta$ 414 y pM $\Delta$ 46 (Fig. 16A) y de las fusiones no tóxicas pM $\Delta$ 1178A1, pM $\Delta$ 1190A1 y pM $\Delta$ 4198A1 (fig 24A) también registraban valores de  $A_{550}$  post-inducción inferiores al cultivo control.



**Figura 27.** Expresión en *E. coli* del dominio extracelular de la gE sin la secuencia TRAPP. **(A)** Cinética de crecimiento de los clones de *E. coli* BL21(DE3) que codifican para los fragmentos A4-A5-A1 de la gE sin la secuencia TRAPP. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. La flecha señala el momento de inducción de los cultivos. El clon pMAL<sup>TM</sup>-c2 constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo de los cultivos pE $\Delta$ 1178A1 y pE $\Delta$ 1190A1, y de las fracciones soluble e insoluble. 'No ind': extracto crudo del cultivo pE $\Delta$ 1178A1 o pE $\Delta$ 1190A1 no inducido. '1h', '2h', '3h': extracto crudo del cultivo a 1, 2 y 3 horas post-inducción. 'sol.': fracción soluble a 3 horas post-inducción (la cantidad de cultivo aplicada en 'sol.' y '3h' es equivalente). 'insol.': fracción insoluble a 3 horas post-inducción. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas denotan la acumulación de la proteína expresada. Condiciones de electroforesis: gel acrilamida 12% y 0.8 mm, 20-30 mA/gel.

El análisis por SDS-PAGE de los extractos crudos de los cultivos inducidos pE $\Delta$ 1178A1 y pE $\Delta$ 1190A1 muestra claramente la presencia de bandas más intensas (Fig. 27B). La migración electroforética de estas bandas se corresponde con la teórica esperada para la proteína recombinante en cada caso: 38.4 kDa para el fragmento de la gE codificado en pE $\Delta$ 1178A1 y 38.8 kDa para el codificado en pE $\Delta$ 1190A1. En definitiva, estos resultados reflejan la acumulación en el citoplasma de *E. coli* del dominio extracelular de la gE del BoHV-1 ( $V_{40}$ - $Q_{366}$ ) sin la secuencia de aminoácidos TRAPP y sin estar fusionado a la MBP. Durante el fraccionamiento de estos cultivos se ha puesto de manifiesto, sin embargo, que la proteína acumulada en ambos casos es insoluble (Fig. 27B) y que forma cuerpos de inclusión en las células productoras. Para estos experimentos no se ha realizado ningún ensayo de expresión de la secuencia  $V_{40}$ - $Q_{366}$  de la gE portadora de la secuencia TRAPP que sirva de contrapunto al resultado obtenido. El único referente lo constituye las pruebas de expresión periplasmática con el plásmido pET22b(+) que se realizaron previamente en nuestro laboratorio (Rebordosa, 1996). En aquella ocasión no se detectó acumulación alguna de la proteína recombinante por SDS-PAGE, y únicamente la transferencia 'Western' revelaba que sí hubo expresión. La

banda detectada tenía una movilidad electroforética de aproximadamente 36 kDa, muy próxima a los 38.5 kDa esperados para la forma procesada de la proteína recombinante.

La producción por vez primera en *E. coli* del dominio extracelular casi íntegro de la gE del BoHV-1 nos ha llevado a realizar algunas pruebas de inmunodetección de la proteína expresada. Estos ensayos se han llevado a cabo sobre transferencia 'Western' utilizando sueros de conejo anti-fragmento A4 y anti-BoHV-1 FM (Rebordosa, 1996), y sueros de terneras convalecientes de IBR. Los resultados obtenidos (no mostrados) indican que el dominio extracelular expresado de la gE sólo es reconocido por el suero de conejo anti-fragmento A4. El hecho de que la proteína expresada no sea reconocida por el antisuero dirigido contra el virus en su estado natural, apunta a que la conformación adoptada no se corresponde con la estructura nativa de la gE. La ausencia de conformación nativa explica en parte el carácter insoluble de la proteína expresada, y sugiere que para preservar la conformación y la inmunogenicidad nativas pueden ser necesarias las glicosilaciones descritas en la gE vírica (Rebordosa, 1996).

Los resultados en la expresión del dominio extracelular de la gE sin la secuencia TRAPP y sin estar fusionado a la MBP, ponen de manifiesto que la delección de TRAPP suprime la inhibición del crecimiento de los cultivos durante la expresión de la secuencia original V<sub>40</sub>-Q<sub>366</sub> de la gE. Asimismo, la delección de la secuencia TRAPP permite la acumulación de la proteína recombinante en el citoplasma de *E. coli*. Por otro lado, el hecho de que la expresión no se haya desarrollado de acuerdo a una estrategia de proteína de fusión evidencia la naturaleza insoluble en *E. coli* del dominio extracelular de la gE del BoHV-1.

### **III.B.2. Expresión de la secuencia TRAPP en proteínas de *E. coli***

En los experimentos de expresión de las formas delecionadas Ct y Nt del fragmento A5, de los productos de fusión de estas formas delecionadas, o del dominio extracelular de la gE sin la MBP, se observa el efecto positivo que la delección de la secuencia TRAPP tiene sobre el cultivo de expresión. La secuencia TRAPP, en caso de ser realmente tóxica, también debiera inhibir el crecimiento si se expresa fusionada en proteínas de *E. coli*. En este sentido se han llevado a cabo algunos experimentos de expresión de proteínas naturales y estables de *E. coli*, en las que se ha incorporado secuencias de 4-7 aminoácidos de la región tóxica de la gE. Algunos de estos péptidos contienen la secuencia TRAPP mientras que en otros esta secuencia está truncada. Las proteínas escogidas para expresar estos péptidos han sido la MBP y la  $\beta$ -galactosidasa.

#### **III.B.2.A. Expresión de la secuencia TRAPP en la MBP**

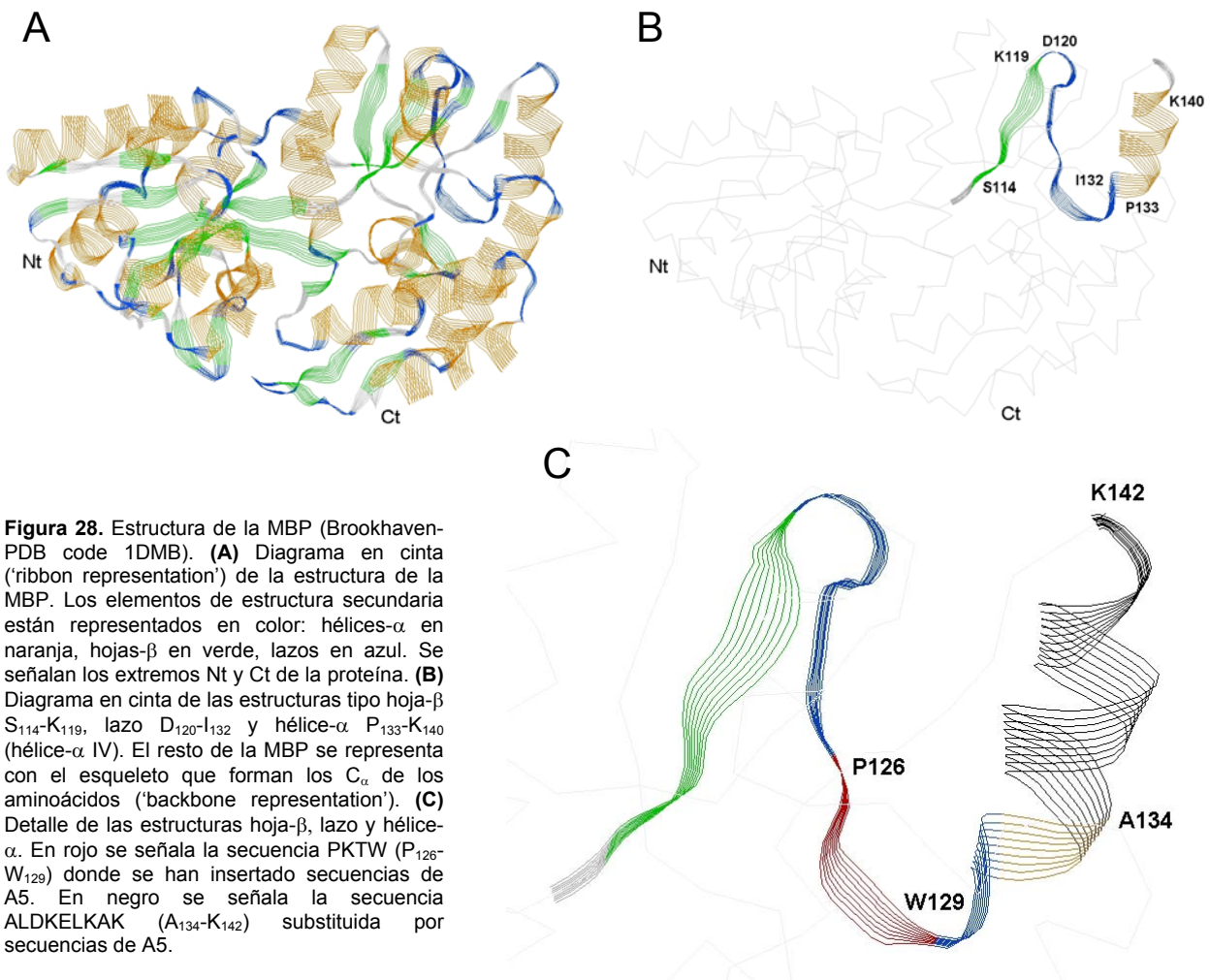
Los péptidos tóxicos del fragmento A5 de la gE se han insertado en la MBP mediante adaptadores. Estos adaptadores se han insertado en el gen *malE*, codificante para la MBP, de dos maneras distintas:

- i) por inserción en una diana de restricción de manera que la MBP expresada presenta una ganancia neta (inserción) de aminoácidos, y
- ii) por inserción entre dos dianas de restricción de manera que la MBP expresada presenta la substitución de parte de su secuencia de aminoácidos.

### III.B.2.A.1. Inserción de la secuencia TRAPP en la MBP

La región de la MBP elegida para la inserción de péptidos corresponde a la estructura tipo lazo formada por la secuencia de aminoácidos  ${}_{120}\text{DLLPNPPKTWEEI}_{132}$ . Esta estructura, relativamente oculta en la MBP, une las estructuras hoja- $\beta$  y hélice- $\alpha$  que forman, respectivamente, las secuencias de aminoácidos  $\text{S}_{114}\text{-K}_{119}$  y  $\text{P}_{133}\text{-K}_{140}$  (Fig. 28A, B). La elección de este lazo para la inserción de péptidos se ha hecho en base a:

- i) su estructura de lazo relativamente rica en aminoácidos prolina, y
- ii) la proximidad de la hélice- $\alpha$  IV, que es una estructura tolerante a inserciones y deleciones de péptidos y sobre la cual se han llevado a cabo numerosos experimentos de expresión y presentación de antígenos peptídicos (Martineau *et al.*, 1992; Rodseth *et al.*, 1990; Coëffier *et al.*, 2001; Clément *et al.*, 1996).



#### III.B.2.A.1.a. Construcción

Las secuencias de aminoácidos expresadas en el lazo  $\text{D}_{120}\text{-I}_{132}$  de la MBP han sido RAPP y TRAPP. Para cada secuencia, se ha diseñado dos tipos de adaptadores: uno en el uso de codón de la gE del BoHV-1, y otro en el uso de codón optimizado de *E. coli*. (Tabla IV). Estos adaptadores se han insertado en la diana

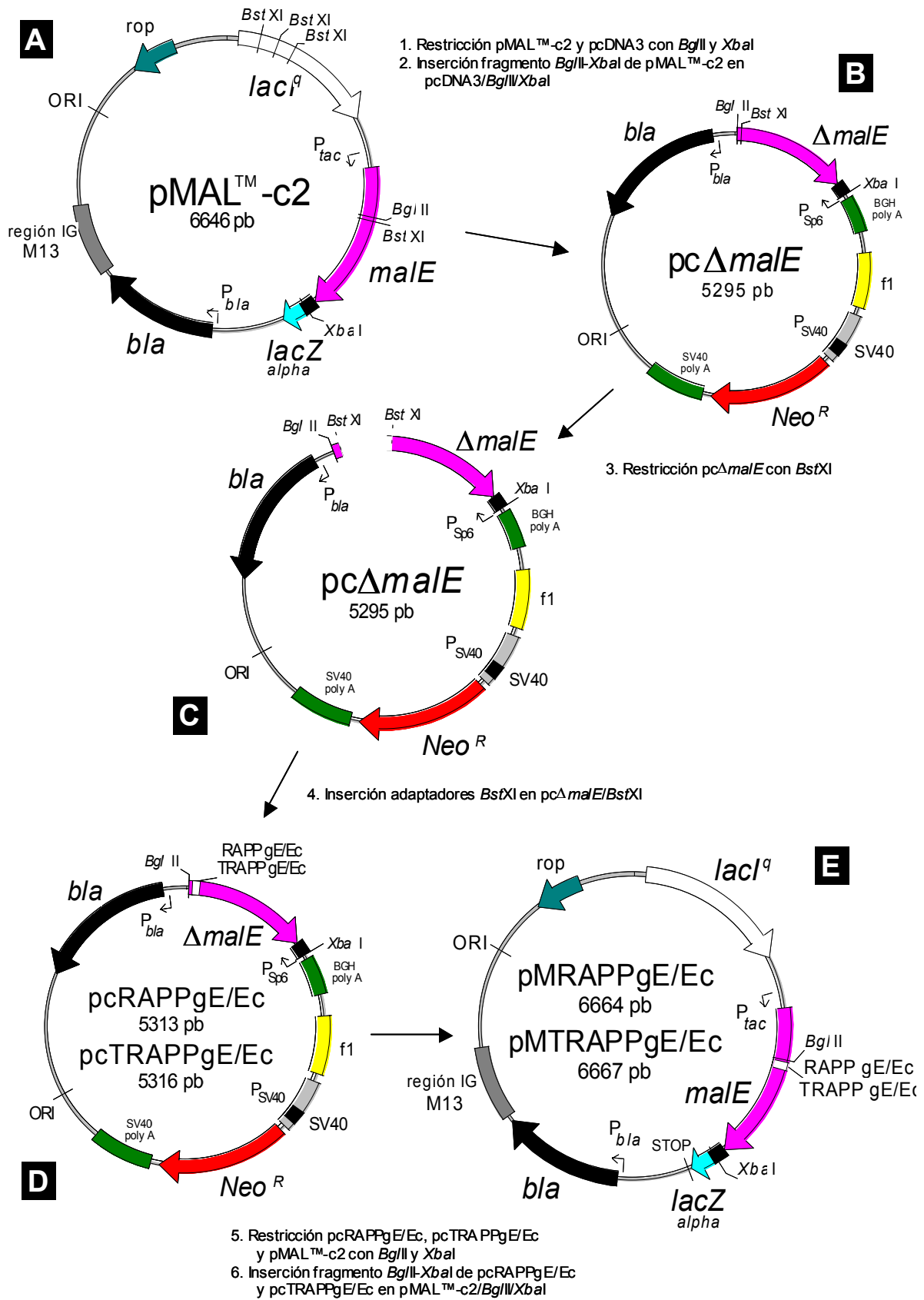
de restricción *Bst*XI contenida en la región del gen *malE* codificante para dicho lazo. Una consecuencia directa de la inserción de los adaptadores es la duplicación parcial de la secuencia de la diana *Bst*XI. Ello se debe a que *Bst*XI es una enzima que genera extremos cohesivos. La diana *Bst*XI ocupa los codones correspondientes a P<sub>126</sub>, K<sub>127</sub> T<sub>128</sub> y W<sub>129</sub> (PKTW) (Fig. 28C), y tras la inserción de los adaptadores los codones codificantes para K<sub>127</sub> y T<sub>128</sub> se duplican. Tras la inserción del adaptador RAPP la secuencia completa a traducir es PKTRAPPKTW, mientras que para el adaptador TRAPP resulta ser PKTTRAPPKTW. Cabe señalar que debido a que la diana *Bst*XI implica al codón codificante para la T<sub>128</sub>, tras la inserción de RAPP la secuencia que se traduce es TRAPP, y que tras la inserción de TRAPP se traduce TTRAPP.

TABLA IV. Adaptadores *Bst*XI para la inserción de las secuencias RAPP y TRAPP en la MBP

Tipo de adaptador	Secuencia	Uso de codón	Plásmido de expresión
RAPP	5' <b>R A P P K</b> acgcgcgccccgaaaac 3' tttg <del>tgcgcgcggggct</del> <i>Bst</i> XI* <i>Bst</i> XI*	gE BoHV-1	pMRAPPgE
	5' <b>R A P P K</b> cgcgcgcgcgcgaaaac 3' ttggg <del>cacgcgcggct</del> <i>Bst</i> XI* <i>Bst</i> XI*	<i>E. coli</i>	pMRAPPEc
TRAPP	5' <b>T R A P P K</b> gacacgcgcgccccgaaaac 3' tttg <del>ctgtgcgcgcggggct</del> <i>Bst</i> XI* <i>Bst</i> XI*	gE BoHV-1	pMTRAPPgE
	5' <b>T R A P P K</b> caccgcgcgcgcgcgaaaac 3' ttggg <del>tgggcacgcgcggct</del> <i>Bst</i> XI* <i>Bst</i> XI*	<i>E. coli</i>	pMTRAPPEc

*Bst*XI\* significa que no se conserva la diana de restricción *Bst*XI tras la inserción del adaptador.

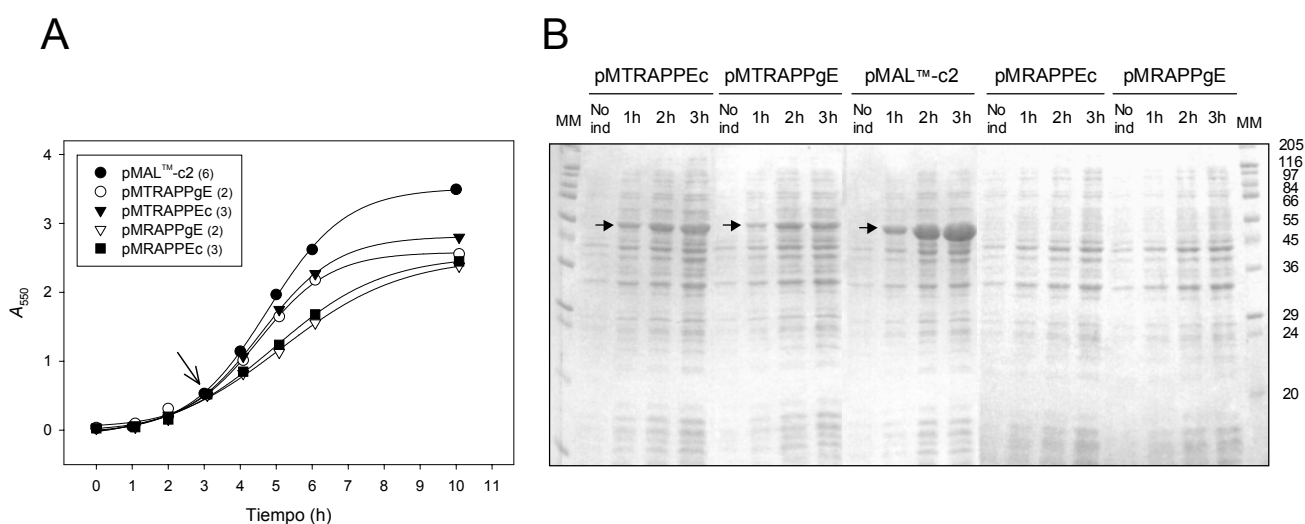
La inserción de los adaptadores en el gen *malE* no se puede realizar directamente sobre el plásmido pMAL<sup>TM</sup>-c2 debido a que la diana *Bst*XI no es una diana única en este plásmido (Fig. 29A). Por consiguiente, ha sido necesario subclonar el gen *malE*. La digestión del plásmido pMAL<sup>TM</sup>-c2 con *Bgl*II y *Xba*I libera un fragmento de DNA de 820 pb codificante para la secuencia de la MBP desde el aminoácido D<sub>120</sub>. Este fragmento, una vez purificado, se ha insertado en el plásmido pcDNA3 previamente linearizado con *Bgl*II y *Xba*I (Fig. 29B). El plásmido obtenido se ha llamado pcΔ*malE*, y presenta como única diana *Bst*XI la contenida en el gen *malE*. El plásmido pcΔ*malE* se ha linearizado con *Bst*XI (Fig. 29C), y una vez purificado se le han insertado los adaptadores *Bst*XI. Los plásmidos resultantes se han denominado pcRAPPgE, pcRAPPEc, pcTRAPPgE y pcTRAPPEc (Fig. 29D). Cabe señalar que debido a que la diana de restricción *Bst*XI no es palindrómica, la inserción de los adaptadores se da únicamente en el sentido deseado haciéndose innecesaria la comprobación posterior de su orientación. Por otro lado, la inserción de estos adaptadores comporta la desaparición de la diana *Bst*XI. La digestión de los plásmidos pcRAPPgE o Ec y pcTRAPPgE o Ec con *Bgl*II y *Xba*I libera, respectivamente, dos fragmentos de 837 y 840 pb codificantes para la secuencia de la MBP desde el aminoácido D<sub>120</sub> portadora de las inserciones RAPP o TRAPP. Estos fragmentos han sido purificados e insertados en el plásmido pMAL<sup>TM</sup>-c2 linearizado con *Bgl*II y *Xba*I (Fig. 29E). Los plásmidos obtenidos se han llamado pMRAPPgE, pMRAPPEc, pMTRAPPgE y pMTRAPPEc, y con ellos se han transformado células competentes XL1-Blue para ensayar la expresión de la MBP que codifican.



**Figura 29.** Diagrama del proceso de inserción de las secuencias RAPP y TRAPP en la MBP. **A, B, C, D y E** representan la secuencia de estadios diferentes del proceso, ordenada según las operaciones descritas 1. a 6. ORFs: *bla* ( $\beta$ -lactamasa), *lacZ* $\alpha$  (subunidad  $\beta$ -galactosidasa de  $\alpha$ -complementación), *lacI*<sup>q</sup> (represor), *malE* y  $\Delta$ *malE* (MBP), *Neo*<sup>R</sup> (resistencia a neomicina); ori (origen replicación); región IG M13 (origen replicación bacteriófago M13); f1 (origen replicación bacteriófago f1); SV40 (origen replicación poliomavirus SV40); *P*<sub>*bla*</sub>, *P*<sub>*lac*</sub>, *P*<sub>*Sp6*</sub>, *P*<sub>*SV40*</sub> (promotores); *Bgl*II, *Bst*XI, *Xba*I (dianas de restricción).

## III.B.2.A.1.b. Expresión

La cinética de crecimiento de los cultivos que expresan la MBP portadora de la secuencia TRAPP o TTRAPP muestra que, tras inducir la expresión, ninguno de ellos alcanza los valores de  $A_{550}$  del cultivo que expresa la MBP nativa (Fig. 30A). Este resultado indica que la expresión de la MBP que contiene TRAPP o TTRAPP reduce la tasa de crecimiento del cultivo. Por otro lado, no se observan diferencias entre las cinéticas de crecimiento de dos cultivos que expresan la misma inserción pero que tienen un uso de codón diferente. Este resultado indica que la toxicidad que desencadenan las secuencias insertadas no viene dada por el uso de codón de la gE sino por la propia naturaleza de la secuencia aminoacídica. A diferencia de lo observado en la expresión de las formas delecionadas Ct o Nt o de sus fusiones, los resultados obtenidos también indican que los cultivos pMTRAPPgE y pMTRAPPEc que expresan la secuencia tóxica más larga (TTRAPP) son los que presentan una menor inhibición del crecimiento, mientras que los cultivos pMRAPPgE y pMRAPPEc que expresan una secuencia más corta (TRAPP) presentan una inhibición mayor del crecimiento.

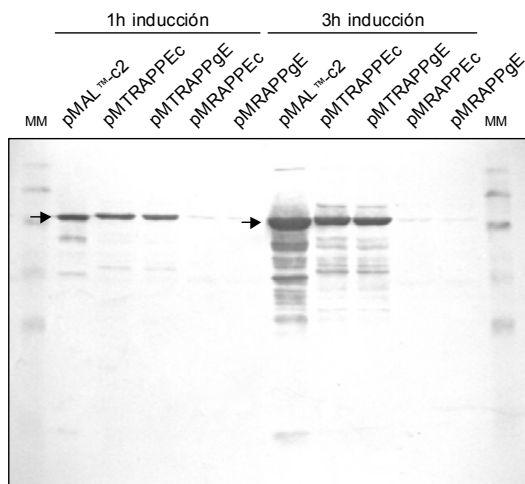


**Figura 30.** Expresión en *E. coli* de la secuencia TRAPP. **Inserción** de TRAPP en el lazo D120-I132 de la MBP. **(A)** Cinética de crecimiento de los clones de *E. coli* **XL1-Blue** que codifican para las inserciones RAPP y TRAPP. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. La flecha señala el momento de inducción de los cultivos. El clon pMAL<sup>TM</sup>-c2 constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo a 1, 2 y 3 horas post-inducción. 'No ind': extracto crudo del cultivo no inducido. '1h', '2h', '3h': extracto crudo del cultivo a 1, 2 y 3 horas post-inducción. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas denotan la acumulación de la proteína expresada. Condiciones de electroforesis: gel acrilamida 15% y 0.8 mm, 15-20 mA/gel.

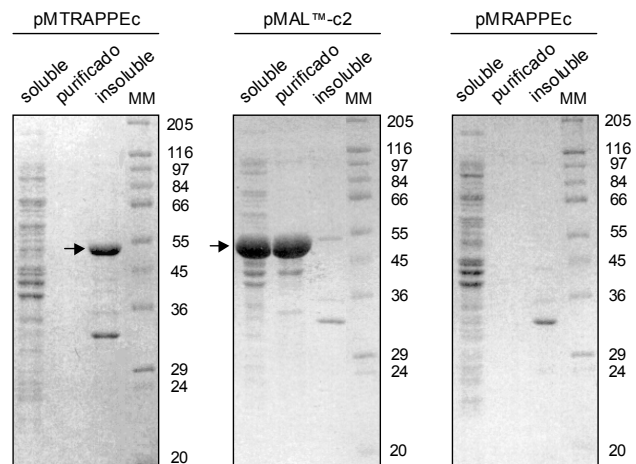
El análisis por SDS-PAGE del extracto crudo de los cultivos de expresión muestra la presencia de bandas intensas únicamente en las muestras de los cultivos inducidos pMAL<sup>TM</sup>-c2 (control), pMTRAPPgE y pMTRAPPEc (Fig. 30B). La migración electroforética de estas bandas se corresponde con la esperada para la proteína recombinante y refleja, por tanto, la acumulación de la MBP nativa en el cultivo control inducido, y de la MBP que contiene la secuencia TTRAPP en los cultivos inducidos pMTRAPPgE y pMTRAPPEc. Por el contrario, no se aprecia acumulación alguna de la MBP portadora de la secuencia TRAPP en los cultivos inducidos pMRAPPgE y pMRAPPEc.



Mediante una transferencia 'Western', y utilizando suero policlonal anti-MBP, se confirma que la proteína acumulada en los cultivos pMAL™-c2, pMTRAPPgE y pMTRAPPEc corresponde a la MBP (Fig. 31). También se pone de manifiesto que en los cultivos de expresión pMRAPPgE y pMRAPPEc sí hay en realidad expresión, aunque mínima, de la MBP portadora de la secuencia TRAPP. Por otro lado, en todas las muestras de inducción es patente un fenómeno de degradación o de síntesis intermedia de la MBP. Este hecho es particularmente acusado en la muestra del cultivo control inducido que expresa la MBP nativa. El fraccionamiento de los cultivos control, pMTRAPPgE y pMTRAPPEc revela que, mientras la práctica totalidad de la MBP nativa expresada en el cultivo control es soluble, la MBP que contiene la secuencia TTRAPP en los cultivos pMTRAPPgE y pMTRAPPEc es insoluble (Fig. 32).



◀ **Figura 31.** Inserción de TRAPP en la MBP. Análisis sobre transferencia 'Western' de la expresión de TRAPP. SDS-PAGE en condiciones reductoras del extracto crudo de XL1-Blue pMRAPPgE/Ec y pMTRAPPgE/Ec a 1 y 3 horas post-inducción: gel acrilamida 15% y 0.8 mm, 10 mA, 150 min. Transferencia: 0.18-0.27 A, 60 min. Inmudetección: dilución 1:5000 de suero policlonal anti-MBP de conejo (Protein Fusion and Purification System pMAL™-2, New England Biolabs) y dilución 1:2000 de suero policlonal anti-conejo de cabra, marcado con peroxidasa (Nordic Immunology). Las flechas denotan la MBP expresada. 'MM': marcador preteñido de peso molecular (SeeBlue™ Pre-Stained Standards, Novex). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM.



▶ **Figura 32.** Inserción de TRAPP en la MBP. Solubilidad de la MBP expresada. SDS-PAGE en condiciones reductoras del extracto crudo fraccionado tras 3 horas de inducción. 'soluble': fracción soluble. 'purificado': extracto purificado de la fracción soluble mediante resina de amilosa (Protein Fusion and Purification System pMAL™-2, New England Biolabs). 'insoluble': fracción insoluble. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas señalan la MBP. Condiciones de electroforesis: gel acrilamida 12% y 0.8 mm, 12.5-15 mA/gel.

El resultado de la expresión de la secuencia TRAPP insertada en el lazo D<sub>120</sub>-I<sub>132</sub> de la MBP constituye una prueba directa de la toxicidad que implica la expresión de dicha secuencia para la célula hospedadora. Por vez primera se demuestra que la expresión de la secuencia de aminoácidos TRAPP en una proteína de *E. coli* conlleva toxicidad. Esta toxicidad se traduce en la inhibición del crecimiento y la no acumulación de la proteína recombinante. Por otro lado, el grado de toxicidad de TRAPP insertada en el lazo D<sub>120</sub>-I<sub>132</sub> de la MBP pudiera depender de la estructura final que adopta la secuencia de aminoácidos. En comparación a TRAPP, la inserción de TTRAPP en el mismo punto, ni es tan inhibitoria para el crecimiento celular ni evita la acumulación, si bien de forma insoluble, de la MBP.

## III.B.2.A.2. Substitución de la secuencia TRAPP en la MBP

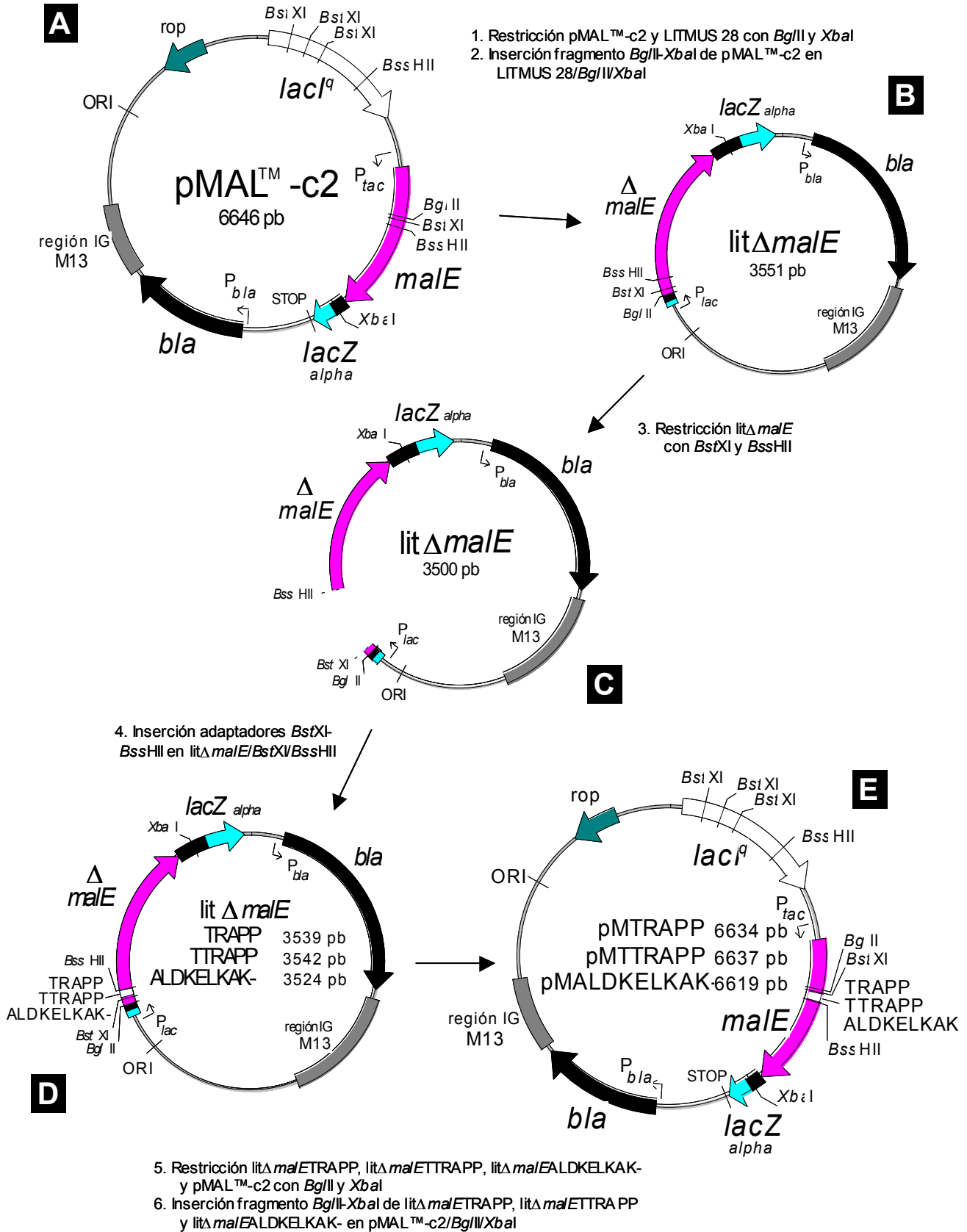
La región de la MBP elegida para llevar a cabo la substitución de una parte de su secuencia de aminoácidos ha sido la hélice- $\alpha$  IV. Esta estructura se halla expuesta en la superficie del dominio C de la MBP, está formada por la secuencia de aminoácidos PALDKELK entre los residuos P<sub>133</sub> y K<sub>140</sub>, y conecta dos hojas  $\beta$  antiparalelas ( $\beta^E$  y  $\beta^F$ ) (Fig. 28B). La elección de esta hélice- $\alpha$  se basa en los resultados descritos en la bibliografía sobre la caracterización de dominios funcionales de la MBP para su exportación al periplasma, la capacidad de asimilación de maltosa y maltodextrinas, o el transporte y unión de maltosa (Duplay *et al.*, 1987; Duplay & Szelcman, 1987). Estos estudios se llevaron a cabo mediante mutagénesis en el gen *malE* por la inserción al azar de oligonucleótidos *Bam*HI. Los diferentes mutantes MBP obtenidos lo son por inserción y deleción de secuencias de aminoácidos, de ahí que se les denomine 'indels' ('insertion/deletion'). Los resultados obtenidos de los diferentes 'indels' indicaban que había regiones de la MBP que no comprometían su estabilidad, la unión de sustrato o su exportación. Posteriormente, el grupo de investigación de M. Hofnung localizó el conjunto de 'indels' sobre la estructura cristalográfica recién resuelta de la MBP (Fig. 28A), y halló que la mayoría implicaban estructuras secundarias hélice- $\alpha$  u hoja- $\beta$  de la MBP (Betton *et al.*, 1993). El mutante MBP 'indel' MalE133 presenta la inserción del dipéptido DP y la deleción de la secuencia ALDKELKAK en la hélice- $\alpha$  IV (Fig. 28C). En numerosos trabajos, esta hélice- $\alpha$  de la MBP ha sido diana para la inserción y expresión de péptidos, y para su presentación como antígeno sin por ello afectar la estructura funcional de la MBP (Martineau *et al.*, 1992; Coëffier *et al.*, 2001; Clément *et al.*, 1996).

## III.B.2.A.2.a. Construcción

En este trabajo experimental se han insertado diversos adaptadores en el gen *malE* para substituir parte de la secuencia de aminoácidos de la hélice- $\alpha$  IV de la MBP. Todos los adaptadores utilizados han sido diseñados para ser traducidos en el uso de codón optimizado para *E. coli*. Los primeros adaptadores utilizados han sido adaptadores *Bst*XI-*Bss*III que codifican para las secuencias TRAPP, TTRAPP y para la deleción de la secuencia ALDKELKAK, respectivamente (Tabla V).

TABLA V. Adaptadores *Bst*XI-*Bss*III para la substitución de la secuencia ALDKELKAK de la MBP

Secuencia	Plásmido de expresión
<pre> W E E I P T R A P P G K 5' ctgggaagagatcccgacccgtgcgccgcggttaagag 3' tttgacccttctctagggctgggcaacgcgccattctcgcgc BstXI BssIII </pre>	pMTRAPP
<pre> W E E I P T T R A P P G K 5' ctgggaagagatcccgacccgtgcgccgcggttaagag 3' tttgacccttctctagggctgggcaacgcgccattctcgcgc BstXI BssIII </pre>	pMTTRAPP
<pre> W E E I P G K 5' ctgggaagagatcccgggtaagag 3' tttgacccttctctagggccattctcgcgc BstXI SmaI BssIII </pre>	pMALDKELKAK
<pre> W E E F P A L D K E L R G R G K 5' ctgggaagaattcccggcgtggataaagaactgcccggcggttaagag 3' tttgacccttcttaaggcgccgacattttcttgacgcccggcgccattctcgcgc BstXI EcoRI NotI BssIII </pre>	pMALEcoNot



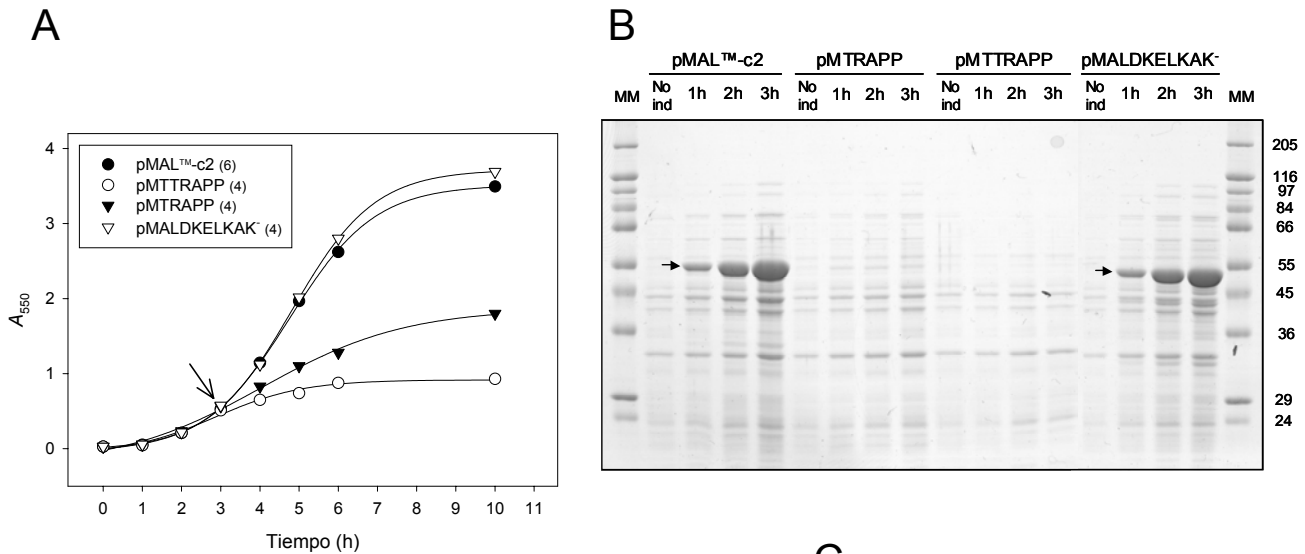
**Figura 33.** Diagrama del proceso de sustitución de la secuencia ALDKELKAK de la MBP (I). **A, B, C, D** y **E** representan la secuencia de estadios diferentes del proceso, ordenada según las operaciones descritas 1. a 6. ORFs: *bla* ( $\beta$ -lactamasa), *lacZ* $\alpha$  (subunidad  $\beta$ -galactosidasa de  $\alpha$ -complementación), *lacI*<sup>q</sup> (represor), *malE* y  $\Delta$ *malE* (MBP); ori (origen replicación bacteriófago M13); *P*<sub>bla</sub>, *P*<sub>lac</sub>, *P*<sub>tac</sub> (promotores); *Bgl*II, *Bss*HII, *Bst*XI, *Xba*I (dianas de restricción).

Para llevar a cabo la inserción de estos adaptadores ha sido necesario subclonar el gen *malE* de manera que las dianas de restricción *Bst*XI y *Bss*HII fueran únicas en el plásmido receptor (Fig. 33A). La digestión del plásmido pMAL<sup>TM</sup>-c2 con *Bgl*II y *Xba*I libera un fragmento de 820 pb codificante para la secuencia de la MBP desde el aminoácido D<sub>120</sub>. Este fragmento se ha purificado e insertado en el plásmido LITMUS 28 previamente linearizado con *Bgl*II y *Xba*I (Fig. 33B). El plásmido obtenido se ha llamado *lit*Δ*malE*, y las únicas dianas *Bst*XI y *Bss*HII que contiene son las del gen *malE* insertado. El plásmido *lit*Δ*malE* digerido con *Bst*XI y *Bss*HII (Fig. 33C) libera un fragmento de 51 pb que ha sido substituido por los diferentes adaptadores *Bst*XI-*Bss*HII. Sobre el fragmento de 3500 pb correspondiente al plásmido *lit*Δ*malE* linearizado, se han insertado los adaptadores *Bst*XI-*Bss*HII codificantes para la delección de ALDKELKAK, para la secuencia TRAPP o para TTRAPP (Tabla V). Los plásmidos resultantes se han llamado *lit*Δ*malE*ALDKELKAK<sup>-</sup>, *lit*Δ*malE*TRAPP y *lit*Δ*malE*TTRAPP, respectivamente (Fig. 33D). Estos plásmidos han sido digeridos posteriormente con *Bgl*II y *Xba*I para recuperar el fragmento de DNA correspondiente al gen *malE* con los adaptadores. Estos fragmentos se han purificado e insertado de nuevo en el plásmido pMAL<sup>TM</sup>-c2 linearizado con *Bgl*II y *Xba*I. Los plásmidos obtenidos se han llamado pMTRAPP, pMTTRAPP y pMALDKELKAK<sup>-</sup> (Fig. 33E), y con ellos se han transformado células competentes XL1-Blue para ensayar la expresión de las distintas MBP mutantes citoplasmáticas. La MBP codificada en los plásmidos pMTRAPP y pMTTRAPP substituye la secuencia de aminoácidos ALDKELKAK por TRAPP y TTRAPP, respectivamente, con la consiguiente pérdida neta de 4 y 3 aminoácidos respecto a la secuencia nativa de la MBP.

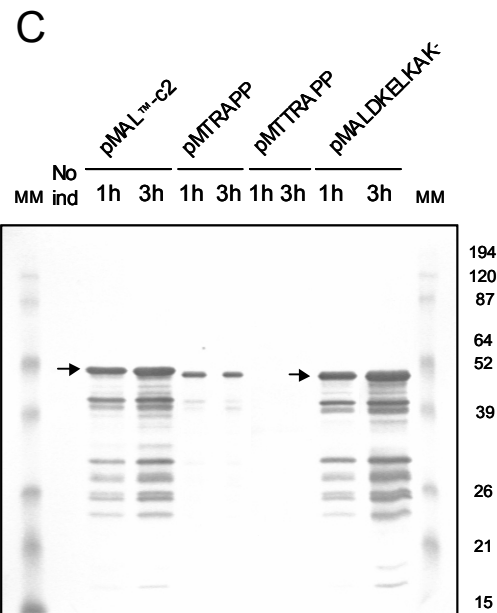
#### III.B.2.A.2.b. Expresión

La cinética de crecimiento de los cultivos de expresión pMTRAPP, pMTTRAPP y pMALDKELKAK<sup>-</sup> muestra que, tras la adición del inductor, el cultivo pMALDKELKAK<sup>-</sup> alcanza niveles de  $A_{550}$  equivalentes al cultivo control pMAL<sup>TM</sup>-c2 (Fig. 34A). Por el contrario, los cultivos pMTRAPP y pMTTRAPP muestran un claro efecto de inhibición post-inducción del crecimiento con respecto al cultivo control. A las dos horas post-inducción, todavía en la fase exponencial del crecimiento, los cultivos pMTRAPP y pMTTRAPP alcanzan un 55% y 38%, respectivamente, de la  $A_{550}$  del cultivo control. En la fase estacionaria del crecimiento, el cultivo pMTTRAPP alcanza únicamente un 25% del valor de  $A_{550}$  del cultivo control.

El análisis mediante SDS-PAGE del extracto crudo de los cultivos de expresión muestra la presencia de bandas intensas únicamente en los cultivos control y pMALDKELKAK<sup>-</sup> (Fig. 34B). La migración electroforética de estas bandas se corresponde con la esperada para la proteína codificada en cada caso, y refleja, respectivamente, la acumulación de la MBP nativa y de la MBP con la hélice- $\alpha$  IV delecionada. Por el contrario, no se aprecia la acumulación de la MBP mutada en los cultivos pMTRAPP ni pMTTRAPP. Mediante una transferencia 'Western', y utilizando suero policlonal anti-MBP, se confirma que la proteína acumulada en los cultivos control y pMALDKELKAK<sup>-</sup> corresponde a la MBP (Fig. 34C). También mediante transferencia 'Western', y sólo a partir de muestras sobrecargadas, se detecta la expresión de la MBP que contiene la secuencia TRAPP en el cultivo de expresión pMTRAPP. Por contra, no se detecta la expresión de la MBP portadora de la secuencia TTRAPP en el cultivo de expresión pMTTRAPP. Por otro lado, se observa de nuevo la acumulación de los intermediarios de síntesis, o bien una notable proteólisis, en aquellas muestras de los cultivos inducidos que acumulan una mayor cantidad de proteína recombinante.



**Figura 34.** Expresión en *E. coli* de la secuencia TRAPP. **Substitución** de la secuencia ALDKELKAK de la MBP (I). **(A)** Cinética de crecimiento de los clones de *E. coli* **XL1-Blue** que codifican para las sustituciones TRAPP y TTRAPP y para la delección de ALDKELKAK. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. La flecha señala el momento de inducción de los cultivos. El clon pMAL™-c2 constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo a 1, 2 y 3 horas post-inducción. 'No ind': extracto crudo del cultivo no inducido. '1h', '2h', '3h': extracto crudo del cultivo a 1, 2 y 3 horas post-inducción. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas denotan la acumulación de la MBP expresada. Condiciones de electroforesis: gel acrilamida 12% y 0.8 mm, 12.5-15 mA/gel. **(C)** Análisis sobre transferencia 'Western'. SDS-PAGE en condiciones reductoras de los extractos crudos a 1 y 3 horas post-inducción; gel acrilamida 12% y 0.8 mm, 15 mA. Transferencia: 0.21-0.24 A, 60 min. Inmunodetección: dilución 1:5000 de suero policlonal anti-MBP de conejo (Protein Fusion and Purification System pMAL™-2, New England Biolabs) y dilución 1:4000 de suero policlonal anti-conejo de oveja, marcado con peroxidasa (The Binding Site). Las flechas señalan la banda que corresponde a la MBP íntegra. 'MM': marcador preteñido de peso molecular (SeeBlue™ Pre-Stained Standards, Novex). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM.



Los resultados obtenidos indican de nuevo que la expresión de la secuencia TRAPP en una proteína de *E. coli* conlleva toxicidad para la célula hospedadora. El hecho de que la delección de la hélice- $\alpha$  IV no afecte a la expresión ni a la estabilidad de la MBP confirma que se trata de una estructura que no compromete la viabilidad de la proteína. La inhibición del crecimiento y la no acumulación de la proteína recombinante en los cultivos de expresión pMTRAPP y pMTTRAPP son, en consecuencia, efecto directo de la expresión de TRAPP en la MBP. La toxicidad observada resulta aún mayor cuando se expresa la secuencia TTRAPP substituyendo la hélice- $\alpha$  IV de la MBP. Este resultado contrasta con el obtenido para la inserción de TTRAPP en el lazo D<sub>120-132</sub> de la MBP, y prueba que el grado de toxicidad que implica su expresión depende del punto de inserción en la MBP. La expresión de las secuencias TRAPP y TTRAPP substituyendo la hélice- $\alpha$  IV de la MBP tiene un efecto equivalente al observado durante la expresión de las formas delecionadas Ct y Nt del fragmento A5 y de las fusiones entre ambas. Ello significa que la expresión y el efecto tóxico de las secuencias TRAPP y TTRAPP en uno y otro caso, se deben haber desarrollado en un entorno molecular o marco estructural y conformacional parecido.

Para confirmar que TRAPP es la mínima secuencia responsable de la toxicidad que se observa cuando se induce su expresión en *E. coli*, se han llevado a cabo nuevas sustituciones en la hélice- $\alpha$  IV de la MBP. Los adaptadores utilizados se han diseñado para codificar las secuencias de aminoácidos que flanquean TRAPP en la gE, y que aparecen en algunos productos de delección Ct y Nt tóxicos y no tóxicos del fragmento A5 (Tabla VI). Estos adaptadores codifican para las secuencias HRTTTRA, TTTRAPP, RAPPRRH y APPRRHG, y lo hacen en el uso de codón optimizado para *E. coli*.

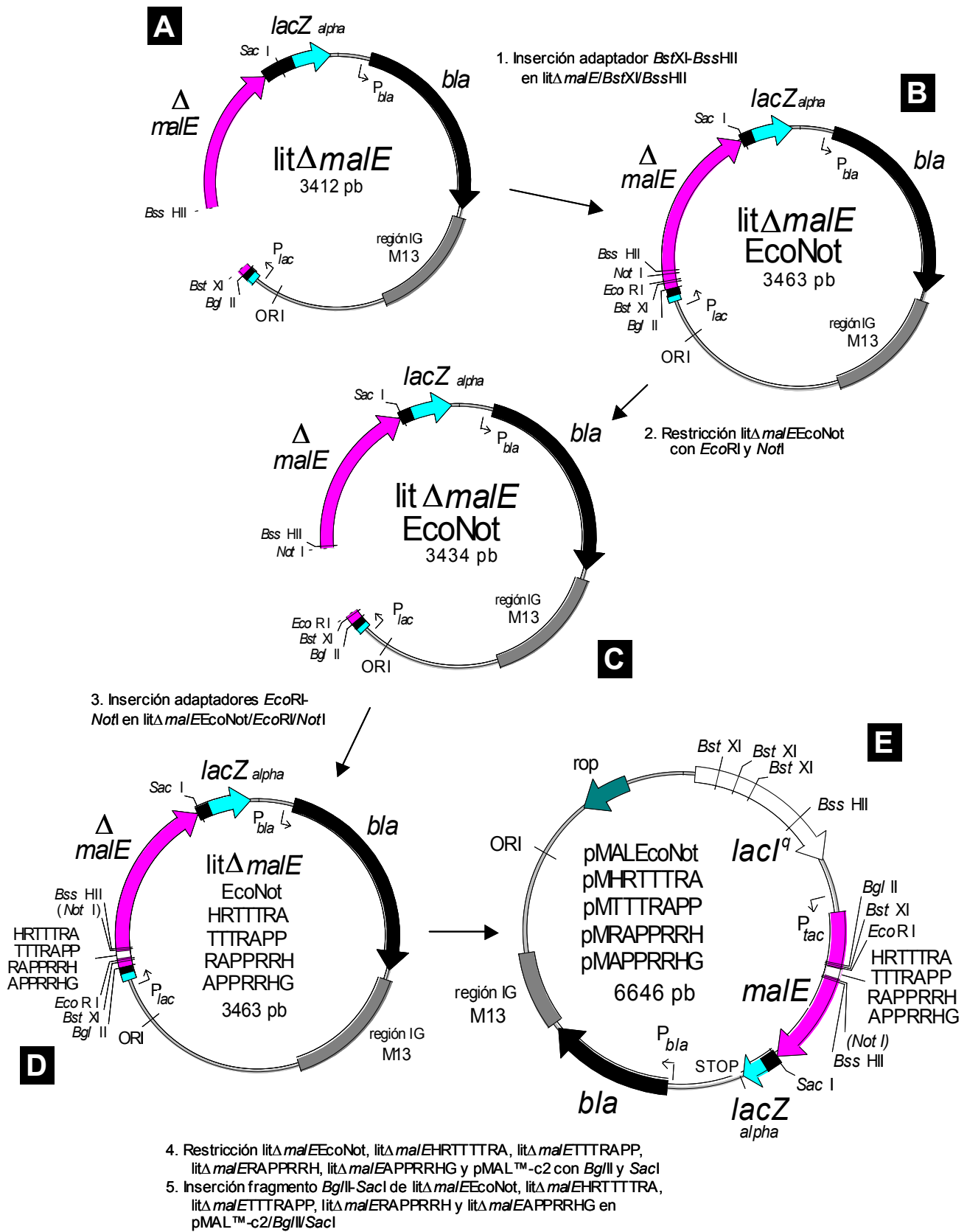
TABLA VI. Adaptadores *EcoRI-NotI* para la sustitución de la secuencia ALDKELKAK de la MBP

Secuencia	Plásmido de expresión
<p style="text-align: center;">F P H R T T T R A</p> <p>5' <b>aattc</b>ccgcatcgtaccaccaccgtgcg 3'</p> <p style="text-align: center;">gggctagcatggtggtggcagc<b>ccgg</b></p> <p><i>EcoRI</i> <span style="float: right;"><i>NotI</i>*</span></p>	pMHRTTTRA
<p style="text-align: center;">F P T T T R A P P</p> <p>5' <b>aattc</b>ccgaccaccaccgtgcgccgccc 3'</p> <p style="text-align: center;">gggctggtggtggcagcggcgg<b>ccgg</b></p> <p><i>EcoRI</i> <span style="float: right;"><i>NotI</i>*</span></p>	pMTTTRAPP
<p style="text-align: center;">F P R A P P R R H</p> <p>5' <b>aattc</b>ccgctgctgcgcgcgcgctcat 3'</p> <p style="text-align: center;">gggctgctgctgctgctgctgct<b>ccgg</b></p> <p><i>EcoRI</i> <span style="float: right;"><i>NotI</i>*</span></p>	pMRAPRRH
<p style="text-align: center;">F P A P P R R H G</p> <p>5' <b>aattc</b>ccggcgccgcccgtgctcatggt 3'</p> <p style="text-align: center;">gggctgctgctgctgctgctgct<b>ccgg</b></p> <p><i>EcoRI</i> <span style="float: right;"><i>NotI</i>*</span></p>	pMAPRRHG

*NotI*\* significa que no se conserva la diana de restricción *NotI* tras la inserción del adaptador.

### III.B.2.A.2.c. Construcción

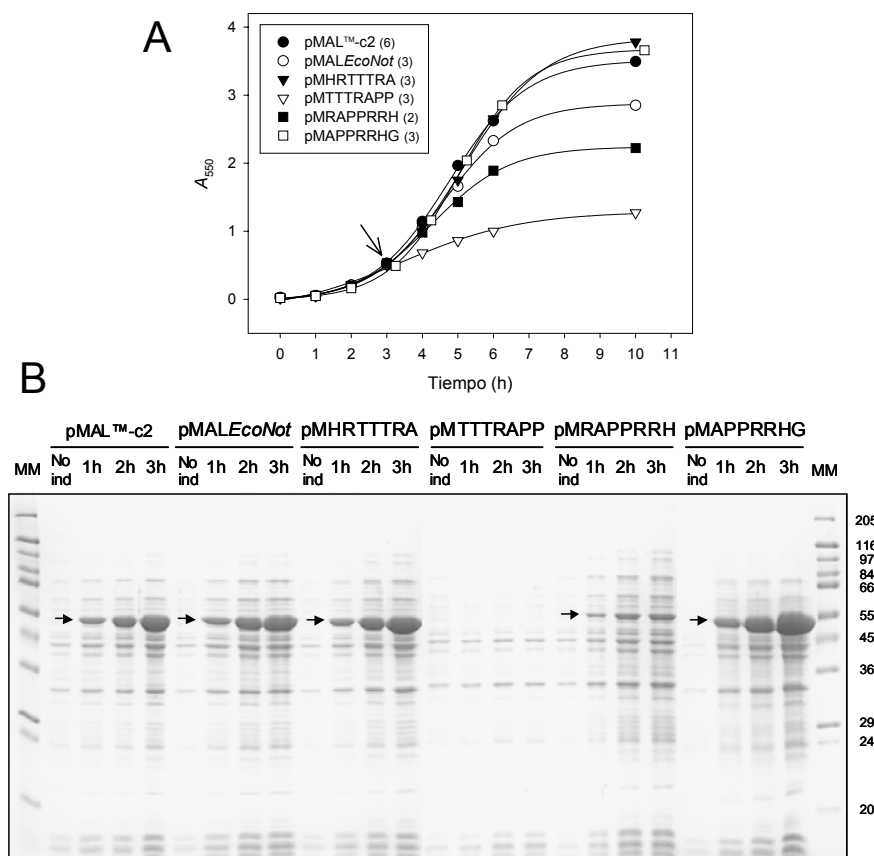
Para insertar los nuevos adaptadores se han creado dos dianas de restricción, *EcoRI* y *NotI*, en la región del gen *malE* que codifica para la hélice- $\alpha$  IV. Estas nuevas dianas se han introducido en el plásmido *lit $\Delta$ malE* mediante un adaptador *BstXI-BssHII* (Tabla V) (Fig. 35A). El plásmido resultante se ha llamado *lit $\Delta$ malEEcoNot* (Fig. 35B). La introducción de las dianas *EcoRI* y *NotI* conlleva una serie de cambios conservativos en la secuencia de aminoácidos de la hélice- $\alpha$  IV: I<sub>132</sub>F, K<sub>140</sub>R, A<sub>141</sub>G y K<sub>142</sub>R. La digestión del plásmido *lit $\Delta$ malEEcoNot* con *EcoRI* y *NotI* (Fig. 35C) libera un fragmento de 29 pb y crea los extremos necesarios para insertar los adaptadores *EcoRI-NotI* HRTTTRA, TTTRAPP, RAPPRRH y APPRRHG. Los plásmidos resultantes se han llamado *lit $\Delta$ malEHRTTTRA*, *lit $\Delta$ malETTTRAPP*, *lit $\Delta$ malERAPRRH* y *lit $\Delta$ malEAPRRHG*, respectivamente (Fig. 35D). Cabe señalar que en estos nuevos plásmidos se conserva la diana de restricción *EcoRI* pero no así la diana *NotI*. La digestión de estos plásmidos con *BglII* y *SacI* libera un fragmento de 750 pb correspondiente al gen *malE* con los adaptadores *EcoRI-NotI*. Estos fragmentos, una vez purificados, se han insertado en el plásmido pMAL<sup>TM</sup>-c2 linealizado con *BglII* y *SacI* (Fig. 35E). Los plásmidos obtenidos se han llamado pMALEcoNot, pMHRTTTRA, pMTTTRAPP, pMRAPRRH y pMAPRRHG, respectivamente, y con ellos se ha transformado células competentes XL1-Blue y BL21(DE3) para ensayar la expresión de la MBP mutagenizada.



**Figura 35.** Diagrama del proceso de sustitución de la secuencia ALDKELKAK de la MBP (II). **A, B, C, D y E** representan la secuencia de estadios diferentes del proceso, ordenada según las operaciones descritas 1. a 5. ORFs: *bla* ( $\beta$ -lactamasa), *lacZ**alpha* (subunidad  $\beta$ -galactosidasa de  $\alpha$ -complementación), *lacI*<sup>q</sup> (represor), *malE* y  $\Delta$ *malE* (MBP); ori (origen replicación); región IG M13 (origen replicación bacteriófago M13); *P*<sub>*bla*</sub>, *P*<sub>*lac*</sub>, *P*<sub>*tac*</sub> (promotores); *Bgl*II, *Bss*HII, *Bst*XI, *Eco*RI, *Not*I, *Sac*I (dianas de restricción).

## III.B.2.A.2.d. Expresión

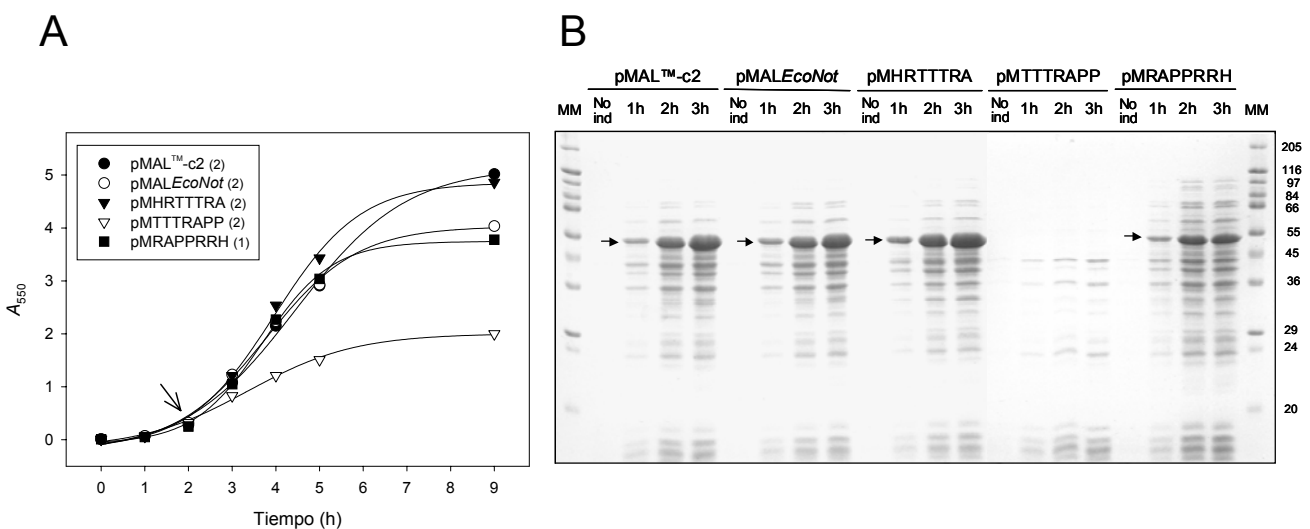
La cinética de crecimiento del cultivo XL1-Blue pMALEcoNot indica que, tras inducir la expresión, el cultivo no alcanza los valores de  $A_{550}$  que registra el cultivo control XL1-Blue pMAL<sup>TM</sup>-c2 (Fig. 36A). Por otro lado, el análisis de expresión por SDS-PAGE muestra que en ambos cultivos se acumula una cantidad comparable de MBP nativa y de MBP(EcoNot), respectivamente (Fig. 36B). A pesar del ligero efecto inhibitorio que tiene la expresión de la MBP(EcoNot) sobre el crecimiento del cultivo, consideramos que esta molécula constituye un control apropiado de expresión no tóxica para valorar la toxicidad de las secuencias adyacentes a TRAPP en la gE. En los cultivos de expresión XL1-Blue pMHRTTTRA y XL1-Blue pMAPRRHG no se observa inhibición post-inducción del crecimiento cuando se compara con los controles XL1-Blue pMAL<sup>TM</sup>-c2 y XL1-Blue pMALEcoNot (Fig. 36A). Por el contrario, los cultivos XL1-Blue pMTTTRAPP y XL1-Blue pMRAPRRH sí muestran inhibición post-inducción del crecimiento. Tras dos horas de inducción, el cultivo XL1-Blue pMRAPRRH registra valores de  $A_{550}$  un 17% inferiores al del cultivo control XL1-Blue pMALEcoNot, mientras que el cultivo XL1-Blue pMTTTRAPP registra valores hasta un 50% inferiores. El análisis de expresión por SDS-PAGE indica que en los cultivos inducidos XL1-Blue pMHRTTTRA y XL1-Blue pMAPRRHG se acumula una cantidad de proteína recombinante comparable, y en ocasiones superior, a la de los cultivos control (Fig. 36B). Cabe señalar, no obstante, que a diferencia de lo que ocurre en los cultivos control, esta acumulación de proteína es insoluble y en forma de cuerpos de inclusión. En el cultivo inducido XL1-Blue pMRAPRRH, la proteína acumulada también forma cuerpos de inclusión y alcanza sólo un 30% de la cantidad acumulada en los cultivos control. En el extracto crudo del cultivo inducido XL1-Blue pMTTTRAPP no se detecta ninguna banda que denote la acumulación de la proteína recombinante.



**Figura 36.** Expresión en *E. coli* de la secuencia TRAPP. **Substitución** de la secuencia ALDKELKAK de la MBP (II). **(A)** Cinética de crecimiento de los clones de *E. coli* XL1-Blue que codifican para las sustituciones HRTTTRA, TTTTRAPP, RAPPRRH y APPRRHG. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. La flecha señala el momento de inducción de los cultivos. El clon pMAL<sup>TM</sup>-c2 constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo a 1, 2 y 3 horas post-inducción. 'No ind': extracto crudo del cultivo no inducido. '1h', '2h', '3h': extracto crudo del cultivo a 1, 2 y 3 horas post-inducción. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas denotan la acumulación de la MBP expresada. Condiciones de electroforesis: gel acrilamida 12% y 0.8 mm, 10-15 mA/gel.



La expresión de las secuencias HRTTTRA, TTTRAPP, RAPPRRH y APPRRHG substituyendo la hélice- $\alpha$  IV de la MBP también se ha llevado a cabo en células BL21(DE3). Ésta es una cepa B de *E. coli*, deficiente en la proteasa citoplasmática Lon (Studier *et al.*, 1990) y en la proteasa de membrana OmpT. Con estas características, la cepa BL21(DE3) se convierte en un hospedador muy adecuado para comprobar si la no acumulación de proteínas que contienen la secuencia TRAPP se debe o no a una actividad mediada por proteasas. En los cultivos de expresión de células BL21(DE3) transformantes para pMAL<sup>TM</sup>-c2, pMALEcoNot, pMHRTTTRA, pMTTTRAPP o pMRAPRRH se ha obtenido, sin embargo, resultados similares a los obtenidos en los ensayos de expresión en células XL1-Blue (Fig. 37A, B). El cultivo de expresión BL21(DE3) pMRAPRRH presenta una inhibición del crecimiento comparable a la del cultivo control BL21(DE3) pMALEcoNot, aunque acumula un 65% menos de proteína recombinante. Por el contrario, el cultivo de expresión BL21(DE3) pMTTTRAPP crece totalmente inhibido y no acumula proteína recombinante. Estos resultados indican que la no acumulación de una proteína que contiene la secuencia TRAPP no se debe inicialmente a la actividad proteolítica de Lon u OmpT. En referencia a la solubilidad de la proteína expresada, únicamente los cultivos BL21(DE3) pMHRTTTRA y BL21(DE3) pMRAPRRH acumulan la proteína de forma insoluble formando cuerpos de inclusión.



**Figura 37.** Expresión en *E. coli* de la secuencia TRAPP. **Substitución** de la secuencia ALDKELKAK de la MBP (III). **(A)** Cinética de crecimiento de los clones de *E. coli* BL21(DE3) que codifican para las substituciones HRTTTRA, TTTRAPP y RAPPRRH. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. La flecha señala el momento de inducción de los cultivos. El clon pMAL<sup>TM</sup>-c2 constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo a 1, 2 y 3 horas post-inducción. 'No ind': extracto crudo del cultivo no inducido. '1h', '2h', '3h': extracto crudo del cultivo a 1, 2 y 3 horas post-inducción. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas denotan la acumulación de la MBP expresada. Condiciones de electroforesis: gel acrilamida 12% y 0.8 mm, 10-12.5 mA/gel.

Los resultados obtenidos en la expresión de las secuencias HRTTTRA, TTTRAPP, RAPPRRH y APPRRHG substituyendo la hélice- $\alpha$  IV de la MBP en células XL1-Blue y BL21(DE3), indican que la expresión de la secuencia TTTRAPP es la que conlleva una mayor toxicidad. A pesar de que la expresión de la secuencia RAPPRRH también se ha mostrado comparativamente tóxica para el crecimiento de la célula hospedadora, los resultados obtenidos corroboran que la toxicidad no reside en las secuencias que flanquean a TRAPP en la gE. En consecuencia, TRAPP es la secuencia mínima tóxica.

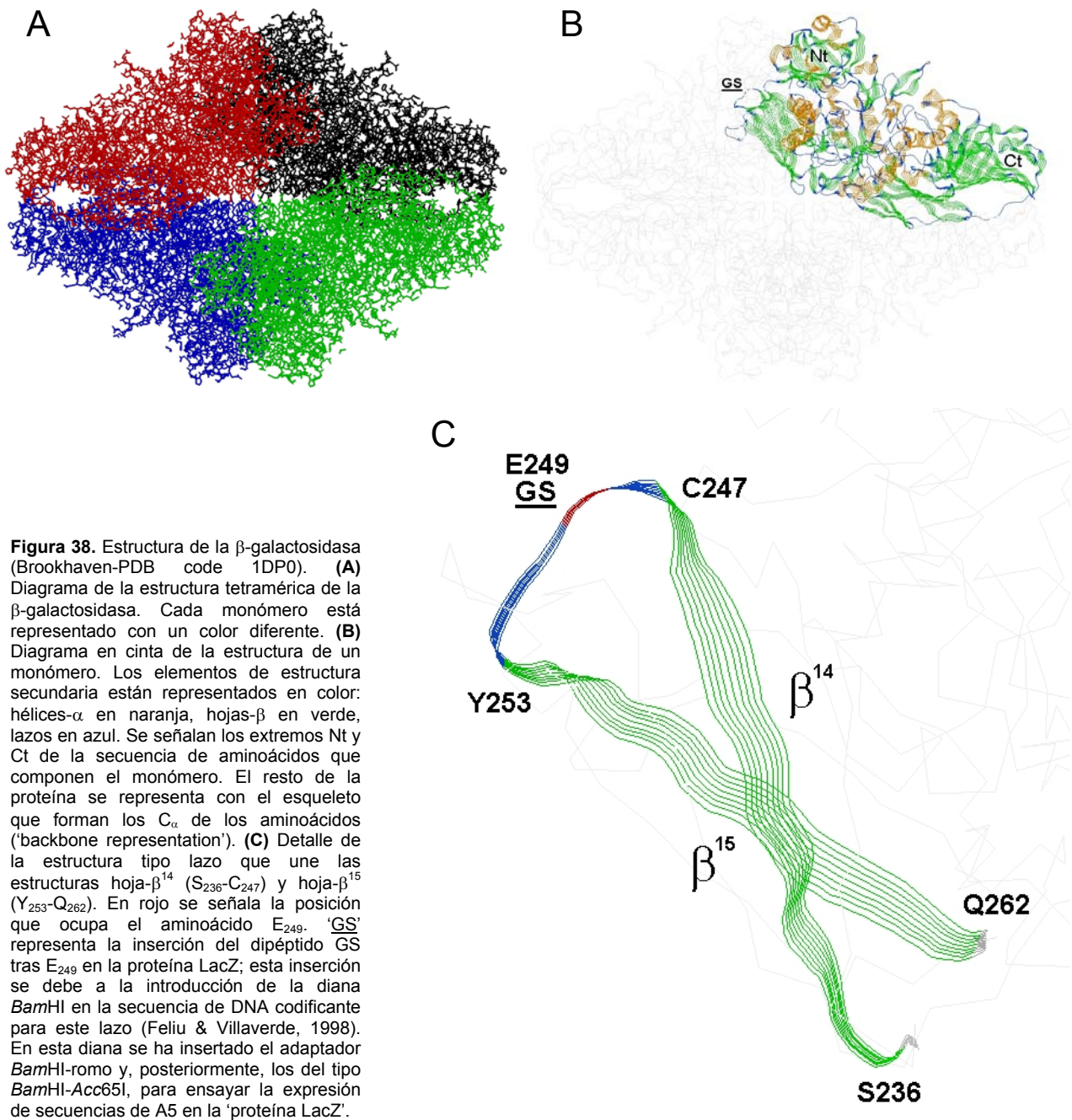
En conjunto, los resultados obtenidos en la expresión de la secuencia TRAPP integrada en la MBP prueban que esta secuencia es tóxica para *E. coli*. Estos resultados apuntan además que la secuencia TRAPP es la secuencia mínima tóxica. No obstante, la diferente repercusión que tiene su inserción en el lazo D<sub>120-132</sub> o su substitución en la hélice- $\alpha$  IV indica que esta toxicidad depende del entorno molecular. El hecho de que la expresión de TRAPP substituyendo la hélice- $\alpha$  IV de la MBP reproduzca la toxicidad observada durante la expresión de las formas delecionadas Nt y Ct del fragmento A5, indica que el efecto tóxico tiene lugar desde un marco estructural y conformacional parecido, probablemente formando parte de un dominio proteico expuesto. En este sentido, el efecto de la secuencia TRAPP integrada en el lazo D<sub>120-132</sub> – sensiblemente menos expuesto en la MBP– debe resultar menos comprometedor para el cultivo de expresión. Por otro lado, la toxicidad y la no acumulación de proteína que conlleva la expresión de la secuencia TRAPP no está en principio relacionada con la actividad de las proteasas Lon u OmpT.

### III.B.2.B. Expresión de la secuencia TRAPP en la $\beta$ -galactosidasa

En este apartado se describen y discuten los resultados obtenidos en la construcción y expresión de la secuencia TRAPP como parte integrante de la  $\beta$ -galactosidasa de *E. coli*. El propósito de los experimentos que se detallan a continuación es corroborar lo observado para la MBP, esto es, que la expresión de la secuencia TRAPP como parte integrante de una proteína de *E. coli* conlleva toxicidad para la célula hospedadora.

La  $\beta$ -galactosidasa es una proteína de elevado peso molecular formada por la unión no covalente de cuatro monómeros iguales codificados en el gen *lacZ* (Fig. 38A, B). El extremo Ct de la  $\beta$ -galactosidasa es poco tolerante a la fusión de proteínas (Corchero *et al.*, 1996; Viaplana *et al.*, 1997a). Por el contrario, el extremo Nt tolera mejor la fusión de otras proteínas o péptidos, y con frecuencia dicha fusión no afecta la actividad enzimática de la quimera (Ullmann, 1992). Por este motivo, la  $\beta$ -galactosidasa se ha utilizado como proteína ‘carrier’ para estabilizar algunas proteínas producto de la expresión heteróloga en *E. coli* (Jacobs *et al.*, 1990; Fuchs *et al.*, 1990). Por otro lado, la resolución de la estructura cristalográfica de la  $\beta$ -galactosidasa (Jacobson *et al.*, 1994), ha permitido identificar algunos dominios permisivos a la inserción, y posterior expresión, de péptidos (Benito *et al.*, 1996; Feliu & Villaverde, 1998). La expresión de una  $\beta$ -galactosidasa que contiene los epítomos P1 y P2 de la gp41 del virus de la inmunodeficiencia humana (VIH), y su posterior aplicación para la detección de anticuerpos anti-VIH (Ferrer-Miralles *et al.*, 2001), es un claro ejemplo de la idoneidad de esta enzima para la presentación de péptidos.

Como en la MBP, la inserción de péptidos tóxicos del fragmento A5 de la gE en la  $\beta$ -galactosidasa se ha realizado mediante adaptadores. Estos adaptadores se han insertado en el gen *lacZ* codificante para la  $\beta$ -galactosidasa del plásmido pJX249 (Feliu & Villaverde, 1998). Cabe decir que a esta  $\beta$ -galactosidasa se la denomina ‘proteína LacZ’ pues, si se compara con la enzima nativa de *E. coli*, carece de los ocho primeros aminoácidos del extremo Nt. El gen *lacZ* del plásmido pJX249 presenta, además, una diana *Bam*HI en la región codificante para el lazo externo de la proteína que une dos estructuras hoja- $\beta$  antiparalelas ( $\beta^{14}$  y  $\beta^{15}$ ), y que está considerado un dominio permisivo (Feliu & Villaverde, 1998) (Fig. 38C). En este lazo es donde se ha insertado los diferentes péptidos del dominio tóxico de la gE.



El primer adaptador utilizado es del tipo *Bam*HI-romo, codifica para el péptido APPRRHG y ha servido, a su vez, para introducir una nueva diana *Acc*65I (Tabla VII). Posteriormente, se ha utilizado las dianas *Bam*HI y *Acc*65I para insertar los adaptadores *Bam*HI-*Acc*65I codificantes para las secuencias HRTTTRA, TTTTRAPP, TTRAPP, TRAPP y RAPRRRH (Tabla VII). Todos los adaptadores utilizados han sido diseñados para ser traducidos en el uso de codón optimizado para *E. coli*.

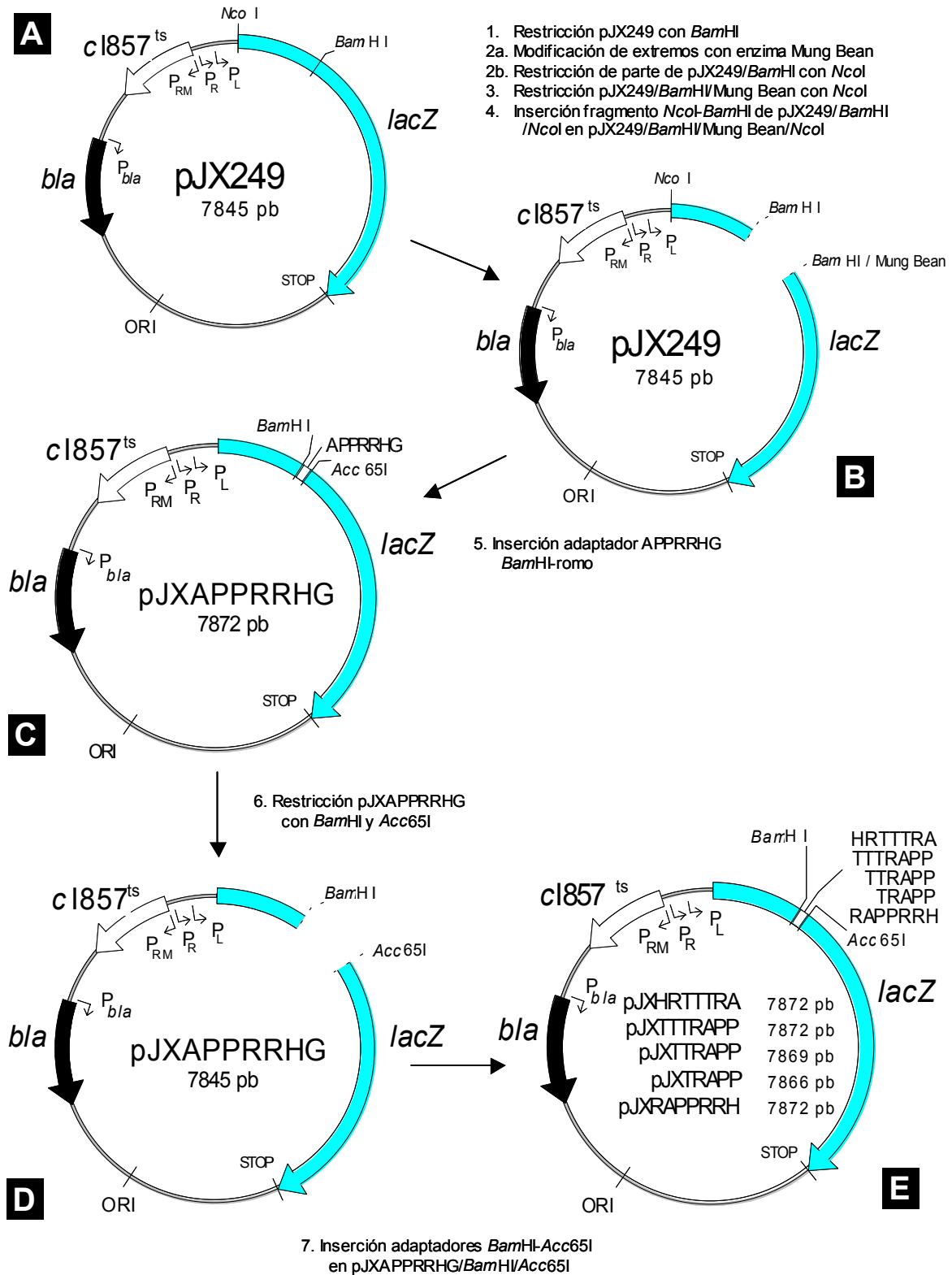
TABLA VII. Adaptadores utilizados para la inserción de secuencias de A5 en la 'proteína LacZ'

Tipo de adaptador	Secuencia	Plásmido de expresión
<i>Bam</i> HI-romo	S A P P R R H G G T 5' <b>gatcc</b> gcgccgcccgcgctcgtcatggt <b>gggtac</b> 3' c ggcggcgccgcagcagtagtacc <b>ccatg</b> g <i>Bam</i> HI romo	pJXAPRRRHG
<i>Bam</i> HI-Acc65I	S H R T T T R A 5' <b>gatccc</b> atcgtaccaccaccgctg <b>cg</b> 3' gtagcatggtggtgggca <b>cgccatg</b> <i>Bam</i> HI Acc65I	pJXHRTTTRA
	S T T T R A P P 5' <b>gatcc</b> accaccaccgctg <b>cgccg</b> 3' gtggtggtgggca <b>cgccatg</b> <i>Bam</i> HI Acc65I	pJXTTTRAPP
	S T T R A P P 5' <b>gatcc</b> accaccgctg <b>cgccg</b> 3' gtggtgggca <b>cgccatg</b> <i>Bam</i> HI Acc65I	pJXTTRAPP
	S T R A P P 5' <b>gatcc</b> accgctg <b>cgccg</b> 3' gtggcagc <b>cgccatg</b> <i>Bam</i> HI Acc65I	pJXTRAPP
	S R A P P R R H 5' <b>gatccc</b> gtgcccgcgctcgtcatg 3' ggcagcggcgccagcagtagtacc <b>ccatg</b> <i>Bam</i> HI Acc65I	pJXRAPRRRH

### III.B.2.B.1.a. Construcción

La digestión del plásmido pJX249 con *Bam*HI lineariza esta molécula de 7845 pb (Fig. 39A). Una fracción del plásmido linearizado se ha sometido a un tratamiento con nucleasa Mung Bean para convertir a romos los extremos *Bam*HI generados. A continuación, el plásmido pJX249 tratado con Mung Bean ha sido digerido con *Nco*I para eliminar el fragmento de 731 pb codificante para los 243 aminoácidos del extremo Nt de la 'proteína LacZ'. La molécula resultante tiene 7114 pb, presenta un extremo romo y otro cohesivo *Nco*I, y codifica para la 'proteína LacZ' a partir del aminoácido L<sub>250</sub>. Paralelamente, otra fracción del plásmido linearizado con *Bam*HI ha sido digerido con *Nco*I para recuperar el fragmento de 731 pb con ambos extremos cohesivos *Nco*I y *Bam*HI. A continuación, se ha llevado a cabo una ligación de tres moléculas: el fragmento *Nco*I-romo de 7114 pb, el fragmento *Nco*I-*Bam*HI de 731 pb y el adaptador *Bam*HI-romo APPRRHG (Fig. 39B,C). El plásmido resultante se ha llamado pJXAPRRRHG, e incorpora una nueva diana de restricción *Acc*65I. Este plásmido codifica para la 'proteína LacZ' con la secuencia GSAPRRRHGGT insertada entre los aminoácidos E<sub>249</sub> y L<sub>250</sub>. Los dipéptidos GS y GT responden a la presencia de las dianas *Bam*HI y *Acc*65I, respectivamente. Finalmente, la digestión del plásmido pJXAPRRRHG con *Bam*HI y *Acc*65I ha permitido insertar los adaptadores *Bam*HI-*Acc*65I HRTTTRA, TTTRAPP, TTRAPP, TRAPP y RPPRRH (Fig. 39D,E). Los plásmidos resultantes se han llamado pJXHRTTTRA, pJXTTTRAPP, pJXTTRAPP, pJXTRAPP y pJXRAPRRRH, y con ellos se ha transformado células BL26 para tratar de expresar la 'proteína LacZ' codificada en cada uno de ellos. La 'proteína LacZ' codificada en pJXTTTRAPP y pJXTRAPP contiene, respectivamente, ocho y siete aminoácidos más que la 'proteína LacZ' codificada en el plásmido pJX249. Los

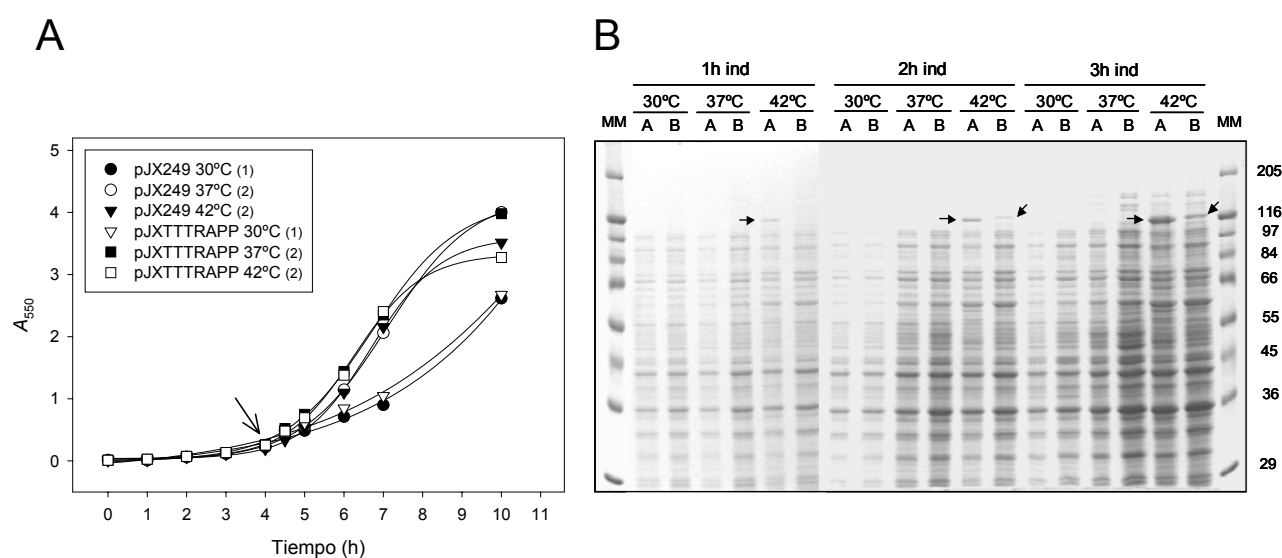
plásmidos pJXAPRRHG, pJXHRTTTRA, pJXTTTRAPP y pJXRAPRRH codifican para una forma de la 'proteína LacZ' que incrementa en nueve el número total de aminoácidos.



**Figura 39.** Diagrama del proceso de inserción de secuencias de A5 en la 'proteína LacZ' (I). **A, B, C, D** y **E** representan la secuencia de estadios diferentes del proceso, ordenada según las operaciones descritas 1. a 7. ORFs: *bla* ( $\beta$ -lactamasa), *lacZ* ('proteína LacZ'), *cI857<sup>ts</sup>* (repressor); ori (origen replicación); *P<sub>bla</sub>*, *P<sub>L</sub>*, *P<sub>R</sub>*, *P<sub>RM</sub>* (promotores); *Acc*65I, *Bam*HI, *Nco*I (dianas de restricción); Mung Bean (exonucleasa).

## III.B.2.B.1.b. Expresión

Inicialmente, se ha llevado a cabo la expresión de la 'proteína LacZ' portadora de la secuencia TTTRAPP en un cultivo de células BL26, y se ha comparado con la expresión de la 'proteína LacZ' nativa. Este ensayo de expresión se ha realizado a tres temperaturas diferentes (30°C, 37°C y 42°C) para tratar de correlacionar la cantidad expresada de la secuencia TTTRAPP y los posibles efectos tóxicos observados. La curva de crecimiento de ambos cultivos registra valores de  $A_{550}$  similares entre sí a cada una de las temperaturas de inducción ensayadas (Fig. 40A). Este resultado indica que la tasa de crecimiento del cultivo no se ha visto afectada por la expresión de TTTRAPP. Sí se observa, en cambio, que la tasa de crecimiento de los cultivos depende de la temperatura de incubación. Así, los cultivos crecidos a 30°C presentan una menor tasa de crecimiento que los cultivos crecidos a 37°C o 42°C. A su vez, el valor de  $A_{550}$  final alcanzado por los cultivos crecidos a 37°C ha sido un 15% mayor que el registrado por los cultivos crecidos a 42°C.



**Figura 40.** Expresión en *E. coli* de la secuencia TRAPP insertada en la 'proteína LacZ' (I). **(A)** Cinética de crecimiento de *E. coli* BL26 que codifica para la inserción TTTRAPP. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon y temperatura. La flecha señala el momento de inducción mediante el cambio de  $T^a$  de los cultivos de 30°C a 37°C y 42°C. El clon pJX249 constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo a 30°C, 37°C y 42°C tras 1, 2 y 3 horas de inducción. 'A': extracto crudo de BL26 pJX249. 'B': extracto crudo de BL26 pJXTTTRAPP. '1h ind', '2h ind', '3h ind': 1, 2 y 3 horas post-inducción. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas denotan la acumulación de la 'proteína LacZ' expresada. Condiciones de electroforesis: gel acrilamida 10% y 0.8 mm, 12.5 mA/gel.

En el análisis por SDS-PAGE de los cultivos inducidos a 30°C no se ha observado ninguna banda que denote la acumulación de la proteína recombinante (Fig. 40B). Este resultado es de esperar pues a esta temperatura el promotor termoinducible  $\lambda P_{RPL}$  de los plásmidos pJX sigue reprimido. Tampoco se aprecia ninguna banda que indique la acumulación de la proteína recombinante en ninguno de los cultivos inducidos a 37°C. Por el contrario, en las muestras de ambos cultivos inducidos a 42°C sí se observa una banda más intensa que presenta un patrón de migración similar al esperado para la 'proteína LacZ'. Cabe subrayar que la acumulación de la 'proteína LacZ' que contiene TTTRAPP se detecta, sin embargo, una hora más tarde que la acumulación de la 'proteína LacZ' nativa. Tras tres horas de inducción a 42°C, la cantidad acumulada

de 'proteína LacZ' portadora de TTTRAPP representa el 20% de la 'proteína LacZ' acumulada en el cultivo control BL26 pJX249.

Los resultados obtenidos en la expresión de la secuencia TTTRAPP insertada en el lazo  $\beta^{14}$ - $\beta^{15}$  de la 'proteína LacZ' indican que no hay inhibición post-inducción del crecimiento del cultivo cuando se compara con el cultivo control BL26 pJX249. Por otro lado, sí hay una acumulación de la 'proteína LacZ' portadora de la secuencia TTTRAPP. Esta cantidad acumulada, aunque menor que la de la 'proteína LacZ' nativa, es considerablemente mayor que la cantidad acumulada de MBP que contiene TTTRAPP (Fig. 36B, 37B). La diferente repercusión que tiene la expresión de la secuencia TTTRAPP en la 'proteína LacZ' y en la MBP, hace sospechar que la activación de algunos sistemas de respuesta 'heat shock' durante la expresión a 42°C haya podido interferir en la toxicidad que implica la expresión de la secuencia TTTRAPP. Ante esta posibilidad, se ha decidido expresar a 37°C los distintos péptidos tóxicos de la gE integrados en la 'proteína LacZ' bajo el control de un promotor inducible por la adición de IPTG. El sistema de expresión escogido ha sido el del promotor  $P_{tac}$  en el plásmido de expresión pMAL-c2Nco.

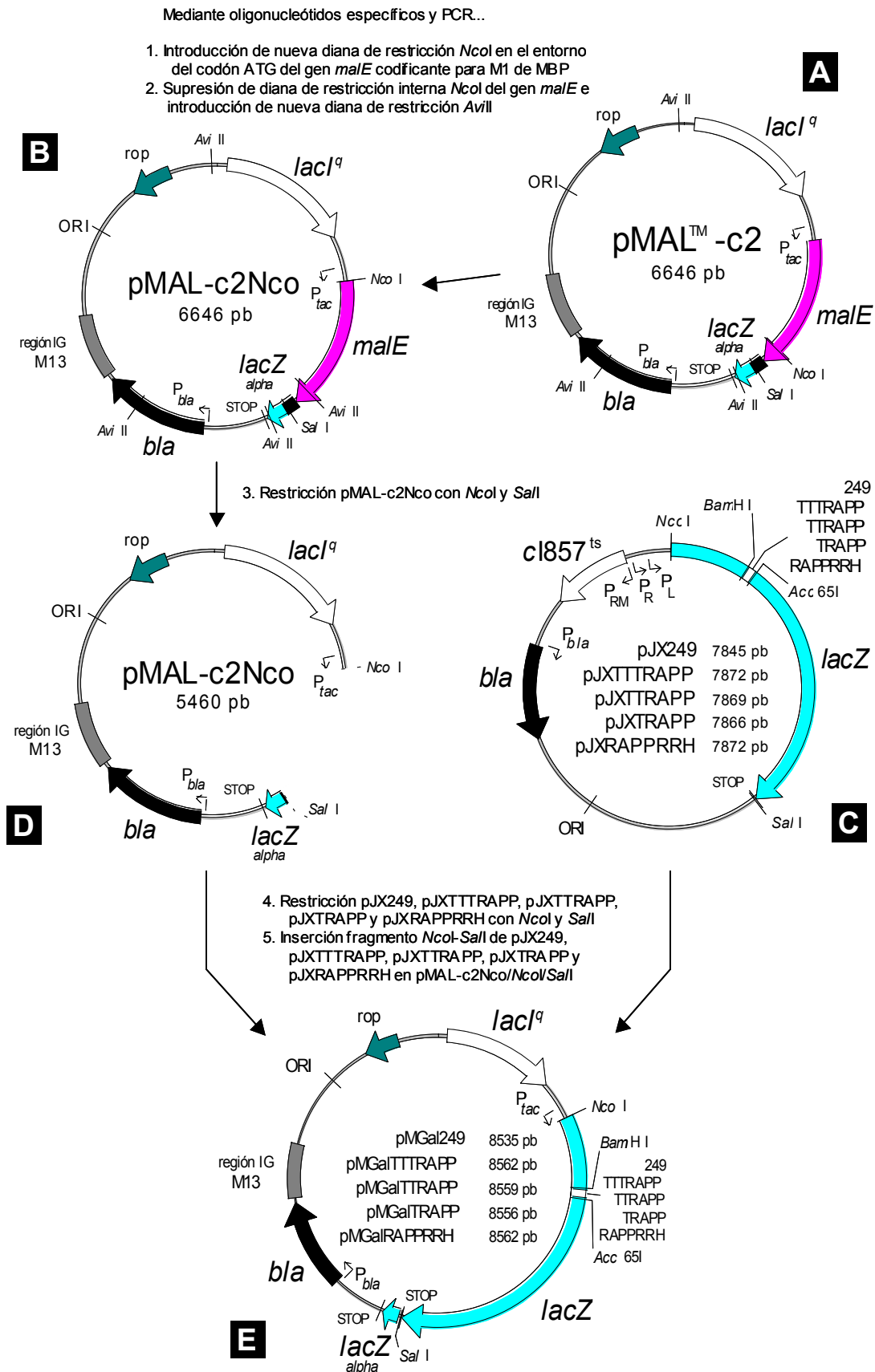
### III.B.2.B.1.c. Construcción

El plásmido pMAL-c2Nco es un plásmido construido expresamente para expresar una secuencia de DNA a partir del promotor  $P_{tac}$ . Este plásmido se ha obtenido mediante mutagénesis dirigida sobre el gen *malE* de pMAL<sup>TM</sup>-c2. Para ello, se ha aplicado la técnica de la PCR y se han utilizado los oligonucleótidos descritos en la Tabla VIII (Fig. 41A,B). El plásmido pMAL-c2Nco presenta una nueva diana de restricción *NcoI* en el codón ATG de inicio de traducción del gen *malE*. La introducción de esta diana conlleva el cambio de aminoácido K<sub>2</sub>E en la MBP codificada. A su vez, se ha eliminado la diana interna *NcoI* del gen *malE* y se ha introducido una nueva diana de restricción *AviI*. Estas dos operaciones no incorporan nuevos cambios en la secuencia de aminoácidos de la MBP.

TABLA VIII. Oligonucleótidos utilizados en la construcción de pMAL-c2Nco

Tarea	Oligonucleótido	Secuencia
Introducción diana <i>NcoI</i>	MBP <i>NcoI</i>	5' ccaacaaggaccatagaccatggaaatcgaagaaggtaaactgg 3' <i>NcoI</i>
	MBPBgIIc	5' ggttcggcagcagatcttgttataaatcagcgataacgc 3' <i>BgII</i>
	MBP <i>NcoI</i> c	5' ccagtttaccttcttcgatttccatggtctatggtccttgttgg 3' <i>NcoI</i>
	MBPEcoRV	5' gcgcggttggtgcggatatctcgg 3' <i>EcoRV</i>
Eliminación diana interna <i>NcoI</i> . Introducción diana <i>AviI</i>	MBP <i>AviI</i>	A A T M E N A Q K G E I M 5' gccgccacgatggaaaaatgcgagaaaggtaaactatgc 3' <i>NcoI</i> * <i>AviI</i>
	pUC/M13 universal	5' gttttccagtcacgac 3'
	MBP <i>AviI</i> c	5' cacctttctgcgcaatcttccatcggtggcggaatacgtgg 3' <i>AviI</i> <i>NcoI</i> *
	MBP <i>NcoI</i>	5' ccaacaaggaccatagaccatggaaatcgaagaaggtaaactgg 3' <i>NcoI</i>

En **negrita** se señalan los cambios de nucleótidos que conllevan la nueva diana. *NcoI*\* significa que se elimina la diana de restricción.



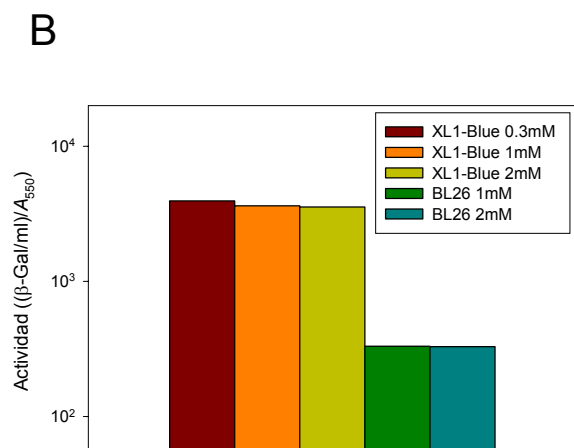
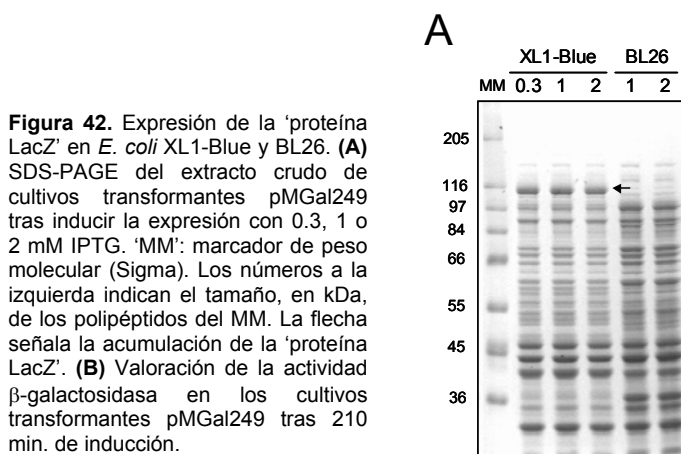
**Figura 41.** Diagrama de la construcción de los plásmidos utilizados para expresar secuencias de A5 en la 'proteína LacZ' a 37°C. **A, B, C, D** y **E** representan la secuencia de estadios diferentes del proceso, ordenada según las operaciones descritas 1. a 5. ORFs: *bla* ( $\beta$ -lactamasa), *lacZ* ('proteína LacZ'), *lacZalpha* (subunidad  $\beta$ -galactosidasa de  $\alpha$ -complementación), *lacI<sup>q</sup>* y *cI857<sup>ts</sup>* (represores), *malE* (MBP); ori (origen replicación); región IG M13 (origen replicación bacteriófago M13); *P<sub>bla</sub>*, *P<sub>L</sub>*, *P<sub>R</sub>*, *P<sub>RM</sub>*, *P<sub>tac</sub>* (promotores); *Avi*II, *Acc65I*, *Bam*HI, *Nco*I, *Sal*I (dianas de restricción).



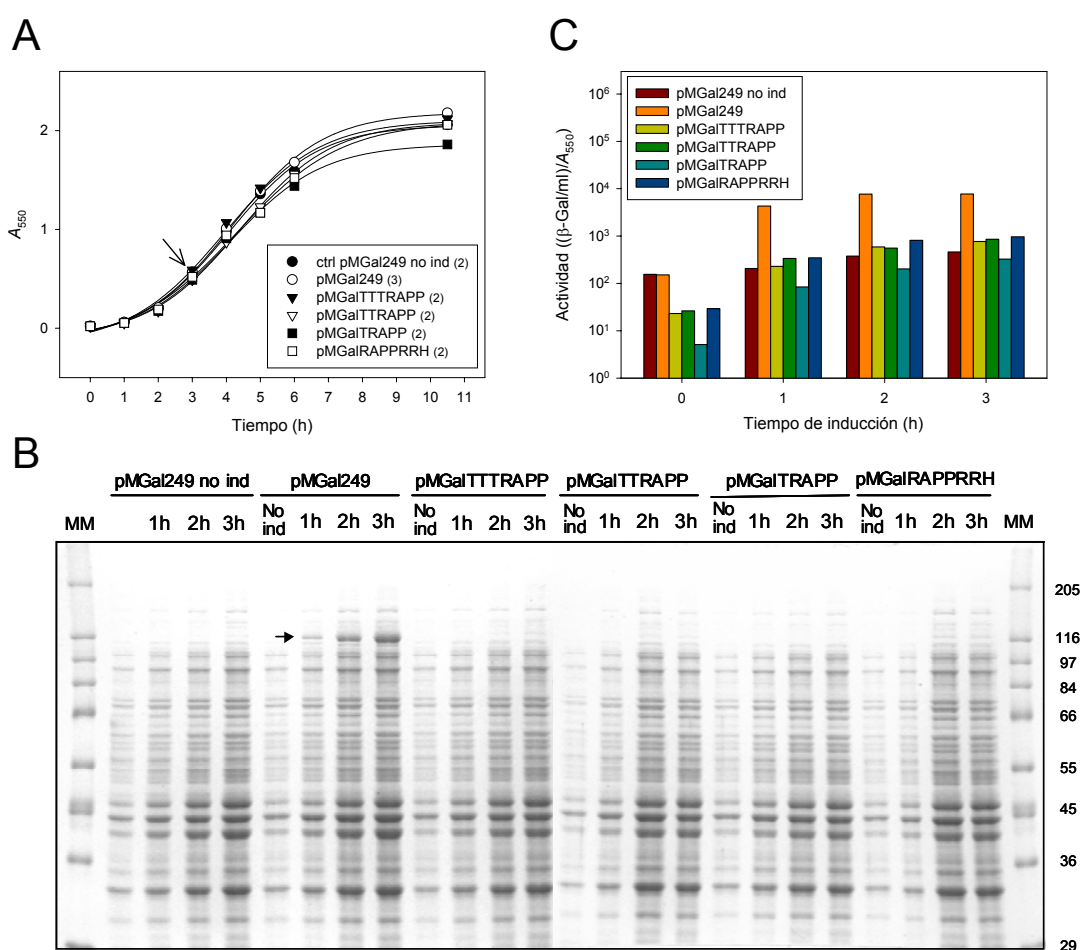
La digestión de los plásmidos pJX249, pJXTTTRAPP, pJXTTRAPP, pJXTRAPP y pJXRAPRRH con *Nco*I y *Sa*II libera un fragmento de 3076-3103 pb, según sea el caso, codificante para la 'proteína LacZ' correspondiente (Fig. 41C). Estos fragmentos, una vez purificados, se han insertado en el plásmido pMAL-c2Nco previamente linearizado con *Nco*I y *Sa*II (Fig. 41D). Los plásmidos resultantes se han llamado pMGal249, pMGalTTTTRAPP, pMGalTTRAPP, pMGalTRAPP y pMGalRAPRRH (Fig. 41E), y con ellos se ha transformado células competentes BL26 y XL1-Blue de *E. coli* para ensayar la expresión de la 'proteína LacZ' a 37°C y a partir del promotor  $P_{tac}$  inducible por IPTG.

### III.B.2.B.1.d. Expresión

Antes de expresar algunos de los péptidos de la gE integrados en el lazo  $\beta^{14}$ - $\beta^{15}$  de la 'proteína LacZ', se ha estudiado la idoneidad de las cepas BL26 y XL1-Blue de *E. coli* para llevar a cabo esta expresión. Para ello, se ha inducido con IPTG la expresión de la 'proteína LacZ' nativa codificada en el plásmido pMGal249 en cultivos de células BL26 y XL1-Blue. El análisis mediante SDS-PAGE del cultivo inducido de células BL26 no muestra acumulación alguna de la 'proteína LacZ'. Este resultado se ha reproducido independientemente de la cantidad de IPTG adicionada al cultivo (Fig. 42A). Por el contrario, el análisis del cultivo inducido XL1-Blue pMGal249 sí revela la presencia de una banda más intensa que presenta un patrón de migración electroforético similar al esperado para la 'proteína LacZ'. Además, la cantidad de 'proteína LacZ' acumulada es la misma independientemente de la concentración de IPTG. Por otro lado, la medida de la actividad  $\beta$ -galactosidasa mediante el método de Miller (Platt *et al.*, 1972) corrobora lo observado en el análisis mediante SDS-PAGE y determina que, en las mismas condiciones de cultivo, las células XL1-Blue pMGal249 expresan hasta diez veces más actividad  $\beta$ -galactosidasa que las células BL26 pMGal249 (Fig. 42B). El hecho que la cepa BL26, pero no así la XL1-Blue, sea una cepa defectiva para las proteasas Lon y OmpT, invita a descartar la idea que cierta actividad proteásica degrade la 'proteína LacZ' acumulada en el cultivo BL26 pMGal249. La no acumulación de 'proteína LacZ' en el cultivo de células BL26 parece hallar una explicación más convincente en su propio fenotipo Lac<sup>-</sup>. Este fenotipo pudiera deberse a una deficiencia del gen *lacY* codificante para la enzima permeasa que no le permite a la célula captar la lactosa del medio de cultivo. De ser así, las células BL26 de un cultivo de expresión también podrían no incorporar el IPTG, análogo de la lactosa, utilizado para iniciar la expresión de la 'proteína LacZ' codificada en el plásmido pMGal249. Independientemente de si esta hipótesis es cierta o no, los resultados obtenidos muestran claramente que la cepa de *E. coli* XL1-Blue, pero no la BL26, es apta para expresar proteínas a partir de un promotor inducible por IPTG como p.e. el  $P_{tac}$ .



Los resultados obtenidos en el cultivo de células XL1-Blue indican que ninguno de los péptidos integrados en la 'proteína LacZ' conlleva inhibición del crecimiento cuando se induce su expresión a 37°C (Fig. 43A). Por otro lado, el análisis mediante SDS-PAGE muestra que únicamente el cultivo control pMGal249 inducido acumula 'proteína LacZ' (Fig. 43B). La valoración de la actividad  $\beta$ -galactosidasa corrobora este resultado, e indica que el cultivo pMGal249 inducido es el cultivo que presenta una mayor actividad a diferentes tiempos de inducción (Fig. 43C). El cultivo pMGal249 inducido presenta en todo momento 17-20 veces más actividad  $\beta$ -galactosidasa que cuando no es inducido. Los cultivos inducidos pMGalTTTRAPP, pMGalTTRAPP y pMGalRAPPRRH registran tan sólo 1-2 veces más actividad  $\beta$ -galactosidasa que el cultivo control no inducido. El cultivo inducido pMGalTRAPP muestra incluso 1.5-2.5 veces menos actividad  $\beta$ -galactosidasa que el cultivo control no inducido. En líneas generales, y a pesar de que sí se observan pequeñas diferencias de actividad  $\beta$ -galactosidasa entre los diferentes cultivos ensayados, los resultados obtenidos no permiten identificar qué péptidos del fragmento A5 insertados en la 'proteína LacZ' y expresados a 37°C son tóxicos para *E. coli*, esto es, inhiben el crecimiento del cultivo toda vez que no permiten la acumulación de la 'proteína LacZ' recombinante.



**Figura 43.** Expresión en *E. coli* de la secuencia TRAPP insertada en la 'proteína LacZ' (II). **(A)** Cinética de crecimiento, a 37°C, de *E. coli* XL1-Blue que codifica para las inserciones TTTRAPP, TTRAPP, TRAPP y RAPPRRH. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. La flecha señala el momento de inducción de los cultivos mediante adición de IPTG. El clon pMGal249 constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo a 1, 2 y 3 horas post-inducción. 'No ind': extracto crudo del cultivo no inducido. '1h', '2h', '3h': extracto crudo del cultivo tras 1, 2 y 3 horas de inducción. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. La flecha denota la acumulación de la 'proteína LacZ' expresada. Condiciones de electroforesis: gel acrilamida 10% y 0.8 mm, 10-12.5 mA/gel. **(C)** Valoración de la actividad  $\beta$ -galactosidasa en los distintos cultivos a 0, 1, 2 y 3 horas post-inducción.

Los resultados obtenidos en la expresión de la secuencia TRAPP insertada en el lazo externo  $\beta^{14}$ - $\beta^{15}$  de la 'proteína LacZ', no corroboran el grado de toxicidad que sí conlleva la expresión de TRAPP en la hélice- $\alpha$  IV, también expuesta, de la MBP. La expresión de TRAPP en la MBP conlleva la inhibición del crecimiento del cultivo y la no acumulación de la MBP, mientras que insertada en la 'proteína LacZ' únicamente conlleva la no acumulación de ésta. Cabe subrayar, sin embargo, que a 42°C sí se acumula la 'proteína LacZ' que contiene la secuencia TRAPP, si bien es cierto que en una cantidad sensiblemente inferior a la de la 'proteína LacZ' nativa. Por otro lado, es posible que a 42°C se hayan activado mecanismos celulares que pueden haber encubierto la posible toxicidad de la secuencia TRAPP. Los resultados obtenidos permiten concluir que la expresión de la secuencia TRAPP es, cuanto menos, contraproducente, y que entraña un cierto grado de toxicidad para el cultivo de expresión.

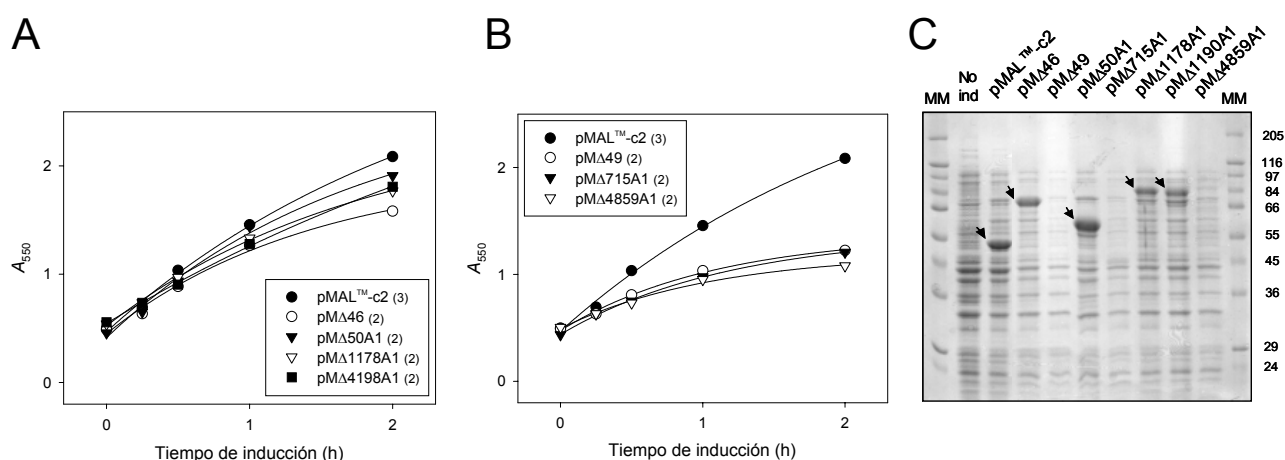
En resumen, los experimentos realizados prueban que la secuencia TRAPP no sólo es tóxica en un contexto que implica la expresión de otros dominios de la gE del BoHV-1, sino que también puede ser tóxica por sí misma, formando parte de proteínas nativas de *E. coli*. La toxicidad inherente a la secuencia TRAPP depende, sin embargo, del entorno molecular en el que se halla, es decir, del marco estructural y conformacional en el que queda integrada la secuencia. Buena prueba de la dependencia del entorno molecular son los diferentes resultados obtenidos durante su expresión cuando se integra en un dominio externo permisivo de la MBP o de la 'proteína LacZ'. Por otro lado, estos mismos resultados ponen de manifiesto que aunque la hélice- $\alpha$  IV de la MBP y el lazo  $\beta^{14}$ - $\beta^{15}$  de la 'proteína LacZ' son ambos dominios externos permisivos, el entorno molecular de uno y otro es marcadamente distinto.

### **III.B.3. Inducción de la respuesta 'heat-shock'**

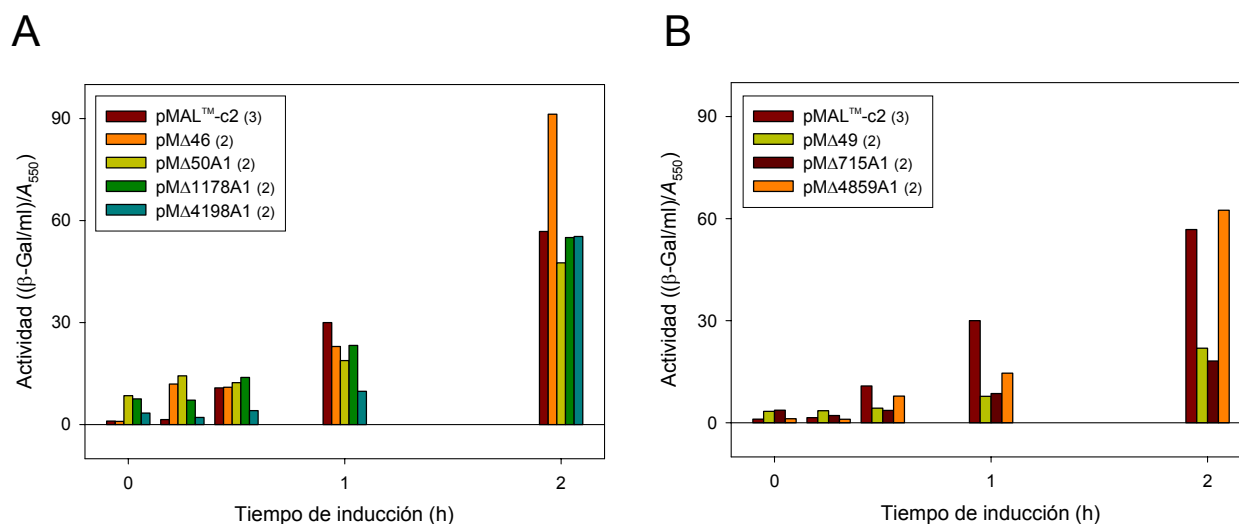
Los resultados hasta ahora expuestos prueban que la expresión de la secuencia TRAPP es, en mayor o menor grado, tóxica en *E. coli*. Cabe pensar que la secuencia TRAPP provoca un cambio de conformación en la proteína capaz de inducir una respuesta 'heat-shock'. Para comprobar esta hipótesis, se ha utilizado la cepa de *E. coli* MC4100 $\lambda$ HS1 (Yano *et al.*, 1987) como hospedadora para la expresión de la secuencia TRAPP contenida en diferentes productos de delección del fragmento A5, y también insertada en la hélice- $\alpha$  IV de la MBP. La cepa MC4100 $\lambda$ HS1 es lisogénica para un bacteriófago derivado de  $\lambda$ pF13 que contiene el promotor termoinducibile *PrpoD<sub>hs</sub>* fusionado al gen *lacZ* ( $\lambda$ pF13-[*PrpoD<sub>hs</sub>*-*lacZ*]). Cuando en la célula MC4100 $\lambda$ HS1 se induce la respuesta 'heat-shock', la RNA polimerasa adquiere, por medio del factor de transcripción  $\sigma^{32}$  (Bukau, 1993; Yura *et al.*, 1993), la especificidad necesaria para transcribir únicamente genes 'heat-shock', o bien aquéllos que como el *lacZ* están fusionados a un promotor 'heat-shock'. La cepa de *E. coli* MC4100 $\lambda$ HS1 permite, pues, medir la intensidad de la transcripción a partir de un promotor 'heat-shock' mediante la valoración de la actividad  $\beta$ -galactosidasa.

Para expresar la secuencia TRAPP en la cepa de *E. coli* MC4100 $\lambda$ HS1, únicamente ha sido necesario transformar dicha cepa con el plásmido de expresión correspondiente. Cuando se induce la expresión de la proteína de fusión en los cultivos MC4100 $\lambda$ HS1 p $\Delta$ 46, p $\Delta$ 50A1, p $\Delta$ 1178A1 y p $\Delta$ 4198A1, todos ellos no codificantes para TRAPP, éstos presentan siempre una ligera inhibición de crecimiento respecto al cultivo control pMAL<sup>TM</sup>-c2 inducido (Fig. 44A), similar a la observada en los cultivos de expresión de células XL1-Blue. Estos cultivos acumulan, además, la proteína de fusión correspondiente (Fig. 44C). Por el contrario, los

cultivos inducidos MC4100 $\lambda$ HS1 pM $\Delta$ 49, pM $\Delta$ 715A1 y pM $\Delta$ 4859A1, todos ellos sí codificantes para TRAPP, presentan siempre un crecimiento notablemente inhibido con respecto al cultivo control (Fig. 44B), y no acumulan la proteína de fusión (Fig. 44C).



**Figura 44.** Expresión de los productos de delección tóxicos y no tóxicos de A5 en *E. coli* MC4100 $\lambda$ HS1 (I). **(A)** Cinética de crecimiento de clones que codifican para productos de delección no tóxicos de A5 fusionados a la MBP. **(B)** Cinética de crecimiento de clones que codifican para productos de delección sí tóxicos de A5 fusionados a la MBP. En **A** y **B**, las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. El clon pMAL<sup>TM</sup>-c2 constituye el cultivo control. **(C)** SDS-PAGE en condiciones reductoras del extracto crudo tras 2 horas de inducción. 'No ind': extracto crudo del cultivo control no inducido. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas denotan la acumulación de la proteína de fusión. Condiciones de electroforesis: gel acrilamida 12% y 0.8 mm, 12.5-15 mA.

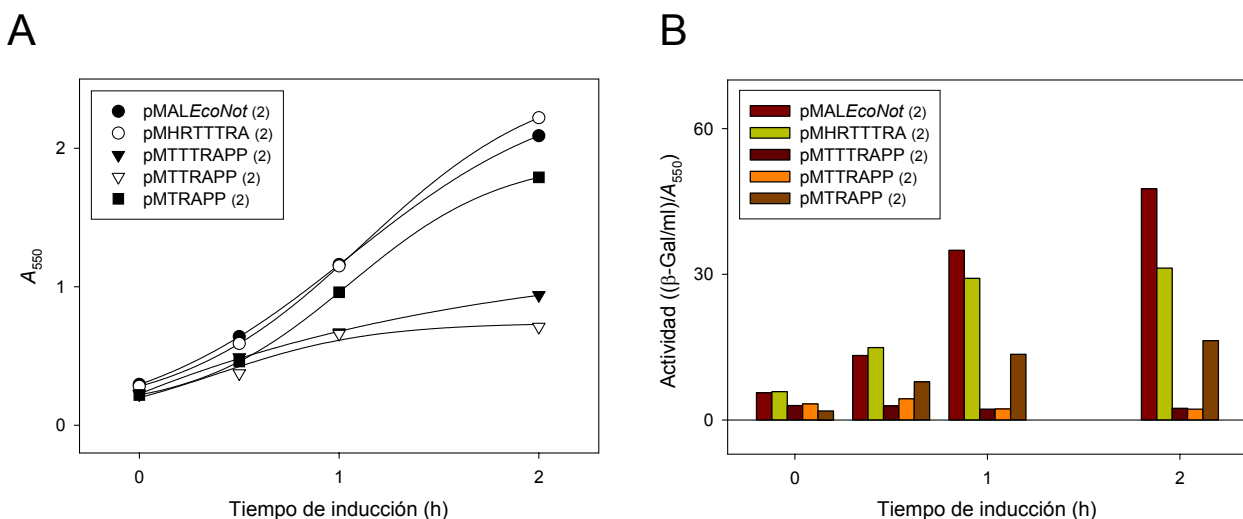


**Figura 45.** Expresión de los productos de delección tóxicos y no tóxicos de A5 en *E. coli* MC4100 $\lambda$ HS1 (II). **(A)** Valoración de la actividad  $\beta$ -galactosidasa a 0, 15, 30, 60 y 120 min. post-inducción en los cultivos que codifican para productos de delección no tóxicos fusionados a la MBP. **(B)** Valoración de la actividad  $\beta$ -galactosidasa en los cultivos que codifican para productos de delección sí tóxicos fusionados a la MBP. En **A** y **B**, la actividad se representa como la media de la actividad valorada en el número de cultivos que se indica entre paréntesis junto a cada clon.

La valoración de la actividad  $\beta$ -galactosidasa indica que, tras 1 hora de inducción, los cultivos que no expresan TRAPP registran un 35-78% de la actividad hallada en el cultivo control (Fig. 45A). En la segunda hora de inducción, los cultivos pM $\Delta$ 46 y pM $\Delta$ 4198A1 experimentan un incremento de la actividad  $\beta$ -galactosidasa, de forma que el cultivo pM $\Delta$ 46 acumula entonces hasta un 60% más de actividad que el

cultivo control pMAL™-c2. Por otro lado, los cultivos que sí expresan TRAPP registran un 26-51% de la actividad  $\beta$ -galactosidasa del cultivo control en la primera hora de inducción (Fig. 45B). Tras 2 horas de inducción, los cultivos pM $\Delta$ 49 y pM $\Delta$ 715A1 registran un 39% y 32%, respectivamente, de la actividad  $\beta$ -galactosidasa del cultivo control, mientras que el cultivo pM $\Delta$ 4859A1 registra un incremento que le lleva a acumular un 6% más de actividad que el cultivo pMAL™-c2.

También se han llevado a cabo cultivos de células MC4100 $\lambda$ HS1 para comprobar si la expresión de secuencias TRAPP substituyendo la hélice- $\alpha$  IV de la MBP inducen o no una respuesta 'heat shock'. Un primer experimento ha constatado que tres cultivos MC4100 $\lambda$ HS1 no inducidos —uno de células sin transformar y los otros dos transformados con pMALEcoNot o pMTTTRAPP—, crecen normalmente, sin inhibición. En todos ellos hay, además, una transcripción basal del promotor *PrpoD<sub>hs</sub>* asociada al crecimiento normal del cultivo. Durante los ensayos de expresión, el cultivo inducido pMHRTTTRA (no codificante para TRAPP) registra una cinética de crecimiento equivalente a la del cultivo control inducido pMALEcoNot (Fig. 46A). Por el contrario, los cultivos inducidos MC4100 $\lambda$ HS1 pMTTTRAPP, pMTTRAPP, y en menor medida pMTRAPP, muestran una inhibición del crecimiento con respecto al cultivo control. La valoración de la actividad  $\beta$ -galactosidasa indica que en los cultivos inducidos pMALEcoNot y pMHRTTTRA, la expresión y acumulación de MBP y MBP portadora de HRTTTRA multiplica por 5 y 3, respectivamente, el nivel basal de transcripción del promotor *PrpoD<sub>hs</sub>* (Fig. 46B). En los cultivos inducidos pMTTTRAPP, pMTTRAPP y pMTRAPP, la actividad  $\beta$ -galactosidasa registrada es en todo momento claramente menor a la del cultivo control inducido pMALEcoNot. Tras 2 horas de inducción, la actividad  $\beta$ -galactosidasa en los cultivos inducidos pMTTTRAPP, pMTTRAPP y pMTRAPP representa, respectivamente, un 5%, 4% y 33% de la actividad hallada en el cultivo control.



**Figura 46.** Expresión de la secuencia TRAPP en *E. coli* MC4100 $\lambda$ HS1 substituyendo la hélice- $\alpha$  IV de la MBP. **(A)** Cinética de crecimiento de clones que codifican para HRTTTRA, TTTTRAPP, TTRAPP y TRAPP en la MBP. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. El clon pMALEcoNot constituye el cultivo control. **(B)** Valoración de la actividad  $\beta$ -galactosidasa en los cultivos a 0, 30, 60 y 120 min. post-inducción. La actividad se representa como la media de la actividad valorada en el número de cultivos que se indica entre paréntesis junto a cada clon.

Los resultados obtenidos indican que la expresión de la secuencia TRAPP, ya sea contenida en los productos de delección Ct o Nt o bien insertada en la hélice- $\alpha$  IV de la MBP, no induce un nivel de transcripción de los promotores de 'heat-shock' mayor que la que induce la expresión de la propia MBP. Los

resultados apuntan a que la activación de la respuesta 'heat-shock' que se observa se debe, antes bien, a la acumulación y solubilidad del producto de expresión. Los cultivos que no acumulan la proteína recombinante son los que generalmente registran una menor actividad 'heat-shock'. Cabe subrayar, sin embargo, que algunos cultivos que no acumulan proteína recombinante como p.e. MC4100 $\lambda$ HS1 pMTRAPP y muy especialmente pM $\Delta$ 4859A1, sí muestran cierta actividad 'heat-shock'. Los resultados parecen indicar que la actividad 'heat-shock' hallada en cultivos tales como MC4100 $\lambda$ HS1 pM $\Delta$ 46, pM $\Delta$ 1178A1 o pM $\Delta$ 4198A1, podría deberse en gran parte a la expresión y acumulación de otras secuencias de la gE, en particular las del fragmento A4. La inducción de expresión de este fragmento, que no su acumulación, también podría ser la causa del incremento de actividad 'heat-shock' en el cultivo MC4100 $\lambda$ HS1 pM $\Delta$ 4859A1. La solubilidad del producto de expresión también está directamente vinculada a la activación de la respuesta 'heat shock'. De hecho, son las proteínas de la respuesta 'heat shock' las que asisten el plegamiento de las proteínas hasta que éstas adquieren la conformación soluble nativa. Este comportamiento explicaría en parte por qué en el cultivo MC4100 $\lambda$ HS1 pMALEcoNot coinciden la expresión soluble de la MBP y una mayor activación de la respuesta 'heat-shock', mientras que en el cultivo MC4100 $\lambda$ HS1 pMHRTTTRA concurren la acumulación de la MBP en cuerpos de inclusión y una activación menor de la respuesta 'heat-shock'.

Los resultados obtenidos de la expresión de la secuencia TRAPP en las células MC4100 $\lambda$ HS1 de *E. coli* corroboran que TRAPP es tóxica, esto es, inhibe el crecimiento del cultivo y conlleva la no acumulación de la proteína expresada. Estos resultados concluyen, además, que el mecanismo por el que la expresión de TRAPP resulta ser tóxica no induce una respuesta 'heat-shock' mediada por un aumento del nivel de transcripción del promotor *PrpoD*. El aumento de transcripción del gen *rpoD*, codificante para el factor de transcripción  $\sigma^{70}$ , está más relacionado con el crecimiento normal de los cultivos y, llegado el caso, con la cantidad y la solubilidad de la proteína expresada acumulada.

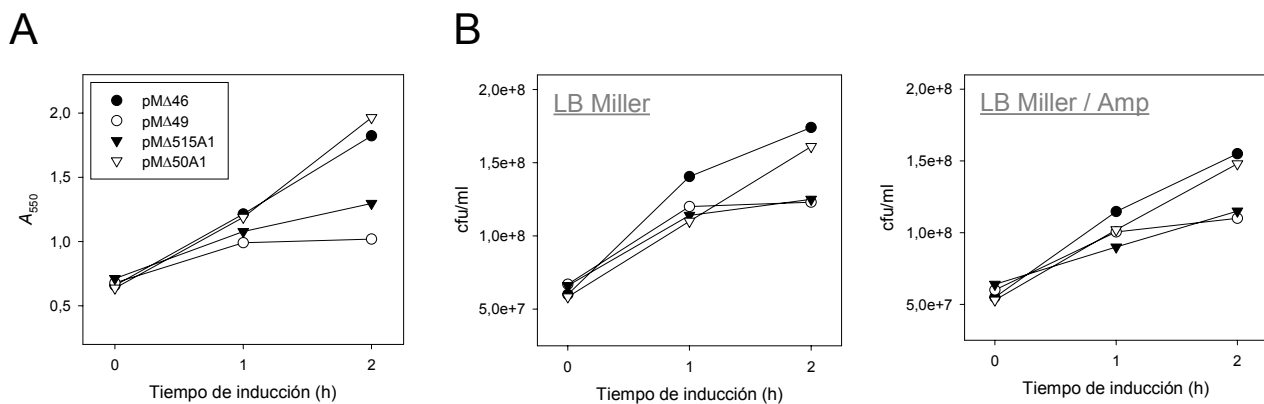
#### III.B.4. Viabilidad del cultivo inhibido

Los resultados expuestos hasta ahora permiten concluir que la expresión de la secuencia de aminoácidos TRAPP es tóxica para *E. coli*. La principal manifestación de esta toxicidad es la inhibición del crecimiento del cultivo de expresión. Asociada a esta inhibición, se evidencia una reducción generalizada de la síntesis proteica en el cultivo y, en particular, del nivel de expresión y de acumulación de la proteína que contiene dicha secuencia.

Para tratar de entender el efecto que la expresión de la secuencia TRAPP tiene sobre el cultivo, se ha estudiado la viabilidad de las células transformantes XL1-Blue pM $\Delta$ 49 y pM $\Delta$ 515A1 –ambas codificantes para TRAPP– y pM $\Delta$ 46 y pM $\Delta$ 50A1 –ambas no codificantes para TRAPP– tras inducir la expresión de la proteína recombinante. Para ello, se ha analizado la  $A_{550}$  y la cantidad de unidades formadoras de colonias (cfu, de 'colony forming units') de los cuatro cultivos a distintos tiempos de inducción. El recuento de cfu se ha efectuado a partir de la siembra de una muestra de los cultivos sobre medio LB Miller y LB Miller/Amp.

Los resultados obtenidos reflejan de nuevo que los cultivos inducidos que expresan TRAPP tienen un crecimiento claramente inhibido si se compara con los cultivos inducidos que no expresan esta secuencia (Fig. 47A). También en el recuento de células viables (Fig. 47B), los cultivos que expresan TRAPP

contabilizan menos cfu/ml que los cultivos que no expresan dicha secuencia. El análisis conjunto de estos resultados sugiere que la cantidad de células viables en los cultivos que expresan TRAPP no es mayor debido a la reducción de la tasa de crecimiento del cultivo. Del mismo modo, en los cultivos que no expresan TRAPP, la mayor tasa de crecimiento implica un mayor número de células viables. Asumiendo que tras la primera hora de inducción todas las células que codifican para TRAPP han iniciado la expresión de dicha secuencia, estos resultados indican que su expresión conlleva la inhibición del crecimiento pero no compromete la viabilidad de las células. La expresión de la secuencia TRAPP tendría, pues, un efecto claramente bacteriostático sobre *E. coli*. En contraposición, el efecto bactericida sería aquél que, inhibiendo el crecimiento celular, disminuiría a su vez la viabilidad celular. Por otro lado, es de destacar la notable coincidencia de recuento de cfu/ml a partir de las siembras sobre medio de cultivo LB Miller y LB Miller/Amp de los cuatro cultivos ensayados. Este resultado sugiere que la pérdida plasmídica en los cultivos inducidos codificantes o no para TRAPP ha sido mínima.



**Figura 47.** Viabilidad de *E. coli* que expresa productos de delección tóxicos de A5. **(A)** Cinética de crecimiento de cultivos inducidos de *E. coli* XL1-Blue que codifican para productos de delección tóxicos o no tóxicos de A5 fusionados a la MBP. Las curvas representadas corresponden a un solo experimento. **(B)** Recuento sobre medio de cultivo LB Miller y LB Miller/Amp de células viables en los cultivos tras 0, 1 y 2 horas de inducción.

### **III.C. La secuencia TRAPP en el proteoma de *E. coli*. Otras secuencias minoritarias**

Los resultados obtenidos en la expresión de los productos de delección de la gE ponen de manifiesto que la expresión de la secuencia de aminoácidos TRAPP conlleva la inhibición del crecimiento de *E. coli* y la no acumulación de la proteína recombinante. Más aún, se ha comprobado que la supresión de dicha secuencia restablece el crecimiento normal del cultivo y la acumulación del producto de expresión. La sospecha de que la secuencia TRAPP es tóxica se ha confirmado tras demostrar que su expresión también inhibe el crecimiento del cultivo y bloquea la acumulación de proteína cuando forma parte de la secuencia de la MBP, especialmente si se expresa expuesta en su superficie. Cabe decir también que no se ha detectado una actividad proteolítica inusual, ni una activación de la respuesta 'heat shock' que denote estrés celular, asociadas a la expresión de TRAPP. Estos resultados convierten a TRAPP, una secuencia de sólo 5 aminoácidos, en una secuencia muy particular que en determinadas circunstancias puede poner en jaque la viabilidad de la célula *E. coli* hospedadora.

La imposibilidad de expresar la secuencia TRAPP de la gE en *E. coli* nos ha llevado a interpretar su no-expresión desde una vertiente probabilística. Se ha analizado la distribución de TRAPP en el proteoma de *E. coli* y en el conjunto de proteínas y marcos abiertos de lectura (ORF, de 'Open Reading Frame') depositados en los bancos de datos informáticos, con la finalidad de justificar su no-expresión por su incompatibilidad natural en *E. coli*. En este apartado se recogen los resultados obtenidos de dichas búsquedas, así como los resultados de expresión de otras secuencias de aminoácidos minoritarias.

#### **III.C.1. La secuencia TRAPP en el proteoma de *E. coli***

Actualmente está secuenciado la totalidad del genoma de *E. coli* (NC\_0000913). En el genoma de la cepa K12 de *E. coli* se ha identificado un total de 4279 ORFs, 2937 de los cuales codifican para polipéptidos o proteínas cuya actividad es conocida, o como mínimo asumida. Los 1342 ORFs restantes son considerados todavía hipotéticos pues para ellos no se ha demostrado la existencia de un producto de expresión.

La búsqueda de la secuencia de aminoácidos TRAPP en los polipéptidos codificados en los 4279 ORFs del genoma de *E. coli* K12, revela que en ninguno de ellos existe dicha secuencia. Sólo la secuencia de aminoácidos FRAPP aparece codificada 1 vez en el conjunto de ORFs del genoma de *E. coli* K12. Este ORF hipotético corresponde al gen *yhfw* en la región intergénica *cysG-trpS*, y codifica para un polipéptido de 408 aminoácidos (44.6 kDa) que puede pertenecer a la familia de las enzimas fosfopentomutasas. La búsqueda de la secuencia TRAPP en un banco de datos no redundante (All non-redundant GenBank CDS translations + PDB + SwissProt + PIR + PRF), indica que dicha secuencia se halla, por lo general, poco representada en el proteoma de especies bacterianas (Tabla IX). Por el contrario, y en comparación a la cantidad de secuencias de origen bacteriano depositadas en los bancos de datos, la secuencia TRAPP se halla codificada con mayor frecuencia en proteínas de animales, vegetales, hongos, virus y también arqueobacterias (Tablas X–XV).



TABLA IX. Proteínas de especies de **eubacterias** que contienen la secuencia TRAPP

Proteína	Especie	Secuencia
<b>Proteínas reguladoras</b>		
gi 27377981 Proteína transportadora ABC de unión de ATP	<i>Bradyrhizobium japonicum</i>	CTRAPP
<b>Proteínas afines a ácidos nucleicos</b>		
gi 10834750 Proteína de replicación	<i>Pseudomonas fluorescens</i>	DTRAPP
<b>Proteínas de secreción</b>		
APA_MYCAV Proteína APA rica en Ala/Pro de unión a fibronectina	<i>Mycobacterium avium</i>	PTRAPP
<b>Proteínas de función desconocida / ORF hipotéticos</b>		
gi 22962256 Proteína hipotética	<i>Rhodopseudomonas palustris</i>	PTAPP
gi 7479312 Proteína hipotética SC10A5.08	<i>Streptomyces coelicolor</i>	PTAPP
gi 21223688 Proteína hipotética	<i>Strep. coelicolor</i>	RTRAPP
gi 23062239 Proteína hipotética	<i>P. fluorescens</i> Pf0-1	RTRAPP
gi 23469238 Proteína hipotética	<i>P. syringae</i>	YTRAPP
gi 15598986 Proteína hipotética	<i>P. aeruginosa</i>	RTRAPP
gi 25169079 Proteína hipotética	<i>Arthrobacter nicotinovorans</i>	ETRAPP
gi 13475880 Proteína hipotética	<i>Mesorhizobium loti</i>	VTRAPP
gi 27375802 Proteína desconocida blr0691	<i>B. japonicum</i>	ITRAPP
gi 22968370 Proteína hipotética	<i>Rhodospirillum rubrum</i>	VTRAPP
gi 23019736 Proteína hipotética	<i>Thermobifida fusca</i>	VTRAPP
gi 16125001 Proteína conservada hipotética	<i>Caulobacter crescentus</i>	MTRAPP
gi 15808034 Proteína conservada hipotética	<i>Deinococcus radiodurans</i>	GTRAPP

TABLA X. Proteínas de especies de **arqueobacterias** que contienen la secuencia TRAPP

Proteína	Especie	Secuencia
<b>Proteínas reguladoras</b>		
gi 14600419 Proteína transportadora ABC de unión de ATP	<i>Aeropyrum pernix</i>	VTRAPP
<b>Proteínas de función desconocida / ORF hipotéticos</b>		
ADEC_METTH Probable adenina aminasa	<i>Methanobacterium thermoautotrophicum</i>	GTRAPP
gi 15790932 Proteína Vng2089h	<i>Pyrococcus horikoshii</i>	ETRAPP
gi 14591475 Proteína hipotética	<i>Halobacterium sp</i>	KTRAPP

TABLA XI. Proteínas de especies de **metazoos** que contienen la secuencia TRAPP

Proteína	Especie	Secuencia
<b>Proteínas reguladoras</b>		
CGB3_CHICK Ciclina B3 específica de transición G2/M (mitosis)	<i>Gallus gallus</i>	ATRAPP
ABC2_HUMAN Proteína cassette de unión de ATP ABCA2, subfamilia A, miembro 2	<i>Homo sapiens</i>	GTRAPP
MYOD_CHICK Proteína homóloga a factor miogénico CMD1 de músculo esquelético	<i>Gall. gallus</i>	HTRAPP
gi 4505613 Proteína respuesta apoptosis prostática 4	<i>H. sapiens</i>	LTRAPP
gi 25145902 Proteína titina	<i>Caenorhabditis elegans</i>	FTRAPP
gi 24620453 Proteína titina 1, 2 MDa	<i>C. elegans</i>	FTRAPP
gi 24620454 Proteína titina 2, 2 MDa	<i>C. elegans</i>	FTRAPP
gi 24620456 Proteína titina 1 301 kDa	<i>C. elegans</i>	FTRAPP
gi 24620457 Proteína titina 2 301 kDa	<i>C. elegans</i>	FTRAPP
gi 3746547 Precursor enzima activador de profenol oxidasa	<i>Manduca sexta</i>	PTAPP
gi 17537419 / Proteína nanOS relacionada con NOS-3	<i>C. elegans</i>	QTRAPP
gi 17537421 RIN2 *		
gi 27716633 Interactor 2 de Ras y Rab / Inhibidor JC265 de Ras	<i>H. sapiens, Mus musculus</i>	CTRAPP
gi 27685605 Proteína similar a proteína adaptora kanadapta	<i>Rattus norvegicus</i>	PTAPP
gi 30158554 Proteína similar a factor transcripción 1 ovárico inducible por gonadotropina	<i>R. norvegicus</i>	KTRAPP
gi 12743763 Proteína similar a factor 4 de transcripción de interferón	<i>H. sapiens</i>	STRAPP
gi 14794520 Proteína similar a factor de transcripción de interferón	<i>H. sapiens</i>	STRAPP
gi 27711714 / Proteína FLJ12168 similar a proteína GTPasa	<i>H. sapiens</i>	PTAPP
gi 6683019 / Factor regulador de complemento GPI	<i>R. norvegicus</i>	GTRAPP
gi 6683021 /		
gi 3873189		
gi 15150298 Receptor adrenérgico alpha2 a2	<i>Takifugu rubripes</i>	RTRAPP
gi 15822701 Receptor adrenérgico alpha2 a2	<i>Danio rerio</i>	RTRAPP
<b>Proteínas afines a ácidos nucleicos</b>		
gi 24657996 Factor GEF64C de intercambio de guanina	<i>Drosophila melanogaster</i>	RTRAPP
gi 17978539		
gi 25151886 Factor Lbc de intercambio de guanina	<i>H. sapiens</i>	VTRAPP
gi 27715485 Proteína similar a proteína de unión a DNA	<i>R. norvegicus</i>	STRAPP

TABLA XI. (cont.)

Proteína		Especie	Secuencia
<b>Proteínas de membrana</b>			
CYA5_*	Adenilato ciclasa tipo V	<i>H. sapiens</i> , <i>Oryctolagus cuniculus</i> , <i>R. norvegicus</i> , <i>M. musculus</i>	STRAPP
YYO1_CAEEL	Receptor 7 transmembrana asociado a proteína G	<i>C. elegans</i>	KTRAPP
gi 28498995	Receptor 7 transmembrana	<i>M. musculus</i>	STRAPP
gi 14550412	Proteína cassette de unión de ATP ABC	<i>H. sapiens</i>	GTRAPP
gi 27661848	Proteína similar a serín proteasa de membrana	<i>R. norvegicus</i>	STRAPP
<b>Proteínas de metabolismo</b>			
C11A_HORSE	Precursor mitocondrial de citocromo P450 11A1	<i>Equus caballus</i>	ITRAPP
G3PT_HUMAN	Gliceraldehído-3-fosfato deshidrogenasa (GADPH-2)	<i>H. sapiens</i>	VTTRAPP
gi 2282013	Proteína 'GADPH-2-like'	<i>H. sapiens</i>	VTRAPP
gi 27659002	Proteína similar a $\beta$ -1,3-galactosiltransferasa	<i>R. norvegicus</i>	ITRAPP
gi 6678716	Proteína 5 de receptor de lipoproteínas de baja densidad	<i>M. musculus</i>	PTTRAPP
gi 15030228	Proteína relacionada con receptor de lipoproteínas de baja densidad	<i>M. musculus</i>	PTTRAPP
gi 3831750	Receptor LR3 de LDL	<i>M. musculus</i>	PTTRAPP
gi 19702209	Proteína similar a transglutaminasa	<i>Paracentrotus lividus</i>	VTRAPP
<b>Proteínas de citoesqueleto</b>			
DYN3_*	Dinamina 3	<i>H. sapiens</i> , <i>R. norvegicus</i>	PTTRAPP
gi 11385416	Proteína Ser/Thr kinasa específica músculo estriado (SPEG)	<i>M. musculus</i>	PTTRAPP
gi 27685249	Proteína similar a SPEG	<i>R. norvegicus</i>	GTRAPP
gi 23510413	Nexina 26	<i>H. sapiens</i>	GTRAPP
gi 26331226	Proteína desconocida, homóloga a dinamina 3	<i>M. musculus</i>	PTTRAPP
gi 24308091	Proteína KIAA0820, homóloga a dinamina 3	<i>H. sapiens</i>	PTTRAPP
gi 20521748	Proteína KIAA1062; Trasportador ABC	<i>H. sapiens</i>	GTRAPP
gi 28526907	Proteína similar a KIAA0303	<i>M. musculus</i>	ATRAPP
gi 28972708	Proteína mKIAA1297	<i>M. musculus</i>	PTTRAPP
gi 28972415	Proteína mKIAA0820	<i>M. musculus</i>	PTTRAPP
<b>Proteínas de matriz extracelular</b>			
YQ36_CAEEL	Colágeno cuticular COL-41	<i>C. elegans</i>	ITRAPP / ATRAPP / GTRAPP
gi 17535575	Quitinasa	<i>C. elegans</i>	ITRAPP
gi 25229108	Quitinasa	<i>Litopenaeus vannamei</i>	ITRAPP
gi 2541918	Proteína quitinasa Pjchi-3	<i>Marsupenaeus japonicus</i>	ITRAPP
gi 25070420 / gi 16716575	Proteína con dominio de unión extracelular	<i>M. musculus</i>	ITRAPP
gi 27660190	Proteína similar a proteína con dominio de unión extracelular	<i>R. norvegicus</i>	ITRAPP
<b>Proteínas de secreción</b>			
SGS3_DROSI	Precursor de proteína salivar de adhesión SGS-3	<i>Dros. simulans</i>	ITRAPP
gi 280630	Precursor de proteína larval Lgp-1	<i>Dros. virilis</i>	ITRAPP
MU5B_HUMAN	Mucina 5 subtipo B traqueobronquial (mucina MG1)	<i>H. sapiens</i>	LTRAPP
gi 4838459	Lectina Tc-ct-4 tipo C	<i>Toxocara canis</i>	ITRAPP
<b>Proteínas de función desconocida / ORF hipotéticos</b>			
gi 3242148	Proteína Pentaxin-like	<i>Dros. melanogaster</i>	ITRAPP
gi 28486744	Proteína hipotética XP 286716	<i>M. musculus</i>	STRAPP
gi 26330736	Proteína desconocida, posible tirosin kinasa	<i>M. musculus</i>	PTTRAPP
gi 29735997	Proteína hipotética XP 299879	<i>H. sapiens</i>	GTRAPP
gi 29742423	Proteína hipotética XP 295187	<i>H. sapiens</i>	ATRAPP
gi 29737843	Proteína hipotética XP 299049	<i>H. sapiens</i>	PTTRAPP
gi 11359886	Proteína hipotética DKFZp434A1010.1	<i>H. sapiens</i>	GTRAPP
gi 7498954	Proteína hipotética F12F3.2	<i>C. elegans</i>	ITRAPP
gi 26985782	Proteína hipotética B0457.2	<i>C. elegans</i>	ATRAPP
gi 21295951	Proteína gnl ENSANGP00000016743	<i>Anopheles gambiae</i>	NTRAPP
gi 21297974	Proteína gnl ENSANGP00000015806	<i>Anoph. gambiae</i>	PTTRAPP
gi 21292371	Proteína gnl ENSANGP00000019479	<i>Anoph. gambiae</i>	ATRAPP
gi 21292909	Proteína gnl ENSANGP00000015023	<i>Anoph. gambiae</i>	HTRAPP
gi 21302332	Proteína gnl ENSANGP00000014375	<i>Anoph. gambiae</i>	ITRAPP
gi 21298649	Proteína gnl ENSANGP00000004463	<i>Anoph. gambiae</i>	PTTRAPP
gi 21291748	Proteína gnl ENSANGP00000017178	<i>Anoph. gambiae</i>	HTRAPP
gi 21295798	Proteína gnl ENSANGP00000017562	<i>Anoph. gambiae</i>	HTRAPP
gi 27731267	Proteína similar a DNAC RIKEN 1810013P09	<i>R. norvegicus</i>	ITRAPP
gi 27697557	Proteína hipotética XP 241562	<i>R. norvegicus</i>	ATRAPP
gi 27662498	Proteína hipotética XP 243026	<i>R. norvegicus</i>	ITRAPP
gi 27675968	Proteína hipotética XP 228119	<i>R. norvegicus</i>	RTRAPP
gi 27708748	Proteína hipotética XP 241842	<i>R. norvegicus</i>	DTRAPP
gi 27691308	Proteína similar a proteína ribosomal L7a	<i>R. norvegicus</i>	PTTRAPP
gi 27687087	Proteína similar a KIAA0303	<i>R. norvegicus</i>	VTRAPP
gi 26328539	Proteína desconocida	<i>M. musculus</i>	ATRAPP
gi 26348331	Proteína desconocida	<i>M. musculus</i>	PTTRAPP
gi 23956308	Proteína hipotética MGC25719	<i>M. musculus</i>	DTRAPP
gi 20894285	Proteína hipotética XP 156170	<i>M. musculus</i>	VTRAPP
gi 27369922	Proteína DNAC RIKEN 9630020E24	<i>M. musculus</i>	PTTRAPP
gi 28972151	Proteína mKIAA0323	<i>M. musculus</i>	ETRAPP
gi 28491881	Proteína DNAC RIKEN 9630021J17	<i>M. musculus</i>	VTRAPP
gi 15991859	Proteína desconocida MGC18222	<i>H. sapiens</i>	LTRAPP
gi 15079463	Proteína desconocida MGC20185	<i>H. sapiens</i>	ATRAPP
gi 14249680	Proteína hipotética MGC14156	<i>H. sapiens</i>	GTRAPP
gi 30158987	Proteína hipotética XP302837	<i>H. sapiens</i>	PTTRAPP
gi 30157025	Proteína hipotética XP303450	<i>H. sapiens</i>	ATRAPP
TRM7_HUMAN	Proteína TRIM7 / GNIP1 de interacción con glicogenina	<i>H. sapiens</i>	ATRAPP
gi 19353648	Proteína desconocida 9630020E24Rik	<i>H. sapiens</i>	PTTRAPP
gi 20521666	Proteína KIAA0820	<i>H. sapiens</i>	PTTRAPP
gi 7959311	Proteína KIAA1522	<i>H. sapiens</i>	GTRAPP

TABLA XI. (cont.)

Proteína		Especie	Secuencia
<b>Proteínas de función desconocida / ORF hipotéticos (cont.)</b>			
gi 20536900	Proteína similar a KIAA1522	<i>H. sapiens</i>	GTRAPP
gi 14250599	Proteína similar a KIAA1062	<i>H. sapiens</i>	GTRAPP
gi 12053343	Proteína hipotética	<i>H. sapiens</i>	CTRAPP
gi 13375952	Proteína hipotética FLJ12168	<i>H. sapiens</i>	PTRAPP
gi 7513172	Proteína F25965_2 homóloga a N-quimerina	<i>H. sapiens</i>	GTRAPP
gi 21064339	Proteína RE07885p	<i>Dros. melanogaster</i>	<b>TTRAPP</b>
gi 18447519	Proteína RE64176p	<i>Dros. melanogaster</i>	ATRAPP
gi 20129117	Proteína CG12446-PA	<i>Dros. melanogaster</i>	FTRAPP
gi 24648397	Proteína CG17203-PA	<i>Dros. melanogaster</i>	<b>TTRAPP</b>
gi 24585436	Proteína CG31687-PA	<i>Dros. melanogaster</i>	ATRAPP
gi 17137552	Proteína desconocida CG3100-PA	<i>Dros. melanogaster</i>	<b>TTRAPP</b>
gi 24585432	Proteína desconocida CG31688-PA	<i>Dros. melanogaster</i>	ATRAPP
gi 27716627	Proteína similar a proteína nuclear de origen bilateral	<i>R. norvegicus</i>	GTRAPP
gi 28278637	Proteína desconocida MGC55452	<i>Dan. rerio</i>	RTRAPP
gi 28201964	Proteína de membrana eritrocitaria	<i>Dan. rerio</i>	ITRAPP
gi 20302741	Proteína desconocida	<i>Branchiostoma floridae</i>	PTTRAPP

TABLA XII. Proteínas de especies de **euglenozoos** que contienen la secuencia TRAPP

Proteína		Especie	Secuencia
<b>Proteínas de citoesqueleto</b>			
gi 255353	Articulina 80 kDa	<i>Euglena gracilis</i>	QTRAPP
gi 99353	Proteína de citoesqueleto	<i>E. gracilis</i>	QTRAPP
<b>Proteínas de función desconocida / ORF hipotéticos</b>			
gi 29250217	Proteína GLP 554 6542 7018	<i>Giardia lamblia</i>	RTRAPP
gi 29246595	Proteína GLP 675 33860 35197	<i>Giar. lamblia</i>	PTRAPP
gi 29245186	Proteína GLP 122 5076 6419	<i>Giar. lamblia</i>	FTRAPP
gi 21629320	Proteína L6202.7	<i>Leishmania major</i>	PTRAPP
gi 6899672	Proteína L6202.7R	<i>L. major</i>	PTRAPP
gi 21629342 / gi 7105524	Proteína L509.4	<i>L. major</i>	LTRAPP
gi 15022392	Proteasa hipotética L3747.06	<i>L. major</i>	DTRAPP
gi 13751814	Proteasa hipotética P883.19	<i>L. major</i>	GTRAPP
gi 6759112	Posible acetiltransferasa L7535.10	<i>L. major</i>	RTRAPP
gi 7271235	Posible proteína transportadora 7-TM	<i>L. major</i>	<b>TTRAPP</b>

TABLA XIII. Proteínas de especies de **plantas** que contienen la secuencia TRAPP

Proteína		Especie	Secuencia
<b>Proteínas de metabolismo</b>			
gi 22775623	Precursor glutamato sintasa dependiente de ferredoxina	<i>Oryza sativana</i>	ATRAPP
PLAS_ARATH	Precursor isoforma menor de plastocianina	<i>Arabidopsis thaliana</i>	ATRAPP
gi 22326585	Proteína At5g03435.1 similar a antranilato fosforibosiltransferasa	<i>Arab. thaliana</i>	PTRAPP
gi 9188547	$\beta$ -1,4-manosa	<i>Chlorella vulgaris</i>	STRAPP
<b>Proteínas de función desconocida / ORF hipotéticos</b>			
gi 21327950	Proteína hipotética P0489G09.11	<i>Oryza sativa</i>	KTRAPP
gi 15223805	Proteína desconocida At1g76970.1	<i>Arab. thaliana</i>	NTRAPP
gi 21739244	Proteína OSJNBa0072F16.23	<i>O. sativa</i>	VTRAPP
gi 20804473	Proteína OSJNBa000016I09.21	<i>O. sativa</i>	WTRAPP
gi 18461260	Proteína P0672C09.18	<i>O. sativa</i>	LTRAPP
gi 19386688	Proteína P0506A10.3	<i>O. sativa</i>	LTRAPP
gi 28875988	Proteína hipotética	<i>O. sativa</i>	<b>TTRAPP</b>
gi 14626283	Proteína hipotética OJ1124 H03	<i>O. sativa</i>	STRAPP
gi 10140738	Proteína inducible por sal y asociada a membrana	<i>O. sativa</i>	DTRAPP
gi 10178049	Proteína desconocida	<i>Arab. thaliana</i>	STRAPP
gi 18414788	Proteína hipotética At5g05100.1	<i>Arab. thaliana</i>	STRAPP
gi 15241285	Proteína desconocida At5g14020.1	<i>Arab. thaliana</i>	VTRAPP
gi 22326585	Proteína At5g03435.1 similar a antranilato fosforibosiltransferasa	<i>Arab. thaliana</i>	PTTRAPP

TABLA XIV. Proteínas de especies de **hongos** que contienen la secuencia TRAPP

Proteína		Especie	Secuencia
<b>Proteínas afines a ácidos nucleicos</b>			
YRA1_YEAST	Proteína de unión/hibricación a RNA	<i>Saccharomyces cerevisiae</i>	NTRAPP
<b>Proteínas de metabolismo</b>			
gi 11595576	Proteína relacionada con fosfatidilserina decarboxilasa	<i>Neurospora crassa</i>	STRAPP
gi 7107367	Celobiohidrolasa	<i>Trichoderma harzianum</i>	<b>TTRAPP</b>

TABLA XIV. (cont.)

Proteína		Especie	Secuencia
<b>Proteínas de función desconocida / ORF hipotéticos</b>			
gi 6474885	Proteína hipotética	<i>Schizosaccharomyces pombe</i>	KTRAPP
gi 19111918	Proteína similar a cd50c de levadura	<i>Schiz. pombe</i>	KTRAPP
gi 19074455	Proteína con dominio zinc-finger	<i>Encephalitozoon cuniculi</i>	ETRAPP
YN18_YEAST	Proteína hipotética de 118.3 kDa en región intergénica ERG24-MET2 / Factor 120 kDa asociado a CCR4	<i>Sacch. cerevisiae</i>	KTRAPP
gi 28919715	Proteína hipotética	<i>N. crassa</i>	LTRAPP
gi 28925956	Proteína hipotética	<i>N. crassa</i>	WTRAPP
gi 28926388	Proteína hipotética	<i>N. crassa</i>	GTRAPP
gi 28917128	Proteína hipotética	<i>N. crassa</i>	FTRAPP
gi 28923809	Proteína hipotética	<i>N. crassa</i>	RTRAPP
gi 28925772	Proteína desconocida	<i>N. crassa</i>	GTRAPP

TABLA XV. Proteínas de especies de **virus** que contienen la secuencia TRAPP

Proteína		Especie	Secuencia
<b>Proteínas afines a ácidos nucleicos</b>			
VE1_HP4V1	DNA helicasa ATP-dependiente E1	Papilomavirus humano tipo 41 (virus DNAds, no RNA)	QTRAPP
gi 9632271	RNA polimerasa RNA dependiente	Virus de <i>Helicoverpa armigera</i> (virus RNAss+, no DNA)	ITRAPP
gi 8925893	DNA polimerasa	Herpesvirus asociado a fibromatosis retroperitoneal	YTRAPP
gi 6469146	DNA polimerasa	ChRV1 (virus DNAds, no RNA)	YTRAPP
gi 28392844	Posible proteína NTPasa	Calicivirus canino (virus RNAss+, no DNA)	STRAPP
<b>Proteínas de membrana y estructurales</b>			
VGLE_HSVBS	Glicoproteína E	BoHV-1 (virus DNAds, no RNA)	<b>TTRAPP</b>
gi 23216981	Proteína E4	Papilomavirus bovino tipo 3	DTRAPP
<b>Proteínas de función desconocida / ORF hipotéticos</b>			
VG65_HSV11	Proteína hipotética gen 65	IcHV-1 (virus DNAds, no RNA)	HTRAPP
YKR4_EBV	Proteína hipotética BKRF4	HHV-4 (virus DNAds, no RNA)	RTRAPP
gi 628146	Proteína hipotética	HHV-4 (virus DNAds, no RNA)	RTRAPP
gi 10444022 / gi 10444024 / gi 10444026 / gi 10444028	Proteína NS5	Virus Edge Hill (virus RNAss+, no DNA)	DTRAPP
gi 3237334	Poliproteína	Virus de la Hepatitis C (HCV) (virus RNAss+, no DNA)	PTRAPP
gi 24528249	Poliproteína	HCV (virus RNAss+, no DNA)	PTRAPP
gi 24528279	Poliproteína	HCV (virus RNAss+, no DNA)	PTRAPP
gi 11559465	Poliproteína	HCV (virus RNAss+, no DNA)	PTRAPP
gi 5821155	Poliproteína	HCV (virus RNAss+, no DNA)	PTRAPP
gi 3237362	Poliproteína	HCV (virus RNAss+, no DNA)	PTRAPP
gi 4753721	Poliproteína	HCV (virus RNAss+, no DNA)	PTRAPP
gi 5918943	Poliproteína	HCV (virus RNAss+, no DNA)	PTRAPP
gi 28395423	Proteína UL17	SuHV-1 (virus DNAds, no RNA)	VTRAPP
gi 27881466	Poliproteína no estructural	Calicivirus canino (virus RNAss+, no DNA)	STRAPP

Conocer la secuencia de aminoácidos de los 4279 polipéptidos codificados en otros tantos ORFs de *E. coli* permite, por un lado, hallar la frecuencia real de uso de un péptido en dicho conjunto de polipéptidos. Así se ha determinado, por ejemplo, que las secuencias de aminoácidos EPI, RAPP y RAQEL se hallan contenidas 145, 1 y 2 veces, respectivamente, en los 4279 polipéptidos codificados de *E. coli*. Por otro lado, también es posible hallar la frecuencia real con que *E. coli* utiliza cada uno de los aminoácidos en el conjunto de sus proteínas (Tabla XVI). Conocida la frecuencia real de uso de cada aminoácido, se puede calcular la frecuencia esperada (número de casos esperados) de hallar una secuencia de aminoácidos determinada en el proteoma de *E. coli* por azar. Este cálculo se ha realizado considerando cada aminoácido con probabilidad independiente de ocurrir en el conjunto de aminoácidos de un polipéptido. En definitiva, cada posible secuencia de aminoácidos lleva asociada una frecuencia real y una frecuencia esperada de suceder en el conjunto de polipéptidos codificados en *E. coli*, y la comparación de ambas frecuencias entre sí da una idea de la divergencia con que esa secuencia de aminoácidos ocurre con respecto a su probabilidad de existir simplemente por azar.

TABLA XVI. Frecuencia de uso de cada aminoácido en el proteoma de *E. coli*

Aminoácido	Frecuencia de uso (%)	Aminoácido	Frecuencia de uso (%)	Aminoácido	Frecuencia de uso (%)	Aminoácido	Frecuencia de uso (%)
A	9.484	G	7.369	M	2.849	S	5.823
C	1.171	H	2.275	N	3.952	T	5.405
D	5.138	I	6.004	P	4.429	V	7.058
E	5.747	K	4.408	Q	4.429	W	1.527
F	3.901	L	10.643	R	5.538	Y	2.850

La capacidad de cálculo que se requiere para analizar las frecuencias observada y esperada de un conjunto de péptidos está en función de la cantidad de estos péptidos. El conjunto de péptidos de 5 aminoácidos lo constituyen 3200000 posibles combinaciones de aminoácidos (todas las combinaciones posibles de los 20 aminoácidos tomados de 5 en 5, esto es  $20^5$ ), y cada una de ellas tiene una probabilidad de hallarse en el proteoma de *E. coli*. Mucho más reducido es el conjunto de péptidos de 4 aminoácidos, formado por 160000 ( $20^4$ ) posibles combinaciones de aminoácidos. La capacidad de cálculo requerida para el conjunto de  $20^5$  péptidos ha resultado ser prohibitiva, y en consecuencia el estudio se ha limitado al conjunto formado por  $20^4$  péptidos de 4 aminoácidos.

Los resultados obtenidos indican que en los 4279 polipéptidos codificados de *E. coli* están contenidos en total 1348142 péptidos de 4 aminoácidos. En ellos se hallan repetidas 2 o más veces el 79.33% de las  $20^4$  posibles secuencias de 4 aminoácidos, y 144411 (90.26%) de todas ellas se hallan codificadas al menos 1 vez en *E. coli*. Por el contrario, las 15589 (9.74%) combinaciones restantes no están codificadas ninguna vez en *E. coli*. La divergencia entre la distribución de las frecuencias observadas y la distribución de las frecuencias esperadas se ha hallado mediante el índice estadístico ji cuadrado ( $\chi^2$ ) (Doménech Massons & Granero, 1998c). Este índice permite medir la distancia entre las frecuencias observada y esperada de una misma secuencia. El índice  $\chi^2$  pondera el cuadrado de la diferencia entre ambas frecuencias dividiendo por la frecuencia esperada. Cuanto más similares son las frecuencias observada y esperada de una secuencia, más tiende a 0 el valor de  $\chi^2$ . Por el contrario, cuanto mayor son las diferencias entre ambas frecuencias, mayor es el índice  $\chi^2$ . El índice  $\chi^2$  para el conjunto de las  $20^4$  posibles secuencias es la suma de los  $\chi^2$  de cada una de las secuencias, y su valor ( $\chi^2 = 318832.51$ ) indica que la distribución de las frecuencias observadas diverge significativamente ( $P < 0.001$ ) de la distribución de las frecuencias esperadas. El resultado obtenido permite concluir que las secuencias de 4 aminoácidos de los polipéptidos codificados en *E. coli* no se deben a una asociación al azar de los aminoácidos entre sí.

Para intentar explicar la escasa presencia de la secuencia RAPP en el proteoma de *E. coli* se ha dividido el conjunto de las  $20^4$  posibles secuencias, y se ha analizado por separado la distribución del índice  $\chi^2$  de las secuencias que se hallan con una frecuencia mayor o menor a la teóricamente esperada si los aminoácidos se combinaran por azar. De las  $20^4$  posibles secuencias, 68984 (43.12%) se hallan con una frecuencia mayor a la esperada ('obs>esp') (Tabla XVII), mientras que las restantes 91016 (56,88%), que incluyen las 15589 combinaciones no codificadas en *E. coli*, se hallan con una frecuencia menor a la esperada ('obs<esp') (Tabla XVIII, XIX).

TABLA XVII. Secuencias de 4 aminoácidos con una frecuencia **mayor** a la esperada por azar en el proteoma de *E. coli*

Secuencia	Frecuencia observada	Frecuencia esperada	$\chi^2$	Secuencia	Frecuencia observada	Frecuencia esperada	$\chi^2$
CIGC	26	0.8179	775.3277	LDEP	67	18.7647	123.9910
GKST	113	13.7825	714.2467	MKKI	28	4.4808	123.4510
HHHH	12	0.3611	375.1116	SGKT	55	13.7825	123.2636
DEPT	67	9.5295	346.5907	CPKC	7	0.3609	122.1295
TPEQ	60	8.2146	326.4612	CCYD	6	0.2707	121.2592
CGHC	10	0.3099	302.9814	WHPH	8	0.4719	120.0960
GSGK	94	18.7906	301.0252	LSGG	119	45.3695	119.4957
CMQH	12	0.4532	294.2067	CPSC	8	0.4768	118.7156
CHHC	5	0.0957	251.3887	FCCD	7	0.3705	118.6147
HCMQ	11	0.4532	245.4544	YTWC	7	0.3713	118.3258
GKTT	66	12.7931	221.2881	WQNM	12	1.0266	117.2991
HWYN	11	0.5275	207.9138	LAGC	48	11.7424	111.9545
QQQR	43	6.4864	205.5431	AEKL	96	34.4726	109.8154
CCDN	9	0.3754	198.1620	CGKS	21	2.9860	108.6753
MQHW	11	0.5910	183.3444	GKPL	68	20.6422	108.6495
QHWY	11	0.5912	183.2727	ALAA	237	122.3980	107.3025
CPHC	6	0.1863	181.4570	MKKL	37	7.9428	106.2995
IHCM	11	0.6143	175.5745	AWQQ	24	3.8298	106.2288
HCDH	9	0.4198	175.3657	WYGW	9	0.6602	105.3527
SGKS	65	14.8484	169.3910	EFCC	7	0.4144	104.6449
KSTL	77	19.9060	163.7560	QYQQ	22	3.3381	104.3313
CPRC	9	0.4534	161.0927	LTPE	65	19.7398	103.7746
CPQC	8	0.3626	160.8523	YRYD	21	3.1158	102.6516
GCGK	28	3.7788	155.2533	LLEQ	102	38.8697	102.5335
QWHF	12	0.8092	154.7701	LRQQ	55	15.5871	99.6584
CPEC	9	0.4705	154.6133	ALRE	109	43.3097	99.6361
NQWH	12	0.8197	152.4842	FWQQ	14	1.5753	97.9964
GCCY	8	0.3882	149.2338	WLAE	46	11.9418	97.1339
QQQL	55	12.4657	145.1315	CIDC	8	0.5703	96.7969
CTGC	11	0.7363	143.0719	MSHQ	17	2.2535	96.4972
CPYC	6	0.2333	142.5108	...	...	...	...
QQAL	87	26.6933	136.2474	...	...	...	...
CSHC	6	0.2449	135.2474	DKAG	26	21.3389	1.0181
CPTC	8	0.4425	129.0628	KDAG	26	21.3389	1.0181

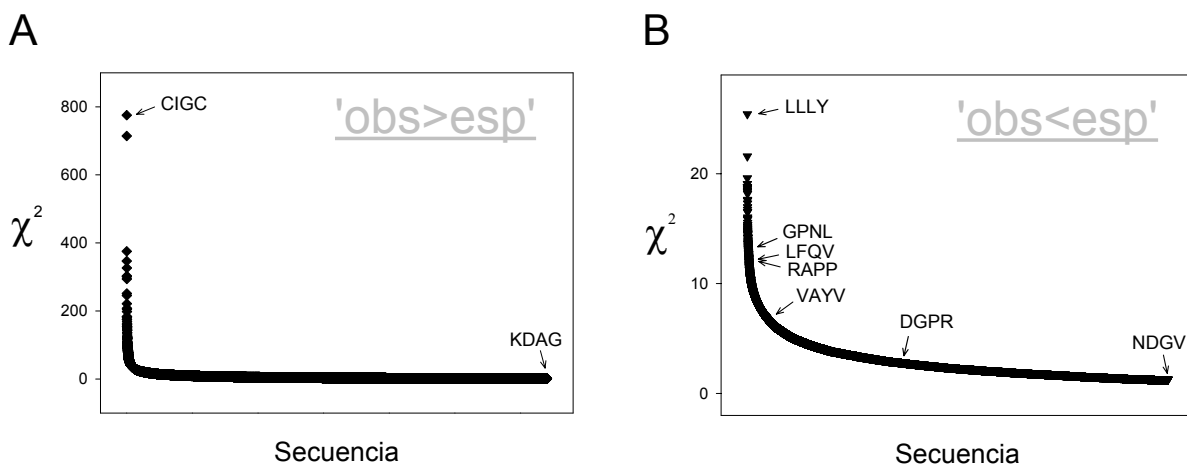
TABLA XVIII. Secuencias de 4 aminoácidos con una frecuencia **menor** a la esperada por azar en el proteoma de *E. coli*

Secuencia	Frecuencia observada	Frecuencia esperada	$\chi^2$	Secuencia	Frecuencia observada	Frecuencia esperada	$\chi^2$
LLLY	12	46.3204	25.4292	IVKL	7	26.8018	14.6300
LILQ	11	40.6079	21.5876	GLQA	19	44.4125	14.5409
VLKL	17	47.5102	19.5931	LGPA	19	44.4125	14.5409
ELIL	21	52.6921	19.0615	QLGA	19	44.4125	14.5409
LGPP	1	20.7405	18.7887	AVVQ	8	28.2096	14.4783
QGPL	1	20.7405	18.7887	LEVQ	18	42.8875	14.4422
DLEL	16	45.0920	18.7692	LGPI	8	28.1161	14.3923
LDLE	16	45.0920	18.7692	DNLG	4	21.4694	14.2146
LILK	13	40.4153	18.5969	QTEL	3	19.7398	14.1957
GPPA	0	18.4819	18.4819	RAGP	5	23.1097	14.1915
LLVK	18	47.5102	18.3298	ELLL	57	93.4048	14.1889
ALGP	16	44.4125	18.1767	DAGP	4	21.4405	14.1868
ALEV	24	55.1968	17.6322	ELGS	13	35.3831	14.1594
GARA	20	49.4858	17.5689	LLIR	24	50.7759	14.1198
ILIQ	3	22.9080	17.3009	GPER	0	14.0037	14.0037
GAAD	18	45.9115	16.9685	ILVR	12	33.6725	13.9490
GPDJ	4	24.0607	16.7257	LIKI	5	22.7994	13.8959
AGPP	1	18.4819	16.5360	LEIG	14	36.4830	13.8553
LEVA	25	55.1968	16.5199	...	...	...	...
TEVL	9	31.4571	16.0320	GPNL	3	18.5068	12.9931
ILIE	8	29.7250	15.8781	...	...	...	...
GRLG	17	43.1489	15.8466	LFQV	3	17.4970	12.0114
LQLG	22	49.8400	15.5511	...	...	...	...
ALQV	17	42.5381	15.3321	<b>RAPP</b>	<b>1</b>	<b>13.8897</b>	<b>11.9617</b>
AVQL	17	42.5381	15.3321	...	...	...	...
EFAL	9	30.5076	15.1627	VAYV	7	18.1525	6.8518
DLLM	4	22.3537	15.0695	...	...	...	...
LVLQ	21	47.7366	14.9748	AFEI	7	17.2102	6.0573
LVRA	25	53.1895	14.9399	...	...	...	...
KLLL	39	71.6423	14.8728	DGPR	7	12.5198	2.4336
EAPL	12	34.6368	14.7943	...	...	...	...
DLGR	9	30.0854	14.7777	...	...	...	...
RTEA	4	21.9946	14.7221	GNDV	10	14.2376	1.2613
LQLE	15	38.8697	14.6582	NDGV	10	14.2376	1.2613

TABLA XIX. Secuencias de 4 aminoácidos **no codificadas** en el proteoma de *E. coli*

Secuencia	Frecuencia observada	Frecuencia esperada	$\chi^2$	Secuencia	Frecuencia observada	Frecuencia esperada	$\chi^2$
GPPA	0	18.4819	18.4819	FTYL	0	8.6222	8.6222
GPER	0	14.0037	14.0037	FFQA	0	8.6176	8.6176
AGSH	0	12.4814	12.4814	RSPP	0	8.5280	8.5280
YAML	0	11.0491	11.0491	KPRS	0	8.4876	8.4876
GDHA	0	11.0131	11.0131	YVTE	0	8.4236	8.4236
RGPP	0	10.7922	10.7922	HDRA	0	8.2767	8.2767
PVQE	0	10.7268	10.7268	IQNS	0	8.2498	8.2498
FSIR	0	10.1824	10.1824	RDTN	0	8.1940	8.1940
KVPT	0	10.0406	10.0406	TDNR	0	8.1940	8.1940
KGPD	0	9.9652	9.9652	KEPT	0	8.1756	8.1756
EFQG	0	9.8643	9.8643	AYEF	0	8.1694	8.1694
FEGP	0	9.8643	9.8643	EFAY	0	8.1694	8.1694
TDQE	0	9.5295	9.5295	FAYE	0	8.1694	8.1694
FQGR	0	9.5056	9.5056	DRTF	0	8.0882	8.0882
IPIK	0	9.4878	9.4878	WDLV	0	7.9454	7.9454
LLWN	0	9.2155	9.2155	TRPP	0	7.9158	7.9158
QDRT	0	9.1830	9.1830	KQDS	0	7.8745	7.8745
MRGS	0	9.1272	9.1272	AYFR	0	7.8723	7.8723
GGCL	0	9.1238	9.1238	RIPN	0	7.8461	7.8461
YIRV	0	9.0169	9.0169	IFPR	0	7.7448	7.7448
RLMF	0	8.8312	8.8312	...	...	...	...
GFQD	0	8.8190	8.8190	...	...	...	...
GPPQ	0	8.6310	8.6310	CCWC	0	0.0331	0.0331
PGPP	0	8.6310	8.6310	WCCC	0	0.0331	0.0331

El índice  $\chi^2$  de los grupos de secuencias 'obs>esp' y 'obs<esp' indica que en ambas distribuciones hay una divergencia estadísticamente significativa ( $P < 0.001$ ) entre las frecuencias observadas y las frecuencias esperadas. En ambas distribuciones las secuencias que merecen una atención especial son aquellas que muestran un mayor valor del índice  $\chi^2$ . El 6.55% de las secuencias 'obs>esp' tiene un índice  $\chi^2$  superior a 10 en una escala 0 a 775.33 (Fig. 48A), mientras que el 0.31% de las secuencias 'obs<esp' tiene un índice  $\chi^2$  superior a 10 en una escala 0 a 25.43 (Fig. 48B). La secuencia de aminoácidos RAPP pertenece a este segundo grupo de secuencias que se hallan representadas en *E. coli* con frecuencias menores a la esperada por una hipotética combinación al azar de los aminoácidos, y en particular de los aminoácidos R, A y P.



**Figura 48.** Distribución del valor de  $\chi^2$  en los conjuntos de secuencias 'obs>esp' y 'obs<esp'. **(A)** Ordenación de mayor a menor  $\chi^2$  del conjunto de secuencias de 4 aminoácidos que en *E. coli* muestran una frecuencia **mayor** a la esperada por azar. Se señalan las secuencias del conjunto que presentan el mayor y menor  $\chi^2$  (Tabla XVII). **(B)** Ordenación de mayor a menor  $\chi^2$  del conjunto de secuencias que en *E. coli* muestran una frecuencia **menor** a la esperada por azar (no se incluyen las secuencias no codificadas en *E. coli*). Se señalan las secuencias del conjunto que presentan el mayor y menor  $\chi^2$ , y otras que han sido objeto de estudio en este trabajo experimental (Tabla XVIII).

La secuencia de aminoácidos TRAPP no se halla codificada en ningún ORF del genoma de *E. coli*. En general, esta secuencia está poco representada entre las eubacterias donde, además, la mayoría de ORFs que la codifican son hipotéticos, es decir, no se ha probado para ellos la existencia del producto de expresión. La secuencia RAPP, en cambio, sí está codificada en el proteoma de *E. coli* aunque con una frecuencia menor a la que se espera por la combinación al azar de los aminoácidos. Esta divergencia se puede explicar por dos motivos. El primero y más sencillo es que la secuencia RAPP es patrimonio de proteínas impropias de *E. coli*, de modo que no tiene porqué hallarse en el proteoma de la bacteria. El segundo es que la expresión de RAPP resulta tóxica para *E. coli*, tal y como se ha demostrado durante la expresión de las secuencias TRAPP, TTRAPP o TTRAPP. En estas circunstancias, la secuencia RAPP pudiera estar formando parte de un dominio proteico incompatible por su estructura y/o función con el crecimiento normal de *E. coli*. La dificultad de expresar la secuencia TRAPP de la gE en *E. coli* se correlaciona pues con una carencia intrínseca de dicha secuencia, y también de RAPP, en los polipéptidos que codifica *E. coli* de forma natural.

### III.C.2. Expresión de otras secuencias minoritarias en *E. coli*

La búsqueda de la secuencia de aminoácidos TRAPP codificada de forma natural en los ORFs de *E. coli*, ha revelado la existencia de un nutrido grupo de péptidos de 4 aminoácidos que como RAPP se hallan codificados con una frecuencia menor a la que se espera por simple azar. En un intento por descubrir si dicha divergencia se debe a la toxicidad que pudiera implicar su expresión, o si es un hecho casual, se ha llevado a cabo la expresión en *E. coli* de algunas de estas secuencias. Los ensayos de expresión se han desarrollado sobre la MBP según las dos estrategias descritas anteriormente: (i) mediante la inserción en la estructura tipo lazo que se extiende entre los aminoácidos D<sub>120</sub> e I<sub>132</sub> y, (ii) mediante la substitución de la secuencia de aminoácidos que compone la hélice- $\alpha$  IV (Fig. 28C).

#### III.C.2.A. Inserción de otras secuencias minoritarias en la MBP

Las secuencias insertadas en el lazo de unión D<sub>120</sub>-I<sub>132</sub> de la MBP han sido GPNL, LFQV, VAYV y AFEI. Estas secuencias tienen índices  $\chi^2$  de divergencia entre la frecuencia observada y la esperada por azar de 12.9930, 12.0113, 6.8518 y 6.0573, respectivamente (Tabla XVIII).

La inserción de las secuencias GPNL, LFQV, VAYV y AFEI en el lazo D<sub>120</sub>-I<sub>132</sub> se ha realizado mediante adaptadores *Bst*XI diseñados para ser traducidos en el uso de codón optimizado de *E. coli* (Tabla XX). La inserción de estos adaptadores se ha llevado a cabo sobre el plásmido pc $\Delta$ malE que presenta como única diana *Bst*XI la de la región codificante para el lazo D<sub>120</sub>-I<sub>132</sub> en el gen *malE*. Los plásmidos resultantes tienen 5313 pb y se han llamado pcGPNL, pcLFQV, pcVAYV y pcAFEI, respectivamente. La digestión de estos plásmidos con *Bg*II y *Xba*I libera un fragmento de DNA de 837 pb codificante para la secuencia de la MBP a partir del aminoácido D<sub>120</sub>, y que incluye la inserción respectiva. Estos fragmentos, una vez purificados, se han insertado en el plásmido pMAL<sup>TM</sup>-c2 linearizado con *Bg*II y *Xba*I. Los plásmidos obtenidos se han llamado pMGPNLi, pMLFQV, pMVAYV y pMAFEI, y con ellos se ha transformado células competentes XL1-Blue para ensayar la expresión de la MBP mutagenizada.

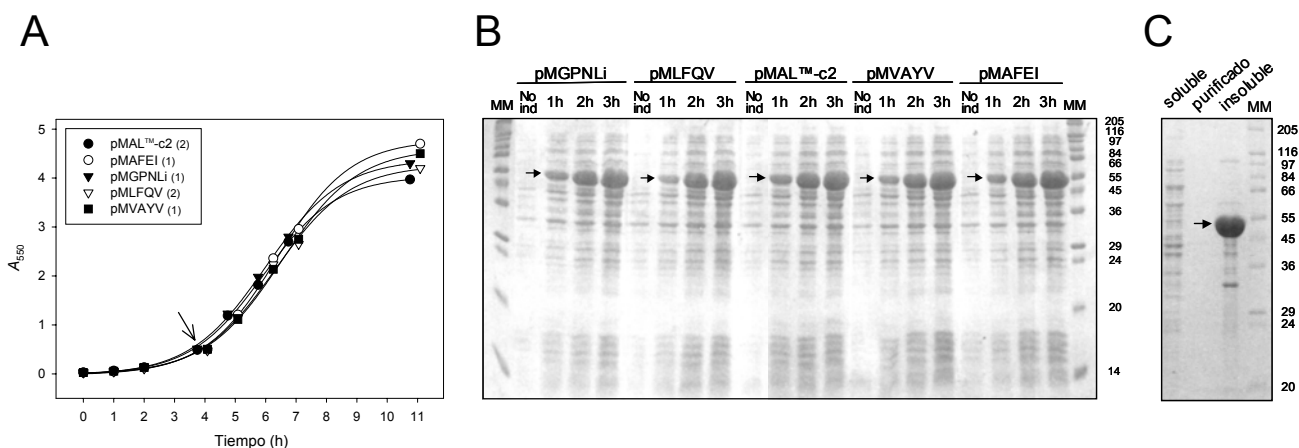


TABLA XX. Adaptadores *Bst*XI para la inserción de secuencias minoritarias en la MBP

Secuencia	Plásmido de expresión
<p style="text-align: center;"><b>G P N L K</b></p> <p>5' <b>cggcccgaacctgaaaac</b> 3'</p> <p><b>tttgg</b>ccgggcttggaact</p> <p><i>Bst</i>XI* <span style="float: right;"><i>Bst</i>XI*</span></p>	pMGPNLi
<p style="text-align: center;"><b>L F Q V K</b></p> <p>5' <b>cctgtttcagggtgaaaac</b> 3'</p> <p><b>tttgg</b>gacaaaagtccaact</p> <p><i>Bst</i>XI* <span style="float: right;"><i>Bst</i>XI*</span></p>	pMLFQV
<p style="text-align: center;"><b>V A Y V K</b></p> <p>5' <b>cgtggcgtatgtgaaaac</b> 3'</p> <p><b>tttgg</b>caccgcatacaact</p> <p><i>Bst</i>XI* <span style="float: right;"><i>Bst</i>XI*</span></p>	pMVAYV
<p style="text-align: center;"><b>A F E I K</b></p> <p>5' <b>cgcgtttgaaattaaaac</b> 3'</p> <p><b>tttgg</b>gccaaactttaat</p> <p><i>Bst</i>XI* <span style="float: right;"><i>Bst</i>XI*</span></p>	pMAFEI

*Bst*XI\* significa que no se conserva la diana de restricción *Bst*XI tras la inserción del adaptador.

La cinética de crecimiento de los cultivos XL1-Blue pMGPNLi, pMLFQV, pMVAYV y pMAFEI muestra que tras la inducción de expresión de la MBP, ninguno de ellos presenta inhibición del crecimiento comparado con el cultivo control XL1-Blue pMAL™-c2 (Fig. 49A). Este resultado indica que la expresión de las secuencias GPNL, LFQV, VAYV o AFEI insertadas en la MBP no interfiere con el crecimiento normal del cultivo. Por otro lado, en estos cultivos sí es de destacar la formación temprana de cuerpos de inclusión, claramente visibles tras la primera hora de inducción. Estas estructuras, que no se observan en el cultivo control, sugieren que la MBP expresada se acumula básicamente de forma insoluble. El análisis por SDS-PAGE de los cultivos inducidos muestra en todos ellos la presencia de una banda mayoritaria cuya migración electroforética coincide con la de la MBP codificada en el cultivo XL1-Blue pMAL™-c2 (Fig. 49B), y que refleja



**Figura 49.** Expresión en *E. coli* de secuencias minoritarias de 4 aminoácidos (I). **Inserción** en el lazo D120-1132 de la MBP. **(A)** Cinética de crecimiento de clones de *E. coli* **XL1-Blue** que codifican para las inserciones GPNL, LFQV, VAYV y AFEI. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. La flecha señala el momento de inducción de los cultivos. El clon pMAL™-c2 constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo a 1, 2 y 3 horas post-inducción. 'No ind': extracto crudo del cultivo no inducido. '1h', '2h', '3h': extracto crudo del cultivo tras 1, 2 y 3 horas de inducción. Las flechas denotan la acumulación de la proteína expresada. Condiciones de electroforesis: gel acrilamida 15% y 0.8 mm, 12.5-15 mA/gel. **(C)** SDS-PAGE en condiciones reductoras del extracto crudo fraccionado del cultivo pMLFQV tras 5 horas de inducción. 'soluble': fracción soluble. 'purificado': extracto purificado de la fracción soluble mediante resina de amilosa (Protein Fusion and Purification System pMAL™-2, New England Biolabs). 'insoluble': fracción insoluble. La flecha señala la MBP que contiene la secuencia LFQV. Condiciones de electroforesis: gel acrilamida 12% y 0.8 mm, 12.5-15 mA. En **B** y **C**, 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM.

la expresión y acumulación en el citoplasma de *E. coli* de la MBP que contiene las secuencias GPNL, LFQV, VAYV y AFEI, respectivamente. Por otro lado, el fraccionamiento de los cultivos control y pMLFQV confirma que mientras la MBP expresada en el cultivo control es soluble, la MBP que contiene la inserción LFQV es mayoritariamente insoluble (Fig. 49C).

### III.C.2.B. Substitución de otras secuencias minoritarias en la MBP

Las secuencias de aminoácidos utilizadas para substituir la secuencia ALDKELKAK en la hélice- $\alpha$  IV de la MBP han sido GPNL y DGPR. También se ha utilizado dichas secuencias para substituir parcialmente la hélice- $\alpha$  IV de la MBP(*EcoNot*). Las secuencias GPNL y DGPR tienen índices  $\chi^2$  de divergencia entre las frecuencias observada y esperada por azar en *E. coli* de 12.9930 y 2.4336, respectivamente (Tabla XVIII).

La substitución de la secuencia ALDKELKAK en la hélice- $\alpha$  IV de la MBP por GPNL y DGPR se ha realizado mediante adaptadores *Bst*XI-*Bss*HII que codifican para dichas secuencias en el uso de codón optimizado para *E. coli* (Tabla XXI). Estos adaptadores han sido insertados en el plásmido *lit* $\Delta$ *malE* linearizado con *Bst*XI y *Bss*HII. Los plásmidos resultantes se han llamado *lit* $\Delta$ *malE*GPNL y *lit* $\Delta$ *malE*DGPR, respectivamente, y su digestión con *Bgl*II y *Xba*I libera el fragmento del gen *malE* con los adaptadores insertados. Una vez purificados, cada uno de los fragmentos ha sido insertado en el plásmido pMAL<sup>TM</sup>-c2 linearizado con *Bgl*II y *Xba*I. Los plásmidos obtenidos se han llamado pMGPNLs y pMDGPR, y con ellos se ha transformado células competentes XL1-Blue para ensayar la expresión de las secuencias GPNL y DGPR que substituyen la hélice- $\alpha$  IV de la MBP.

TABLA XXI. Adaptadores *Bst*XI-*Bss*HII y *Eco*RI-*Not*I para la substitución de secuencias minoritarias en la MBP

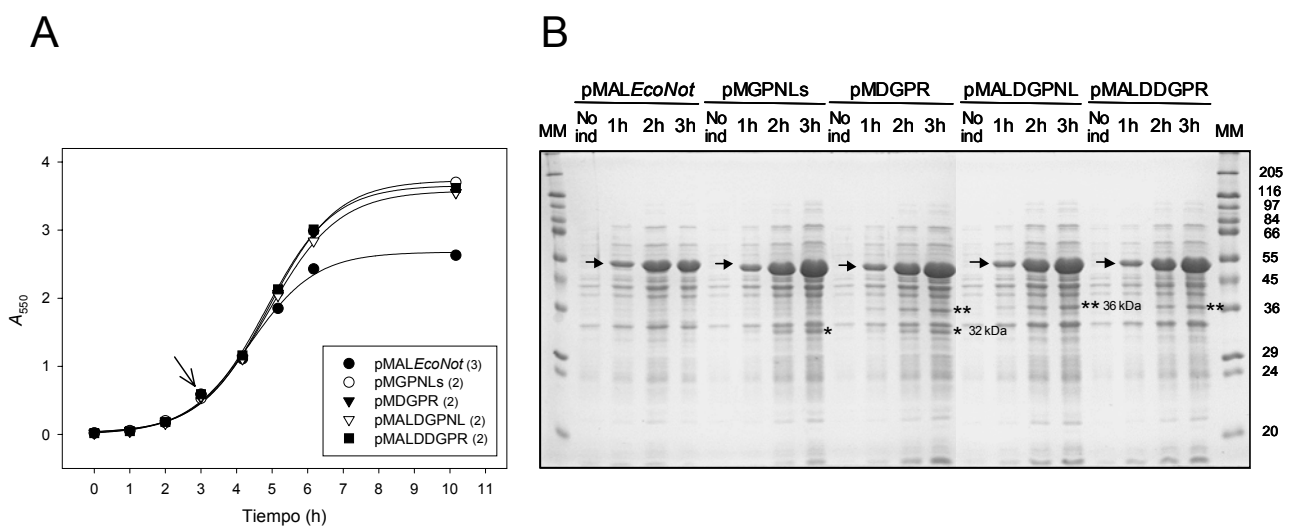
Tipo de adaptador	Secuencia	Plásmido de expresión
<i>Bst</i> XI- <i>Bss</i> HII	<pre> W E E I P G P N L G K 5' ctgggaagagattccgggccccaacctgggtaagag 3' tttggacccttctctaaggccgggcttggaccattctcgcgc BstXI BssHII </pre>	pMGPNLs
	<pre> W E E I P D G P R G K 5' ctgggaagagattccgggatggtccgcgtggtaagag 3' tttggacccttctctaaggcctaccaggcgcaaccattctcgcgc BstXI BssHII </pre>	pMDGPR
<i>Eco</i> RI- <i>Not</i> I	<pre> F P A L D G P N L 5' aattccggcgctggatggtccgaacctg 3' gggccgcgacctaccaggcttggaccggg EcoRI NotI* </pre>	pMALDGPNL
	<pre> F P A L D D G P R 5' aattccggcgctggatgatggtccgcgt 3' gggccgcgacctactaccaggcgcaaccggg EcoRI NotI* </pre>	pMALDDGPR

*Not*I\* significa que no se conserva la diana de restricción *Not*I tras la inserción del adaptador.

La substitución parcial de la secuencia ALDKELR en la hélice- $\alpha$  IV de la MBP(*EcoNot*) se ha realizado mediante adaptadores *Eco*RI-*Not*I que codifican para las secuencias ALDGPNL y ALDDGPR en el uso de codón optimizado de *E. coli* (Tabla XXI). La inserción de estos adaptadores se ha realizado en el plásmido

lit $\Delta$ malE*EcoNot* linearizado con *EcoRI* y *NotI*. Los plásmidos resultantes se han llamado lit $\Delta$ malEALDGPNL y lit $\Delta$ malEALDDGPR, respectivamente. La digestión de estos plásmidos con *BglII* y *SacI* libera un fragmento de 750 pb correspondiente al gen *malE* que contiene las nuevas secuencias. Cada fragmento, una vez purificado, se ha insertado en el plásmido pMAL<sup>TM</sup>-c2 linearizado con *BglII* y *SacI*. Los plásmidos así obtenidos se han llamado pMALDGPNL y pMALDDGPR, y con ellos se ha transformado células competentes XL1-Blue para ensayar la expresión de la MBP que codifican.

La cinética de crecimiento de los cultivos inducidos XL1-Blue pMGPNLs, pMDGPR, pMALDGPNL y pMALDDGPR muestra que ninguno de ellos presenta inhibición del crecimiento con respecto al cultivo control XL1-Blue pMALE*EcoNot* (Fig. 50A). El valor final de  $A_{550}$  que alcanzan estos cultivos es un 30% superior al del cultivo control. Estos resultados indican que la expresión de las secuencias GPNL y DGPR en la superficie de la MBP no interfiere con el crecimiento normal del cultivo. En todos los cultivos a excepción del cultivo control se detecta, además, la formación de cuerpos de inclusión. El análisis por SDS-PAGE de los cultivos inducidos muestra una banda intensa que denota la acumulación de la MBP expresada (Fig. 50B). La cantidad de MBP acumulada, que a tenor de los cuerpos de inclusión observados es en buena parte en forma insoluble, es similar en los distintos cultivos y un 20% superior con respecto a la del cultivo control. En el extracto crudo de los cultivos inducidos pMGPNLs, pMDGPR, pMALDGPNL y pMALDDGPR se observan además algunas bandas minoritarias que no se hallan en el extracto crudo del cultivo control. Entre ellas destacan dos bandas cuya movilidad electroforética indica que están formadas por polipéptidos de 32 y 36 kDa, respectivamente. El análisis de la secuencia aminoacídica de estos polipéptidos revela que ambos son fragmentos de la MBP expresada. El polipéptido de 32 kDa equivale al fragmento que se extiende desde el extremo Nt hasta aproximadamente el segundo tercio de la hélice- $\alpha$  X de la MBP. El polipéptido de 36 kDa responde al fragmento de la MBP que se extiende desde el aminoácido D<sub>136</sub> en la hélice- $\alpha$  IV (punto de inserción de estas secuencias) hasta el extremo Ct de la proteína. El hecho reiterado de hallar un residuo



**Figura 50.** Expresión en *E. coli* de secuencias minoritarias de 4 aminoácidos (II). **Substitución** de la hélice- $\alpha$  IV de la MBP. **(A)** Cinética de crecimiento de clones de *E. coli* XL1-Blue que codifican para las sustituciones GPNL y DGPR. Las curvas representadas son la media del número de cinéticas individuales que se indica entre paréntesis junto a cada clon. La flecha señala el momento de inducción de los cultivos. El clon pMALE*EcoNot* constituye el cultivo control. **(B)** SDS-PAGE en condiciones reductoras del extracto crudo a 1, 2 y 3 horas post-inducción. 'No ind': extracto crudo del cultivo no inducido. '1h', '2h', '3h': extracto crudo del cultivo tras 1, 2 y 3 horas de inducción. 'MM': marcador de peso molecular (Wide Molecular Weight Range Sigma Marker, Sigma). Los números a la derecha indican el tamaño, en kDa, de los polipéptidos que componen el MM. Las flechas denotan la acumulación de la MBP expresada. Se indica la presencia de polipéptidos de degradación de 32 (\*\*\*) y 36 kDa (\*\*\*) de la MBP expresada. Condiciones de electroforesis: gel acrilamida 12% y 0.8 mm, 10-12.5 mA/gel.

Asp (D) en alguno de los extremos de estos fragmentos, así como de hallar fragmentos similares en el extracto crudo de otros cultivos, apuntan a que estos fragmentos puedan deberse a una proteólisis inducida por el calor aplicado durante la preparación de la muestra para su análisis mediante SDS-PAGE.

La búsqueda de la secuencia de aminoácidos TRAPP en el conjunto de proteínas de *E. coli* y de aquellas depositadas en bancos de datos informáticos refleja que se trata de una secuencia poco habitual en *E. coli*, y en general entre los organismos procariontes. Esta secuencia, y otras más identificadas, constituyen un grupo de péptidos que se hallan en *E. coli* en una frecuencia real siempre menor a la esperada por simple azar. A diferencia de TRAPP, algunas de estas secuencias sí son expresadas y se acumulan en *E. coli*. Este resultado prueba que la discrepancia entre la frecuencia observada y esperada no significa que la expresión de un péptido determinado vaya a ser imposible en *E. coli*. Por otro lado, no obstante, la acumulación de la proteína que contiene estas secuencias se ha dado siempre en forma de cuerpos de inclusión, lo cual denota una inestabilidad de la proteína, su incompatibilidad en el entorno citoplasmático de *E. coli*, o ambos a la vez. Los resultados de expresión de estas secuencias no permiten concluir qué mecanismos se activan y hacen imposible la acumulación de la secuencia TRAPP, ni tan siquiera como forma agregada insoluble.

### **III.D. Construcción de un virus recombinante**

Los resultados expuestos hasta ahora demuestran que la expresión de la gE del BoHV-1 es tóxica para *E. coli*, y que la secuencia de aminoácidos TRAPP en particular es la principal responsable de esta toxicidad. La secuencia TRAPP es, además, la secuencia de aminoácidos más corta de la gE cuya expresión resulta tóxica. La supresión de esta secuencia restablece el crecimiento normal del cultivo así como la expresión y acumulación de la gE. Cabe señalar que la secuencia TRAPP también interfiere con el crecimiento normal del cultivo y/o la expresión de proteínas cuando se halla integrada en proteínas de *E. coli* tales como la MBP o la 'proteína LacZ'. Estos resultados adquieren un interés especial cuando se comprueba que TRAPP, y también la secuencia RAPP, tienen muy escasa representación en proteínas de origen procarionta, mientras que son más frecuentes en proteínas de origen eucariota o vírico. En el genoma de *E. coli* íntegramente secuenciado, la secuencia RAPP aparece codificada en un único ORF correspondiente a una hipotética proteína de 44.5 kDa situado en la región intergénica *cysG-trpS*. El conjunto de estos datos convierte la secuencia TRAPP en una secuencia de aminoácidos muy particular, cuya presencia en la gE pudiera ser relevante para la estructura/función de esta glicoproteína.

Los estudios llevados a cabo en nuestro grupo de investigación demostraron la importancia de la gE en la transmisión directa 'célula-a-célula' del BoHV-1 (Rebordosa *et al.*, 1996). Esto ya había sido descrito para la gE de otros miembros de la subfamilia *alphaherpesvirinae* como p.e. el SuHV-1 (Zsak *et al.*, 1992; Jacobs, 1994) o el HHV-1 (Dingwell *et al.*, 1994; Dingwell *et al.*, 1995). Actualmente se sabe que la gE se une a la gI y que el complejo gE/I participa en dos procesos relacionados entre sí: la morfogénesis y la transmisión 'célula-a-célula' de los alphaherpesvirus. En concreto, el complejo gE/I participa en el transporte de la vesícula que ha de formar la envoltura vírica hacia las regiones de la membrana plasmática que contactan con las células adyacentes (Johnson & Huber, 2002). Esta función, aunque no resulta esencial para la viabilidad del BoHV-1, sí es específica del complejo gE/I. Por otro lado, también se ha descrito que la gE del HHV-1, HHV-3 y SuHV-1 puede unir la fracción constante (Fc) de las IgG. Esta actividad FcR permite a los viriones escapar de la neutralización mediada por anticuerpo y por complemento, y a las células infectadas escapar de la respuesta inmune citotóxica dependiente de anticuerpo (Dubin *et al.*, 1991; Favoreel *et al.*, 1997; Lubinski *et al.*, 2002). Cabe señalar, no obstante, que esta función de modulación de la respuesta inmunológica no ha sido descrita para otros alphaherpesvirus, entre ellos el BoHV-1, y que por tanto podría tratarse de un caso particular de una actividad más general de unión de receptores celulares.

Con el objetivo de estudiar la aportación de la secuencia TRAPP a la estructura/función de la gE del BoHV-1, se ha construido una cepa viral BoHV-1 recombinante defectiva para esta secuencia. En este apartado se recogen los resultados obtenidos de dicha construcción y de su caracterización en términos de infección/propagación sobre cultivos celulares en medio líquido y en medio semisólido.

#### **III.D.1. Construcción de un virus BoHV-1 recombinante defectivo para TRAPP**

Nuestro grupo de investigación puso a punto un método para la obtención de cepas recombinantes del BoHV-1 durante la construcción de una cepa defectiva para la gE (BoHV-1 gE<sup>-</sup>) (Rebordosa, 1996). Este

método es una adaptación del descrito inicialmente por X. Liang y colaboradores (Liang *et al.*, 1991), y se basa en la recombinación *in vivo* entre el DNA viral y un DNA plasmídico co-transfectados en líneas celulares. La molécula de DNA plasmídico, o plásmido lanzadera, es portadora de una copia de los fragmentos de secuencia del genoma del virus entre los cuales se desea llevar a cabo la modificación genética. Estos fragmentos de DNA viral constituyen lo que se denomina 'brazos' del plásmido lanzadera y flanquean la modificación que se desea introducir. Esta modificación puede ser por inserción, sustitución o delección de secuencia. Tras la co-transfección de las células con el DNA viral y el plásmido lanzadera, cabe esperar un episodio de recombinación homóloga dentro de cada uno de los brazos entre ambas moléculas. Este episodio de doble recombinación incorpora el cambio deseado en el genoma vírico.

Para construir el virus BoHV-1 defectivo para la secuencia TRAPP, se ha transfectado células GBK con DNA del virus BoHV-1 gE<sup>-</sup> y un plásmido lanzadera que contiene una copia del gen *gE*, en el que la secuencia de DNA que codifica para RAPP<sup>r</sup> ha sido substituida por otra que codifica SHEFP. La cepa vírica resultante se ha llamado BoHV-1 RAPP<sup>r</sup>. A continuación se describen los resultados obtenidos en la construcción de este virus.

#### III.D.1.A. El virus BoHV-1 gE<sup>-</sup>

El virus BoHV-1 gE<sup>-</sup> es una cepa derivada del BoHV-1 FM, defectiva para la gE (Rebordosa, 1996). El virus BoHV-1 gE<sup>-</sup> se utiliza actualmente como vacuna viva pues induce una respuesta inmune protectora frente al virus salvaje BoHV-1 FM. Esta vacuna es, además, una vacuna marcadora pues su utilización conjunta con un ensayo de detección de anticuerpos anti-gE en suero permitiría la diferenciación serológica entre animales vacunados y animales infectados.

El genoma del BoHV-1 gE<sup>-</sup> contiene una secuencia de DNA de 3072 pb correspondiente al gen *lacZ* casi íntegro de *E. coli*, que substituye una secuencia de 1620 pb del gen *gE*. La secuencia del gen *gE* substituida codifica para 540 aminoácidos del extremo Ct de la gE, esto es, el 93.9% de la secuencia de aminoácidos de la gE. El fragmento insertado del gen *lacZ* codifica para 1016 aminoácidos, y representa la secuencia íntegra de la enzima β-galactosidasa a excepción de los 10 primeros aminoácidos de su extremo Nt. La inserción del gen *lacZ* en el BoHV-1 gE<sup>-</sup> es en la orientación inversa a la que ocupa el gen *gE* en el BoHV-1 FM. Esta orientación, así como la inclusión de codones de finalización de traducción (señales STOP), evitan la síntesis de la β-galactosidasa a partir de la región promotora, no delecionada, del gen *gE* en el BoHV-1 gE<sup>-</sup>.

La inserción del gen *lacZ* en el genoma de BoHV-1 gE<sup>-</sup> permitió aislar y purificar mediante hibridación *in situ* este virus defectivo. El virus obtenido se caracterizó genéticamente mediante un mapa de restricción y transferencia Southern (Rebordosa, 1996).

III.D.1.B. Construcción del plásmido lanzadera

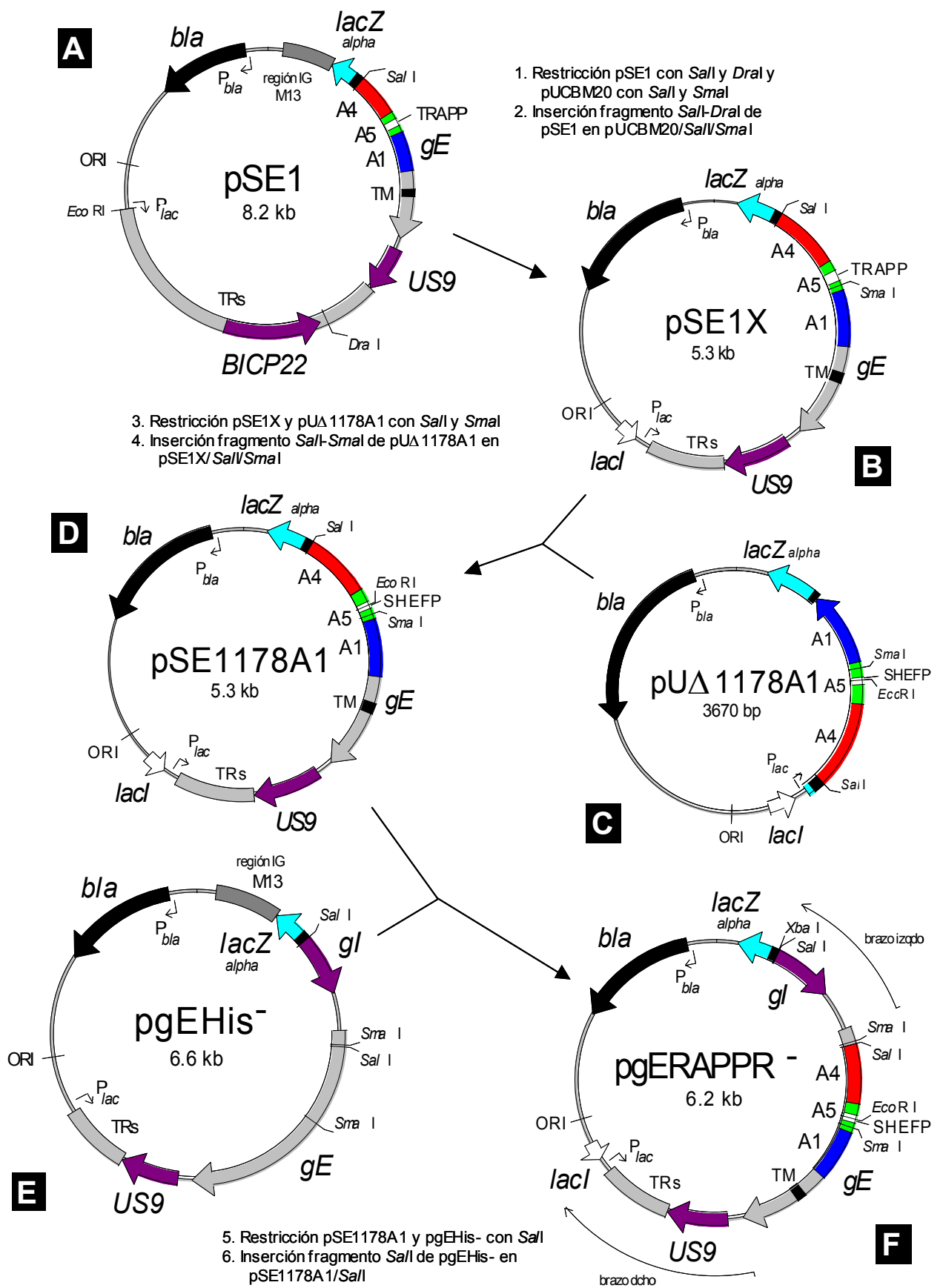
Para obtener una cepa viral BoHV-1 FM defectiva en la secuencia TRAPP de la gE se ha construido un plásmido lanzadera que contiene íntegro el marco de lectura del gen *gE* excepto en la secuencia codificante para TRAPP.

Los clones transformantes de *E. coli* pMΔ1178A1, pMΔ1190A1 y pMΔ4198A1 codifican para el dominio extracelular de la gE fusionado al extremo Ct de la MBP. Todos ellos son clones no tóxicos para la expresión de la gE en la que se ha truncado la secuencia original <sub>212</sub>TRAPP<sub>216</sub>. En los plásmidos pMΔ1178A1 y pMΔ1190A1 se ha substituido, respectivamente, las secuencias <sub>213</sub>RAPPR<sub>217</sub> y <sub>213</sub>RA<sub>214</sub> por SHEFP, mientras que en el plásmido pMΔ4198A1 la secuencia <sub>205</sub>PPPHRTTT<sub>212</sub> ha sido substituida por HPMEANS. A pesar de que cualquiera de estas tres construcciones da pie a construir un plásmido lanzadera carente de TRAPP, se ha optado por incorporar el cambio de secuencia del plásmido pMΔ1178A1 ya que es el único que conserva el número de aminoácidos del dominio extracelular de la gE. La cepa viral resultante será por tanto defectiva en la secuencia TRAPP en tanto que la secuencia original <sub>213</sub>RAPPR<sub>217</sub> es substituida por SHEFP.

Los 'brazos' izquierdo y derecho del plásmido lanzadera se han obtenido, respectivamente, a partir de los plásmidos pgEHis<sup>-</sup> y pSE1 construidos por X. Rebordosa y colaboradores durante la construcción del virus defectivo BoHV-1 gE<sup>-</sup> (Rebordosa, 1996). El plásmido pgEHis<sup>-</sup> (6.6 kb) contiene un fragmento *Sall-DraI* de 3.5 kb del BoHV-1 FM que se extiende 0.7 y 1 kb, respectivamente, por los extremos 5' y 3' del ORF del gen *gE*, y que codifica para la gE y para la también proteína de membrana US9 (Fig. 51E). El plásmido pSE1 (8.2 kb) contiene un fragmento *Sall-EcoRI* de 5.1 kb que se extiende desde la base T118 del ORF del gen *gE* hasta 3.5 kb desde su extremo 3', y que codifica para la proteína US9 y para la proteína BICP22 represora de genes  $\alpha$  y genes  $\gamma^2$  (Fig. 51A).

La digestión del plásmido pSE1 con *Sall* y *DraI* libera un fragmento de DNA de 2.6 kb codificante para la gE desde la V<sub>40</sub>, y para la proteína US9 (Fig. 51A). Este fragmento, una vez purificado, ha sido insertado en el plásmido pUCBM20 previamente linearizado con *Sall* y *SmaI*. El plásmido resultante se ha llamado pSE1X y codifica para la secuencia TRAPP en el gen *gE* (Fig. 51B). Por otro lado, la digestión del plásmido pUΔ1178A1 con *Sall* y *SmaI* libera un fragmento de 582 pb codificante para la secuencia V<sub>40</sub>-P<sub>235</sub> de la gE en la que SHEFP substituye la secuencia original RAPPR (Fig. 51C). Este fragmento, una vez resuelto y purificado, ha sido insertado entre las dianas *Sall* y *SmaI* del plásmido pSE1X y reemplaza así la secuencia original del gen *gE* del BoHV-1 FM. El plásmido resultante se ha llamado pSE1178A1 (Fig. 51D). La digestión del plásmido pgEHis<sup>-</sup> con *Sall* libera un fragmento de 839 pb (Fig. 51E). Este fragmento, una vez purificado, ha sido insertado en el plásmido pSE1178A1 previamente linearizado con *Sall*. Mediante una digestión con *SmaI*, se ha seleccionado la construcción que presenta el fragmento *Sall* correctamente orientado. La construcción finalmente seleccionada es el plásmido lanzadera, el cual codifica para la substitución de RAPPR por SHEFP y presenta unos brazos izquierdo y derecho de 721 y 1045 pb, respectivamente. Al plásmido lanzadera se le ha llamado pgERAPPR<sup>-</sup> (Fig. 51F).

**Figura 51.** (Página 117) Diagrama de la construcción del plásmido lanzadera pgERAPPR<sup>-</sup>. **A, B, C, D, E y F** representan la ► secuencia de estadios diferentes del proceso, ordenada según las operaciones descritas 1. a 6. ORF: *bla* ( $\beta$ -lactamasa), *lacZalpha* (subunidad  $\beta$ -galactosidasa de  $\alpha$ -complementación), *lacI* (represor), *gE*, *gl*, *US9* y *BICP22* (proteínas codificadas en la región US del BoHV-1); *ori* (origen replicación); región IG M13 (origen replicación bacteriófago M13); TRs (secuencia repetida terminal de US); TM (transmembrana); P<sub>*bla*</sub>, P<sub>*lac*</sub> (promotores); *DraI*, *EcoRI*, *Sall*, *SmaI*, *XbaI* (dianas de restricción).





### III.D.1.C. Co-transfección, selección y enriquecimiento de viriones recombinantes

El DNA del BoHV-1 gE<sup>-</sup> y el vector lanzadera pgERAPPR<sup>-</sup> –previamente linearizado con *Xba*I– han sido co-transfectados en células GBK mediante electroporación (VI.F.1.A.1.).

El cultivo de células GBK co-transfectadas ha mostrado un efecto citopático total a las 72 horas post-transfección. A partir del sobrenadante de este cultivo, se han infectado distintos cultivos de células MDBK en medio semisólido. La selección de las calvas formadoras de viriones recombinantes se ha realizado mediante hibridación *in situ* (VI.F.1.A.3.) con dos DNA sonda distintos:

*gE*\*DIG: corresponde al fragmento *Hind*III de 1015 pb del plásmido pXJ marcado con digoxigenina y que contiene la secuencia de 997 pb codificante para la secuencia V<sub>40</sub>-A<sub>373</sub> (fragmentos A4, A5 y A1) del dominio extracelular de la gE, y

*lacZ*\*BIO: representa la casi totalidad del gen *lacZ* marcado con biotina. El marcaje se ha llevado a cabo sobre el fragmento *Bam*HI de 3072 pb del plásmido pΔgELac<sup>-</sup> codificante para el 99% de la secuencia de aminoácidos de la β-galactosidasa de *E. coli* (Rebordosa, 1996).

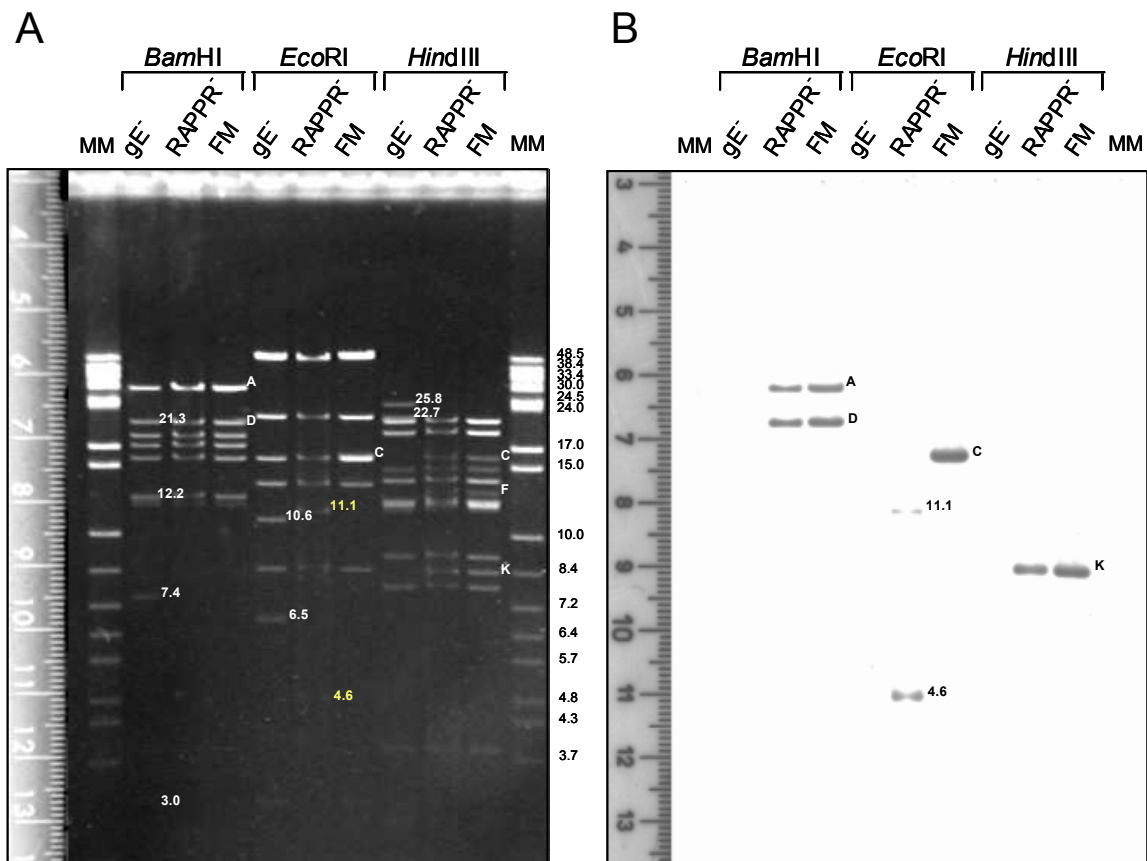
Las calvas que se forman por la infección de un virión defectivo BoHV-1 gE<sup>-</sup> han hibridado con la sonda *lacZ*\*BIO pues la progenie de dicho virión también contiene el gen *lacZ*. Por el contrario, las calvas que se forman por la infección de un virión BoHV-1 gE<sup>-</sup> que ha reincorporado el gen *gE* del plásmido lanzadera, han hibridado con la sonda *gE*\*DIG. La frecuencia obtenida de virus con reincorporación del gen *gE* ha sido del 0.1%, esto es, 1 de cada 1000 calvas contadas. Esta frecuencia es sensiblemente inferior al 0.5% descrito durante la construcción del BoHV-1 gE<sup>-</sup> (Rebordosa, 1996). Esta diferencia se puede explicar por el diferente tamaño de los ‘brazos’ de los plásmidos lanzadera utilizados. El tamaño de los ‘brazos’ determina la frecuencia de virus recombinantes obtenida pues condiciona la probabilidad de que en cada uno de ellos se dé un episodio de recombinación con el DNA viral. Así, a menor longitud de brazo menor es la secuencia hibridada y menor la probabilidad de recombinar. El plásmido lanzadera pΔgELac<sup>-</sup> utilizado para la construcción del virus BoHV-1 gE<sup>-</sup> tiene unos ‘brazos’ izquierdo y derecho de 1170 y 1703 pb, respectivamente, mientras que el plásmido lanzadera pgERAPPR<sup>-</sup> presenta unos ‘brazos’ izquierdo y derecho de 721 y 1045 pb, respectivamente. Por otro lado, durante la construcción del plásmido pgERAPPR<sup>-</sup> se consideró que la presencia de ‘brazos’ excesivamente largos podía complicar su manipulación e incluso comprometer su estabilidad durante la construcción en el hospedador procariota, y muy especialmente si, como es el caso, se trata de incorporar ‘brazos’ con un alto %G+C.

Tras seleccionar las calvas formadoras de viriones que han reincorporado el gen *gE*, se ha procedido a enriquecer la población de estos virus recombinantes. El objetivo ha sido conseguir una suspensión viral recombinante homogénea, es decir, una suspensión a partir de la cual todas las calvas formadas hibridan con la sonda *gE*\*DIG y no con la sonda *lacZ*\*BIO. Para ello han sido necesarias hasta cuatro rondas de purificación, y el porcentaje de virus recombinante obtenido en cada una de ellas ha sido del 0.1%, 10%, 33% y 100%, respectivamente. A partir del medio semisólido que cubría una calva *gE*\*DIG aislada en la última ronda de purificación, se ha infectado un cultivo de células MDBK en medio líquido sobre una placa de cultivo de 60 mm de diámetro. Finalmente, a partir del sobrenadante de infección obtenido, se ha infectado una monocapa de células MDBK en un frasco tipo Falcon de 75 cm<sup>2</sup> de superficie y posteriormente en un frasco de 175 cm<sup>2</sup>. La cantidad de partículas virales recombinantes producidas ha sido de 1.6x10<sup>9</sup> pfu/Falcon de 175 cm<sup>2</sup>, y la cantidad de DNA viral extraído 12 μg/Falcon de 175 cm<sup>2</sup>.

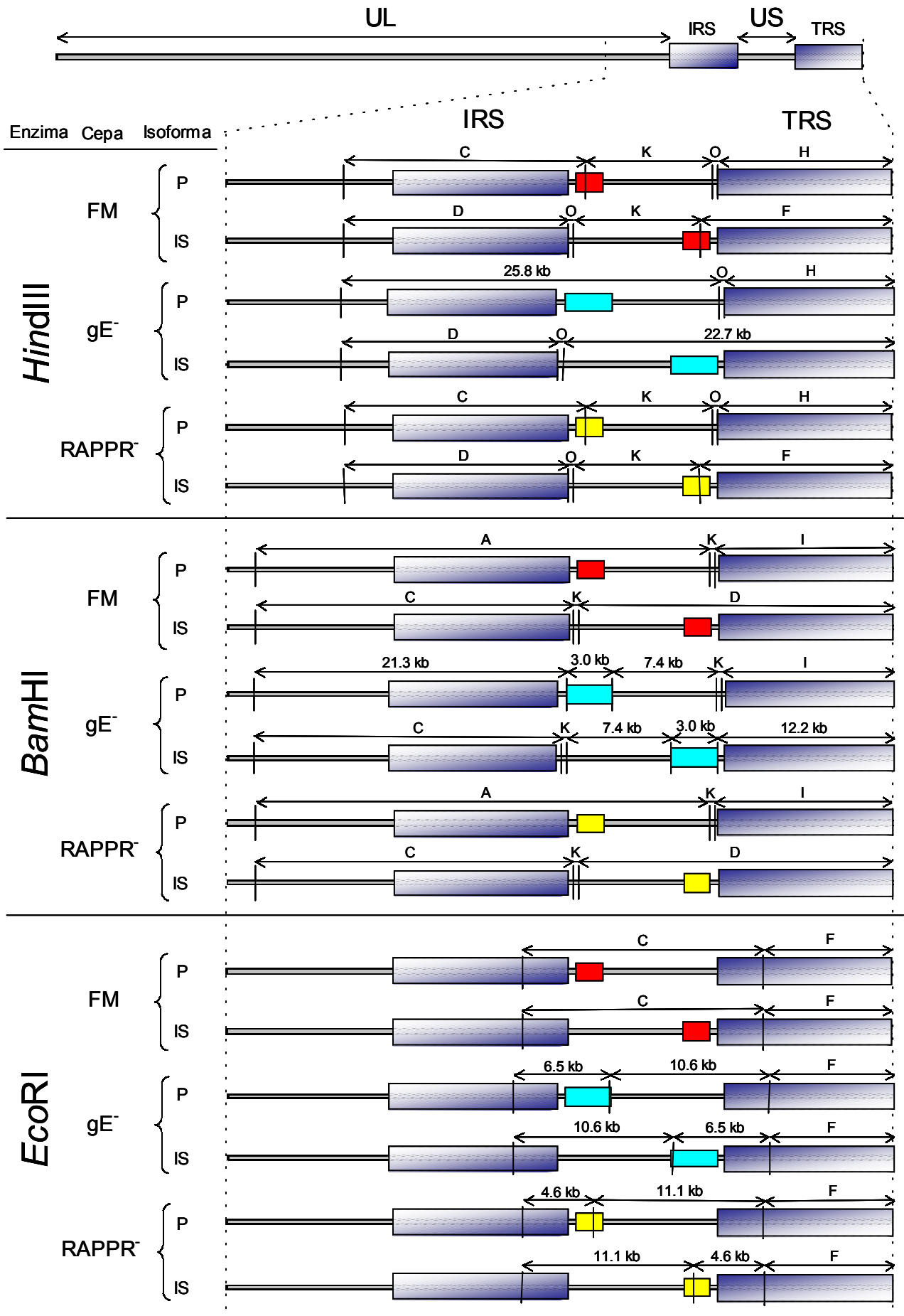
III.D.1.D. Caracterización del virus recombinante

El virus recombinante producto de la reincorporación del gen *gE* se ha caracterizado comparando su mapa de restricción genómico con el de las cepas víricas BoHV-1 *gE*<sup>-</sup> y BoHV-1 FM. También se ha llevado a cabo una transferencia Southern para confirmar la presencia y hallar de una manera más precisa la localización del gen *gE* reincorporado.

Los enzimas de restricción utilizados para caracterizar los genomas del virus recombinante defectivo para la secuencia TRAPP, del BoHV-1 *gE*<sup>-</sup> y del BoHV-1 FM, han sido *Hind*III, *Bam*HI y *Eco*RI. Cabe señalar que los cambios en el mapa de restricción del genoma de estos virus resultan sensiblemente magnificados pues afectan a la región US, la cual presenta dos orientaciones diferentes (P e IS) correspondientes a cada una de las isoformas mayoritarias del genoma del BoHV-1 (Mayfield *et al.*, 1983; Schynts *et al.*, 2003). El virus BoHV-1 *gE*<sup>-</sup>, construido a partir del BoHV-1 FM, tiene la mayor parte del marco de lectura del gen *gE* reemplazado por el del gen *lacZ*, en orientación invertida. En términos de dianas de restricción, esta construcción conlleva la desaparición de una diana *Hind*III y la aparición de dos nuevas dianas *Bam*HI y una



**Figura 52.** Caracterización genómica de las cepas FM, *gE*<sup>-</sup> y RAPPR<sup>-</sup> del BoHV-1. **(A)** Electroforesis de DNA genómico digerido con *Bam*HI, *Eco*RI e *Hind*III, teñida con bromuro de etidio. 'MM': marcador de peso molecular ( $\lambda$ BstEII +  $\lambda$  Monocut Mix, New England Biolabs). Los números a la derecha indican el tamaño, en kb, de los fragmentos de DNA doble hebra que componen el MM. Sobre la imagen se señalan las diferencias entre las tres cepas en cada una de las digestiones. La nomenclatura de bandas en los carriles FM es la establecida por J. E. Mayfield y colaboradores (Mayfield *et al.*, 1983) (Fig. 53). En los carriles *gE*<sup>-</sup> se señala el tamaño de los nuevos fragmentos de restricción que conlleva la sustitución del gen *gE* por el *lacZ*. En el carril RAPPR<sup>-</sup> de la digestión *Eco*RI se señala el tamaño de los nuevos fragmentos de restricción que conlleva la nueva diana *Eco*RI. Condiciones de electroforesis: gel de agarosa 0.5%, 22.5 V, 17 h. **(B)** Transferencia Southern del gel de electroforesis de **A**. La sonda utilizada corresponde al fragmento *Sal*I-*Hind*III de 997 pb del plásmido pXJ (Fig. 13A), marcada con DIG.



diana *EcoRI* (Fig. 53). Al comparar el patrón de restricción *HindIII* del BoHV-1 gE<sup>-</sup> y del BoHV-1 FM se constata la desaparición de las bandas *HindIII*-C (15.9 kb), *HindIII*-F (12.8 kb) y *HindIII*-K (8.4 kb) a la vez que se detecta la presencia de dos nuevas bandas de 25.8 kb y 22.7 kb (Fig. 52A). En el patrón de restricción *BamHI* desaparecen las bandas *BamHI*-A (29.7 kb) y *BamHI*-D (20.9 kb), a la vez que aparecen cuatro nuevas bandas de 21.3 kb, 12.2 kb, 7.4 kb y 3 kb (Fig. 52A). En el patrón de restricción *EcoRI* desaparece la banda *EcoRI*-C (15.7 kb) y aparecen dos nuevas bandas de 10.6 kb y 6.5 kb (Fig. 52A). El mapa de restricción del BoHV-1 gE<sup>-</sup> para los enzimas *HindIII*, *BamHI* y *EcoRI* concuerda pues con la incorporación del fragmento del gen *lacZ* reemplazando el gen *gE*, resultado que se corroboró mediante una transferencia e hibridación Southern (Rebordosa, 1996).

La estrategia utilizada para obtener un virus BoHV-1 defectivo en la secuencia TRAPP ha consistido en reemplazar el gen *lacZ* del BoHV-1 gE<sup>-</sup> por un gen *gE* que codifica para el cambio de secuencia de RAPPOR por SHEFP. En definitiva, la construcción de este nuevo virus debe reestablecer la organización del genoma del BoHV-1 FM con el gen *gE*, y con ello también su patrón de restricción genómico. Los resultados de restricción con *HindIII* o *BamHI* del DNA del virus obtenido y del BoHV-1 FM prueban lo dicho. (Fig. 52A, 53). Según el patrón de restricción *EcoRI* aparecen, en cambio, dos nuevas bandas de 11.1 kb y 4.6 kb, respectivamente (Fig. 52A, 53). En este punto cabe recordar que la secuencia SHEFP surge de fusionar la forma delecionada Ct del fragmento A5 contenida en el plásmido pUΔ46, con la forma delecionada Nt del plásmido pUΔ714A1, y que en esta operación se ha generado una nueva diana *EcoRI*. Así pues, los resultados obtenidos mediante restricción del DNA genómico indican que en el virus construido sí se ha reemplazado el gen *lacZ* por el gen que codifica para la gE que contiene la secuencia TRAPP.

Las restricciones de los diferentes DNA víricos han sido analizadas mediante transferencia Southern, utilizando como DNA sonda el fragmento *Sall-HindIII* de 997 pb del plásmido pXJ marcado con DIG y codificante para el dominio extracelular de la gE (Rebordosa, 1996) (Fig. 52B). Los resultados obtenidos revelan la presencia del gen *gE* en las bandas *HindIII*-K, *BamHI*-A y *BamHI*-D del BoHV-1 FM y del nuevo virus. También se detecta la presencia del gen *gE* en la banda *EcoRI*-C del BoHV-1 FM y en las bandas de 11.1 kb y 4.6 kb del virus obtenido. Por el contrario, el DNA sonda no hibrida con ningún fragmento de restricción del DNA del virus defectivo BoHV-1 gE<sup>-</sup>. Estos resultados confirman que el virus construido ha incorporado el gen *gE* que codifica para el cambio de secuencia de RAPPOR por SHEFP, y que su localización es análoga a la del gen *gE* en el virus salvaje BoHV-1 FM.

El análisis del patrón de restricción *HindIII*, *BamHI* y *EcoRI* y los resultados de la transferencia Southern prueban que el virus construido es un virus BoHV-1 FM que contiene un gen *gE* no codificante para la secuencia TRAPP. Este virus se ha llamado BoHV-1 RAPPOR<sup>-</sup> y se ha utilizado, posteriormente, en diferentes experimentos de infección *in vitro* para determinar la relevancia de la secuencia TRAPP en la gE.

◀ **Figura 53.** (Página 120) Diagrama de los mapas de restricción *BamHI*, *EcoRI* e *HindIII* de la región US de las cepas FM, gE<sup>-</sup> y RAPPOR<sup>-</sup> del BoHV-1. En la parte superior se muestra un esquema del genoma del BoHV-1. Se señalan las regiones UL y US, y las secuencias repetidas interna (IRS) y terminal (TRS) que flanquean US. Se ha obviado señalar las secuencias repetidas de apenas 24 pb que flanquean UL. El diagrama detallado de la región US muestra una caja roja (gen *gE*) para la cepa FM, una caja celeste (gen *lacZ*) para la cepa gE<sup>-</sup>, y una caja amarilla (gen *gE* no codificante para RAPPOR<sup>-</sup>) para la cepa RAPPOR<sup>-</sup>. Para cada enzima de restricción (*BamHI*, *EcoRI* e *HindIII*) se muestra en detalle la distribución de dianas (líneas verticales 'I') en la región US y secuencias IRS y TRS del genoma de las tres cepas del BoHV-1. Cada genoma, además, está representado en sus dos isoformas mayoritarias: P e IS. Sobre cada isoforma se identifican los fragmentos de restricción de acuerdo a la nomenclatura establecida por J. E. Mayfield y colaboradores (Mayfield *et al.*, 1983). Tomando como referencia la cepa FM, los fragmentos *HindIII* a analizar son el C, D, F, H, K y O; los fragmentos *BamHI* son el A, C, D, I y K; los fragmentos *EcoRI* son el C y F. Los fragmentos inalterados de las cepas gE<sup>-</sup> o RAPPOR<sup>-</sup> se identifican con la misma nomenclatura. Por el contrario, los cambios en el patrón de restricción de gE<sup>-</sup> y RAPPOR<sup>-</sup> están representados por el tamaño de los nuevos fragmentos.

### III.D.2. Caracterización *in vitro* del virus BoHV-1 RAPPR<sup>-</sup>

Los ensayos de infección *in vivo* llevados a cabo con el virus BoHV-1 gE<sup>-</sup> indicaban que éste presenta una virulencia muy reducida a la vez que mantiene la inmunogenicidad de la cepa FM del BoHV-1 (Rebordosa, 1996). Paralelamente, otros grupos de investigación publicaron unos resultados similares con otras cepas del BoHV-1 defectivas para el gen *gE*, y destacaban el hecho que la administración de dexametasona en animales vacunados no conllevaba la reactivación del virus. La razón aludida era que el virus no podía reactivarse porque no había alcanzado los puntos del sistema nervioso en los que establece latencia (Kaashoek *et al.*, 1994). Ciertamente, mediante técnicas de PCR se determinó que el ganglio trigémino de terneras infectadas con el virus defectivo contenía 2.5-10 veces menos copias de DNA viral que el de terneras infectadas con el virus salvaje (van Engelenburg *et al.*, 1994). Estos resultados apuntaban a que la disminución de la virulencia del BoHV-1 gE<sup>-</sup> podía deberse a una propagación deficiente del virus en el sistema nervioso del animal vacunado, y que esta propagación dependía de la gE.

Por otro lado, también se había descrito que el complejo gE/I del HHV-1 tenía actividad FcR, esto es, de unión a la fracción Fc de las IgG (Johnson *et al.*, 1988). Se postuló entonces que esta actividad protegía a los viriones de la neutralización mediada por anticuerpo y por complemento (Frank & Friedman, 1989). En el BoHV-1 gE<sup>-</sup>, la ausencia de la gE y de una hipotética actividad FcR explicarían la propagación deficiente del virus. Sin embargo, esta hipótesis fue rechazada pues la reducción de la cantidad de viriones BoHV-1 gE<sup>-</sup> secretados en los animales infectados era incluso anterior a la aparición de anticuerpos específicos anti-BoHV-1.

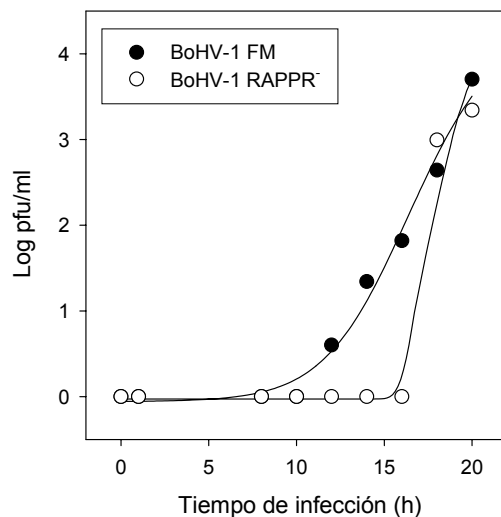
Finalmente, para caracterizar el nuevo virus BoHV-1 gE<sup>-</sup> y para determinar la importancia del sistema inmunológico en la disminución de su virulencia, se llevaron a cabo ensayos de infección *in vitro*. Estos ensayos fueron la infección de células en medio líquido, y la infección de células en medio semisólido. Los resultados obtenidos sugerían que, con independencia de unos posibles mecanismos de evasión inmunológica, la ausencia de la gE conllevaba una propagación más lenta del virus en el cultivo celular, y apuntaban por vez primera a que la gE del BoHV-1 participa en un mecanismo distinto de propagación de la infección vírica: la transmisión directa del virus 'célula-a-célula' (Rebordosa *et al.*, 1996). Este mecanismo, que ya había sido descrito para el SuHV-1 (Zsak *et al.*, 1992), sugería que el complejo gE/I promovía la fusión de la membrana plasmática de la célula infectada con la de sus células vecinas, de manera que los viriones propagaban la infección sin pasar por el medio extracelular. Actualmente se sabe que la transmisión directa 'célula-a-célula' es la vía de propagación *in vivo* más importante de los herpesvirus, y que el complejo gE/I tiene una implicación directa en ella pues promueve el ensamblaje de los viriones en las regiones próximas de la membrana plasmática desde donde inician la infección de las células adyacentes.

#### III.D.2.A. Cultivo de infección en medio líquido

Para caracterizar el virus BoHV-1 RAPPR<sup>-</sup> y comparar su cinética de crecimiento con la de la cepa salvaje FM, se han infectado dos monocapas de células MDBK de 175 cm<sup>2</sup> crecidas en medio líquido con

sendos virus a una multiplicidad de infección (m.o.i.) de 0.001, y se ha hallado la cantidad de viriones liberados al sobrenadante en distintos tiempos de infección.

Los resultados obtenidos indican que la cinética de producción de viriones BoHV-1 RAPPPr<sup>-</sup> en medio líquido es similar a la de viriones FM. La cantidad de viriones RAPPPr<sup>-</sup> hallada en el sobrenadante del cultivo a partir de las 20 primeras horas de infección equivale a la de viriones FM. Este resultado apunta a que la substitución de la secuencia RAPPPr de la gE no interfiere con la proliferación normal del BoHV-1. Sin embargo, sí se observan diferencias cuando se halla la cantidad de viriones en el sobrenadante del cultivo durante las 20 primeras horas de infección (Fig. 54). Desde las 12 hasta las 18 horas post-infección únicamente se contabilizan viriones FM, mientras que los viriones RAPPPr<sup>-</sup> aparecen sólo a partir de las 18 horas post-infección. Además, la cantidad de viriones RAPPPr<sup>-</sup> en el sobrenadante de infección tras 18 horas, y también en las horas posteriores de infección, equivale a la cantidad hallada de viriones FM. Estos resultados indican que no existen diferencias en la tasa de replicación de los virus BoHV-1 RAPPPr<sup>-</sup> y FM, pero sí en la tasa de liberación al medio de cultivo de los viriones de uno y otro tipo. El súbito aumento de título de los viriones RAPPPr<sup>-</sup> parece indicar que los viriones producidos ven ralentizada hasta entonces su salida al medio extracelular.



**Figura 54.** Cinética de producción de viriones BoHV-1 en medio líquido. Infección de células MDBK a una m.o.i. de 0.001.

Algunos autores han observado que la presencia de la gE conlleva una acumulación mayoritaria de los viriones en el espacio intercelular basolateral de los epitelios polarizados (Johnson *et al.*, 2001). Ésta es la base de la transmisión 'célula-a-célula' de los alphaherpesvirus. Los viriones BoHV-1 FM no son una excepción y se acumulan de forma mayoritaria en este espacio intercelular, al menos mientras dura el ciclo de infección (aprox. 18 horas). La aparición de viriones FM en el sobrenadante desde las 12 horas de infección es consecuencia de la saturación de este espacio y de un transporte intracelular deficiente, que llevan a algunos viriones a situarse en la superficie apical de la célula desde donde pasan al sobrenadante. En el caso de los viriones RAPPPr<sup>-</sup>, cabe interpretar que la substitución de la secuencia RAPPPr de la gE conlleva una acumulación más eficiente de los viriones en la superficie celular basolateral. También cabe pensar que dicha substitución induce un cambio conformacional que lleva a la gE, o al complejo gE/I, a interactuar con otras proteínas de la membrana plasmática o de la matriz extracelular con una mayor afinidad. Por otro lado, descartamos la idea de que la acumulación de viriones RAPPPr<sup>-</sup> pueda ser

intracelular, ya que en ningún caso se ha modificado el dominio intracelular de la gE responsable de localizar la proteína cerca de la membrana y, con ello, de promover la salida de los viriones. Finalmente, cabe apuntar que la cinética de liberación de los viriones BoHV-1 RAPPR<sup>-</sup> en un cultivo de infección es claramente opuesta a la de los viriones gE<sup>-</sup>. Éstos, que aparecen en el sobrenadante de infección hasta 4 horas antes que los viriones FM (Rebordosa *et al.*, 1996), tienen una presencia aún mayor en la superficie apical celular como consecuencia de una morfogénesis deslocalizada. La ausencia del dominio citoplasmático de la gE sitúa la fase final de la morfogénesis del virión gE<sup>-</sup> lejos de los dominios basolaterales de la célula. Igualmente, la ausencia del dominio extracelular de la gE priva al virión gE<sup>-</sup> de reconocer los complejos de unión de la célula infectada con las células adyacentes. La delección de la gE refleja, en cualquier caso, una capacidad ciertamente limitada de transmisión ‘célula-a-célula’ del BoHV-1.

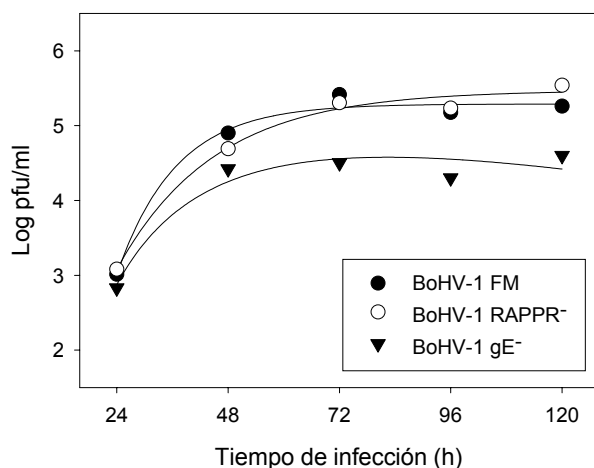
El cultivo de infección en medio líquido determina, pues, que no existen diferencias en la capacidad de proliferación de los virus BoHV-1 RAPPR<sup>-</sup> o FM, pero sí entre las respectivas pautas de liberación de los viriones. Los resultados obtenidos reflejan que la sustitución de la secuencia RAPPR no conlleva la pérdida de funcionalidad de la gE en la transmisión del virus. Por ello, ni las secuencias RAPPR o TRAPP, ni la región de la gE de la que ambas forman parte, parecen ser relevantes para mantener la función de la gE. Con todo, no obstante, la sustitución de la secuencia RAPPR sí puede haber modificado la interacción de los viriones a través de la gE con, p.e., los complejos de unión celular.

#### III.D.2.B. Cultivo de infección en medio semisólido

La caracterización del virus BoHV-1 RAPPR<sup>-</sup> en presencia de medio semisólido ha tenido como objeto comprobar si la sustitución de la secuencia de aminoácidos RAPPR de la gE altera la capacidad de transmisión ‘célula-a-célula’ del BoHV-1. Para ello, se han infectado tres monocapas de células MDBK crecidas en placas de cultivo de 60 mm de diámetro con virus BoHV-1 RAPPR<sup>-</sup>, FM o gE<sup>-</sup>, y se ha hallado la cantidad de viriones liberados y retenidos en el medio semisólido en distintos tiempos de infección. Para garantizar que esta cantidad de viriones refleja la capacidad de transmisión ‘célula-a-célula’ inherente de cada tipo de virus, se ha aplicado una m.o.i. inicial de 0.0001, 10 veces inferior a la aplicada en el cultivo de infección en medio líquido. En estas condiciones, y tras varias rondas de infección, los virus con una capacidad de transmisión menor habrán producido menos viriones. Por otro lado, también se ha teñido la monocapa celular con solución de Giemsa para comprobar cómo progresa la infección de cada tipo de virus.

A las 24 horas post-infección, únicamente son visibles las calvas formadas por los virus BoHV-1 RAPPR<sup>-</sup> y FM. El tamaño de las calvas RAPPR<sup>-</sup> es apenas superior al de las calvas FM. Por el contrario, las calvas que forma el virus defectivo BoHV-1 gE<sup>-</sup> sólo son visibles a partir de las 48 horas de infección, y su tamaño es marcadamente menor en comparación al de las calvas FM y RAPPR<sup>-</sup>. *A priori*, estos resultados reflejan que el virus BoHV-1 RAPPR<sup>-</sup> tiene una capacidad de transmisión y de producción equivalente a la del virus FM y superior a la del defectivo gE<sup>-</sup>. Ciertamente, la cantidad de viriones que se extrae del medio semisólido que cubre los cultivos infectados con el BoHV-1 RAPPR<sup>-</sup> es en todo momento muy parecida a la de viriones FM y, a partir de las 48 horas de infección, siempre superior a la de viriones gE<sup>-</sup> (Fig. 55). El título hallado a las 24 horas de infección, muy similar para los tres tipos de virus, es consecuencia directa de un número todavía reducido de rondas de infección. La cantidad de viriones gE<sup>-</sup> “deslocalizados” que ha pasado

al medio semisólido a las 24 horas de infección resulta ser parecido al “excedente” de viriones RAPPR<sup>-</sup> o FM que no ha podido situarse en los espacios basolaterales de las calvas respectivas. A partir de las 72 horas de infección, la cantidad producida de viriones RAPPR<sup>-</sup> o FM es hasta 7-9 veces superior a la de viriones gE<sup>-</sup>.



**Figura 55.** Cinética de producción de viriones BoHV-1 en medio semisólido. Infección de células MDBK a una m.o.i. de 0.0001.

Los resultados obtenidos indican claramente que en un cultivo crecido en medio semisólido, el virus BoHV-1 RAPPR<sup>-</sup> se propaga más y mejor que el defectivo gE<sup>-</sup>. En estas condiciones de cultivo, en las que se favorece la capacidad de transmisión ‘célula-a-célula’, el virus RAPPR<sup>-</sup> se transmite con una efectividad similar a la de la cepa salvaje FM. Estos resultados sugieren que, a pesar de la sustitución de la secuencia RAPPR, la gE de los viriones RAPPR<sup>-</sup> todavía ayuda a situarlos mayoritariamente en los espacios basolaterales de las células epiteliales infectadas. Desde aquí, los viriones FM y también los RAPPR<sup>-</sup> tienen más facilidades para infectar las células adyacentes propagando así mejor la infección. Los viriones gE<sup>-</sup>, por el contrario, se hallan más deslocalizados sobre la superficie celular y sólo algunos de ellos alcanza la superficie de la célula adyacente.

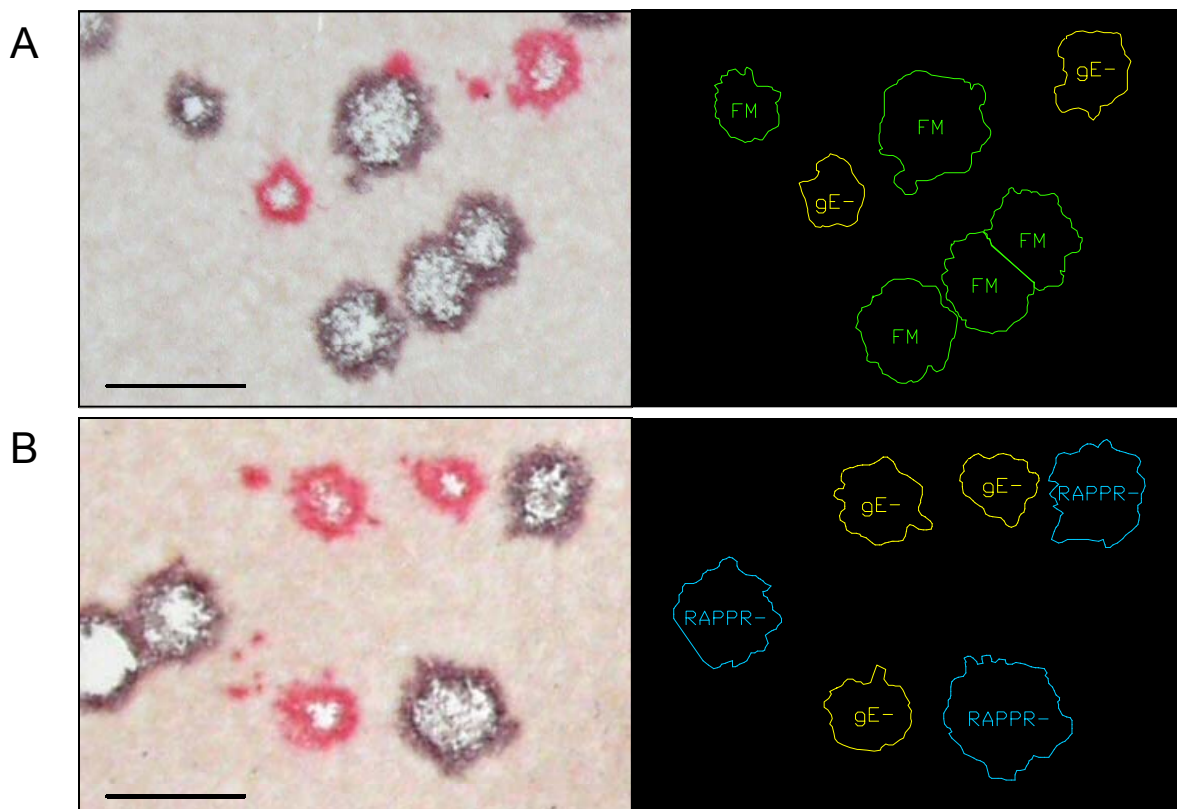
Tal y como se ha apuntado anteriormente, los dominios intra- y extracelulares de la gE del BoHV-1 desarrollan funciones distintas. El dominio intracelular determina el transporte basolateral de la gE y participa en la morfogénesis de los viriones. El dominio extracelular, en cambio, parece reconocer los complejos de unión de la célula infectada con las células adyacentes. En conjunto, ambas funciones determinan la capacidad de transmisión ‘célula-a-célula’ del BoHV-1. La cantidad de viriones hallada en un cultivo de infección en medio semisólido sugiere que la sustitución de la secuencia RAPPR de la gE no ha alterado la capacidad de transmisión del virus. Aparentemente, el cambio de secuencia en la gE no ha interferido, pues, en el transporte basolateral de la gE y tampoco en el reconocimiento de los complejos de unión celulares. Sin embargo, el resultado del cultivo de infección en medio líquido, sí asocia la sustitución de la secuencia RAPPR de la gE a una tasa de liberación diferente de los viriones al medio de cultivo. Cabe pensar, pues, que si bien la sustitución de la secuencia induce un pequeño cambio en las características de la gE, éste apenas repercute en su funcionalidad o en el nivel de proliferación del virus. De hecho, no se observa diferencias de título entre las cepas FM y RAPPR<sup>-</sup> en el cultivo de infección en medio líquido. Para comprobar si este cambio de secuencia de la gE afecta algún otro parámetro de la transmisión ‘célula-a-célula’ del



BoHV-1, se ha medido el área de las calvas que forman las cepas FM, RAPPR<sup>-</sup> y gE<sup>-</sup> en un cultivo de infección en medio semisólido.

### III.D.2.C. Determinación del área de calvas en medio semisólido

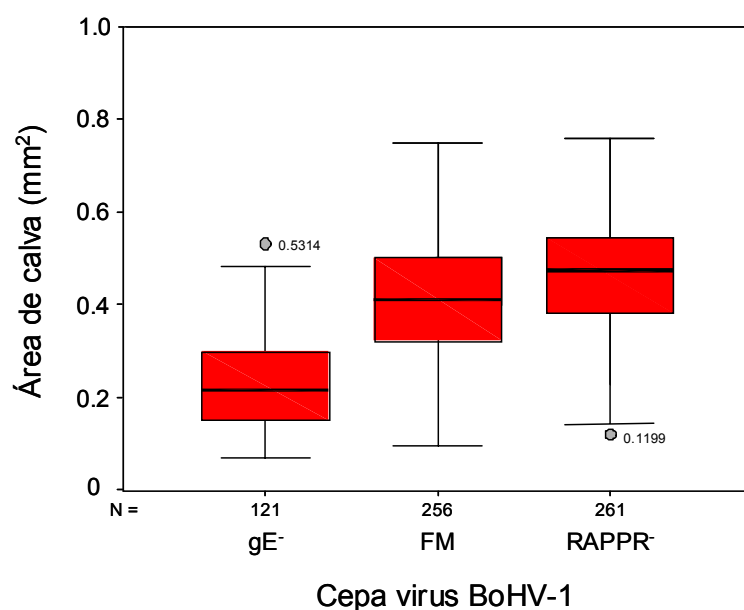
El área de la calva o núcleo de infección que forma un virión de alphaherpesvirus cuando infecta una monocapa celular crecida en medio semisólido, refleja la capacidad de transmisión 'célula-a-célula' del virus en cuestión (Rebordosa *et al.*, 1996). Esta transmisión, igual que la infección de la célula desde el medio extracelular, depende siempre de un conjunto de interacciones entre las proteínas de la envoltura del virión y las proteínas de la membrana plasmática de la célula a infectar. En el caso de la gE, parece probado que el dominio extracelular permite a los viriones reconocer los complejos de unión celular. La diferente tasa de liberación observada para el BoHV-1 RAPPR<sup>-</sup> sugiere que la sustitución de la secuencia RAPPR del dominio extracelular de la gE modifica, en parte, esta interacción con los complejos de unión o con otras proteínas de la matriz extracelular. Planteamos, pues, la medida del área de la calva RAPPR<sup>-</sup> y su comparación con la de las cepas FM y gE<sup>-</sup> como una primera aproximación para determinar si la secuencia RAPPR tiene alguna implicación en la transmisión 'célula-a-célula' del BoHV-1.



**Figura 56.** Transmisión 'célula-a-célula' del BoHV-1. Determinación del área de calva de las cepas FM, gE<sup>-</sup> y RAPPR<sup>-</sup> en medio semisólido. **(A)** Calvas FM (marrón) y gE<sup>-</sup> (rojo) sobre un cultivo de células MDBK a las 72 h post-infección. A la derecha se muestra el perímetro trazado de las calvas FM y gE<sup>-</sup> de la imagen. **(B)** Calvas RAPPR<sup>-</sup> (marrón) y gE<sup>-</sup> (rojo) sobre un cultivo de células MDBK a las 72 h post-infección. A la derecha se muestra el perímetro trazado de las calvas de la imagen. En **A** y **B**, células infectadas a una m.o.i. de 0.0001. Calvas gE<sup>-</sup> hibridadas con la sonda *lacZ*\*BIO y reveladas con sustrato Fast Red de la fosfatasa alcalina. Calvas FM y RAPPR<sup>-</sup> hibridadas con la sonda gE\*DIG y reveladas con sustrato NBT/BCIP de la fosfatasa alcalina. Barras 1mm.

La medición del área de calva se ha realizado sobre dos monocapas de células MDBK infectadas conjuntamente con viriones FM y gE<sup>-</sup>, o RAPPPr<sup>-</sup> y gE<sup>-</sup>, respectivamente, y sobre las que se ha realizado una hibridación *in situ* para detectar los tres tipos de calvas formados (Fig. 56A,B).

Un primer análisis determina que no existen diferencias de área significativas ( $P > 0.05$ ) entre las calvas gE<sup>-</sup> de las dos monocapas infectadas. Este resultado indica que las condiciones del cultivo en ambas placas han sido muy parecidas. En particular, el grado de confluencia celular en ambas monocapas ha sido muy similar y no ha condicionado el área de unas u otras calvas. Tras 72 horas de infección, el área media de la calva que forman los virus BoHV-1 FM, RAPPPr<sup>-</sup> y gE<sup>-</sup> es de 0.407 mm<sup>2</sup> (DE=0.124 mm<sup>2</sup>), 0.459 mm<sup>2</sup> (DE =0.126 mm<sup>2</sup>) y 0.233 mm<sup>2</sup> (DE =0.099 mm<sup>2</sup>), respectivamente (Fig. 57, Tabla XXII).



**Figura 57.** Diagrama de caja ('Boxplot') de la variable área de calva de las cepas FM, gE<sup>-</sup> y RAPPPr<sup>-</sup> del BoHV-1. Los tres tipos de calvas se han medido sobre un cultivo de células MDBK en medio semisólido a 72 h post-infección.

TABLA XXII. Parámetros descriptivos de la variable área de calva de las cepas FM, gE<sup>-</sup> y RAPPPr<sup>-</sup> del BoHV-1

Cepa	Número de casos	Mediana (mm <sup>2</sup> )	Máximo (mm <sup>2</sup> )	Mínimo (mm <sup>2</sup> )	Amplitud intercuartil (mm <sup>2</sup> )	Media (mm <sup>2</sup> )	Desviación estándar (DE) (mm <sup>2</sup> )	Intervalo de confianza (IC) del 95% de la media (mm <sup>2</sup> )
gE <sup>-</sup>	121	0.2163	0.5314	0.0718	0.1456	0.2329	0.0997	0.2150 a 0.2509
FM	256	0.4115	0.7492	0.0981	0.1783	0.4068	0.1243	0.3915 a 0.4221
RAPPPr <sup>-</sup>	261	0.4748	0.7590	0.1199	0.1654	0.4589	0.1258	0.4436 a 0.4743

El resultado del análisis estadístico (Análisis de la variancia con contrastes) determina que existen diferencias significativas ( $P < 0.05$ ) entre el área media de los tres tipos de calvas (Tabla XXIII). Las calvas FM tienen, en promedio, un área 0.174 mm<sup>2</sup> mayor (IC 95%: 0.150 a 0.197 mm<sup>2</sup>) que las calvas gE<sup>-</sup>. Así, las calvas FM resultan ser un 74.6% mayores que las calvas gE<sup>-</sup>. Las calvas RAPPPr<sup>-</sup> tienen un área 0.226 mm<sup>2</sup> mayor (IC 95%: 0.202 a 0.249 mm<sup>2</sup>) que las calvas gE<sup>-</sup>, esto es, son hasta un 96.9% más grandes que las

gE<sup>-</sup>. Finalmente, las calvas RAPPPr<sup>-</sup> ocupan un área 0.052 mm<sup>2</sup> mayor (IC 95%: 0.030 a 0.074 mm<sup>2</sup>) que las calvas FM, y son por ello hasta un 12.7% mayores.

TABLA XXIII. Resumen del análisis de la variancia con contrastes del área media de las calvas FM, gE<sup>-</sup> y RAPPPr<sup>-</sup> del BoHV-1

Contrastes	Estadístico de contraste (t)	Diferencia de medias (mm <sup>2</sup> )	Error estándar (EE) de la diferencia de medias (mm <sup>2</sup> )	Intervalo de confianza (IC) del 95% de la diferencia de medias (mm <sup>2</sup> )
FM vs gE <sup>-</sup>	14.564	0.1739*	0.0119	0.1503 a 0.1974
RAPPPr <sup>-</sup> vs gE <sup>-</sup>	18.911	0.2260*	0.0119	0.2024 a 0.2495
RAPPPr <sup>-</sup> vs FM	4.736	0.0521*	0.0110	0.0305 a 0.0737

\*La diferencia entre las medias es significativa con un grado de significación P<0.05

Los resultados obtenidos indican que la capacidad de transmisión ‘célula-a-célula’ de las cepas FM, RAPPPr<sup>-</sup> y gE<sup>-</sup> del BoHV-1 difieren entre sí. La cepa gE<sup>-</sup> forma calvas un 42% y 49% menores, respectivamente, que las de las cepas FM y RAPPPr<sup>-</sup>. La cepa gE<sup>-</sup> presenta, pues, una capacidad de transmisión ‘célula-a-célula’ marcadamente menor. Esta menor capacidad de transmisión se explica por la deslocalización de los viriones de los espacios intercelulares basolaterales inducida por la ausencia de la gE, y justifica la menor proliferación observada en los cultivos de infección líquido y semisólido. La cepa RAPPPr<sup>-</sup> es, por el contrario, la cepa que forma las calvas más grandes —un 12.7% mayores que las de la cepa salvaje FM— y presenta por ello una capacidad de transmisión ‘célula-a-célula’ mayor. Esta mayor capacidad de transmisión, conjuntamente con la menor tasa de liberación de viriones hallada en el cultivo de infección en medio líquido, sugieren que la localización basolateral de los viriones RAPPPr<sup>-</sup> es más efectiva que la de los viriones FM. Cabe señalar, no obstante, que estos resultados son igualmente congruentes con el hecho que los viriones RAPPPr<sup>-</sup> infectan con mayor rapidez las células adyacentes a la célula infectada sin necesidad de acumularse en el espacio basolateral. Esta posibilidad se basa en el hecho de que en un cultivo de infección en medio semisólido, tras 24-36 horas de infección y a simple vista, las calvas RAPPPr<sup>-</sup> ya son ligeramente mayores que las calvas FM. En este caso se puede interpretar que mientras la mayoría de los viriones FM se acumulan en el espacio basolateral y algunos incluso se deslocalizan hacia la superficie apical de la célula, los viriones RAPPPr<sup>-</sup> ya infectan las células contiguas.

El hecho que la substitución de la secuencia de aminoácidos RAPPPr de la gE del BoHV-1 no conlleve la pérdida de funcionalidad de la proteína, indica que las secuencias de aminoácidos RAPPPr y TRAPP, y muy probablemente el dominio del que ambas forman parte, no están directamente implicados en la función de la gE. El incremento observado en la capacidad de transmisión ‘célula-a-célula’ del BoHV-1 RAPPPr<sup>-</sup> puede deberse a un cambio en la interacción de la gE con los complejos de unión o con otras proteínas de la matriz extracelular. Así pues, la secuencia de aminoácidos TRAPPPr y el dominio de la gE del que forma parte, juegan un papel todavía por determinar en la estructura/función de esta glicoproteína.

### **III.E. Obtención de anticuerpos monoclonales anti-gE y diseño de un test ELISA de bloqueo**

La vacunación contra el BoHV-1, si bien contribuye a reducir la infección y los signos clínicos derivados de la misma, no elimina la posibilidad de futuras infecciones por BoHV-1 en el animal vacunado. Por ello, en varios países, el control del BoHV-1 pasa por desarrollar programas de erradicación basados en la identificación serológica y sacrificio de los animales infectados. Para identificar qué animales han sido infectados después de la vacunación, generalmente se aplica un test ELISA de bloqueo que detecta la presencia de anticuerpos dirigidos contra la/s proteína/s del BoHV-1 ausentes en la vacuna marcadora administrada. Nuestro grupo de investigación ha diseñado un test ELISA de bloqueo para utilizarlo conjuntamente a la aplicación de la cepa vacunal defectiva BoHV-1 gE<sup>-</sup>. Para ello, se ha obtenido una pequeña colección de anticuerpos monoclonales (MAbs) anti-gE.

#### **III.E.1. Caracterización de los anticuerpos monoclonales**

Para obtener los MAbs se han llevado a cabo tres fusiones independientes con esplenocitos procedentes de ratones Balb C inmunizados con BoHV-1 FM y células de mieloma NSI-D8. En total se contabilizaron 652 pocillos de ELISA con uno o más hibridomas, pero tan sólo en 86 de estos pocillos (13.2%) se detectó la producción de anticuerpos anti-BoHV-1 mediante ELISA indirecto. Este porcentaje relativamente bajo de pocillos positivos anti-BoHV-1 permite asegurar, primero, que la mayoría de ellos contiene un único hibridoma anti-BoHV-1 y, segundo, que al no coincidir más de un hibridoma por pocillo no pasarán desapercibidos aquéllos que produzcan anticuerpos anti-gE. Por otro lado, y dado que el clonaje de un solo hibridoma puede prolongarse durante semanas o meses (Harlow & Lane, 1988), se ha preferido posponer el clonaje de los hibridomas obtenidos hasta realizar un primer análisis que demostrase la idoneidad de algún posible candidato.

A partir de los sobrenadantes de los 86 pocillos positivos, se han identificado los posibles hibridomas anti-gE. Esta identificación se ha realizado mediante un ELISA indirecto que utiliza el virus purificado BoHV-1 gE<sup>-</sup> (Rebordosa *et al.*, 1995) como antígeno a reconocer. Con sólo 5 de los 86 sobrenadantes ensayados (1H12, 1E2, 2C9, 3F9 y 2D4) se obtuvieron valores de  $A_{405}$  iguales o muy próximos al control negativo del ensayo (pocillos sin BoHV-1 gE<sup>-</sup>) (Tabla XXIV). Dado que el BoHV-1 gE<sup>-</sup> no contiene la gE, la ausencia de respuesta al utilizar estos sobrenadantes indica que el hibridoma correspondiente produce un MAb dirigido contra la gE del BoHV-1. Por otro lado, la baja frecuencia de hibridomas anti-gE en el panel obtenido (5.8%) es acorde con el amplio repertorio antigénico del virión del BoHV-1.

Se ha valorado, asimismo, la capacidad de los 5 MAbs anti-gE para detectar un posible bloqueo del epítipo que reconoce cada uno de ellos. Para ello, se ha ensayado un ELISA gE de bloqueo en el que el antígeno adsorbido (BoHV-1 FM completo) ha sido incubado con suero de animales infectados por BoHV-1 FM (sueros gE<sup>+</sup>) antes de aplicar el MAb anti-gE (Tabla XXIV). Como control negativo, el antígeno se ha incubado con suero de animales vacunados con una dosis de BoHV-1 gE<sup>-</sup> (sueros gE<sup>-</sup>). Los resultados obtenidos son muy dispares. Al utilizar el sobrenadante 2D4 no se detecta ningún tipo de bloqueo, ni con

TABLA XXIV. Valoración mediante ELISA indirecto y ELISA de bloqueo de los hibridomas anti-gE

Tipo de ensayo	1H12	1E2	2C9	3F9	2D4
ELISA indirecto <sup>(a)</sup>					
A <sub>405</sub> BoHV-1 FM	2.398	0.334	2.627	0.497	0.531
A <sub>405</sub> BoHV-1 gE <sup>-</sup>	0.102	0.101	0.115	0.084	0.106
ELISA de bloqueo <sup>(b)</sup>					
Bloqueo suero gE <sup>+</sup> (%) <sup>(c)</sup>	83	65	74	81	10
Bloqueo suero gE <sup>-</sup> (%) <sup>(c)</sup>	84	57	54	48	1
Diferencia <sup>(d)</sup>	-1	8	20	33	9

<sup>(a)</sup> En el ELISA indirecto, los pocillos sensibilizados con antígeno de la cepa BoHV-1 FM (wt) o BoHV-1 gE<sup>-</sup> (defectivo) han sido incubados con el sobrenadante de los hibridomas obtenidos. En la tabla se muestran los hibridomas con cuyo sobrenadante se han registrado los valores de A<sub>405</sub> más bajos frente al antígeno BoHV-1 gE<sup>-</sup>, comparables al registro en los pocillos control (sin antígeno) (A<sub>405</sub>=0.073-0.092). Método: 1) adsorción del antígeno en tampón carbonato 0.2M pH 9.6, o/n 4°C, 2) aspirar contenido de los pocillos y escurrir la placa, 3) incubar los pocillos con 300 µl solución de bloqueo (PBS 1x 1.5% (p/v) Albumin Fraction V), mínimo 2 h a 37°C, 4) aspirar contenido de los pocillos y escurrir la placa, 5) dispensar 50 µl del sobrenadante de hibridoma en cada pocillo, 6) incubar 1 h a 37°C, 7) por 3 veces, lavar los pocillos con 300 µl PBS 1x 0.02% (v/v) Tween<sup>®</sup>20, y escurrir la placa, 8) dispensar, en cada pocillo, 50 µl de dilución según fabricante de anticuerpo secundario anti-mouse/AP en solución de bloqueo, 9) incubar 1 h a 37°C, 10) repetir paso 7, 11) revelar la placa, parar la reacción y leer la A<sub>405</sub>.

<sup>(b)</sup> En el ELISA de bloqueo, los pocillos han sido sensibilizados exclusivamente con antígeno de la cepa BoHV-1 FM. Antes de incubar el antígeno con el sobrenadante de los hibridomas, los pocillos han sido incubados 1 h a 37°C con suero de animales seropositivos para el BoHV-1 (sueros gE<sup>+</sup>), o bien de animales vacunados con el BoHV-1 gE<sup>-</sup> y por tanto seronegativos para la gE (sueros gE<sup>-</sup>).

<sup>(c)</sup> El % de bloqueo está calculado según la fórmula  $100 \times ((A_{405}bl - A_{405}sp) / A_{405}sp)$ , donde A<sub>405</sub>bl es el valor medio de absorbancia a 405 nm de tres pocillos incubados con solución de bloqueo (bl) durante el tiempo en el que el resto de los pocillos eran incubados con los sueros problema (sp).

<sup>(d)</sup> Resultado de restar el % de bloqueo del suero gE<sup>-</sup> del % de bloqueo del suero gE<sup>+</sup>.

sueros gE<sup>+</sup> ni con sueros gE<sup>-</sup>, lo cual sugiere que los animales infectados por BoHV-1 FM no generan anticuerpos contra el epítipo que reconoce el MAb 2D4. Por el contrario, los MAbs 1H12, 1E2, 2C9 y 3F9 detectan claramente el bloqueo de su epítipo producido por los sueros gE<sup>+</sup>. Cabe señalar, no obstante, que en estos casos los sueros gE<sup>-</sup> también producen un mayor o menor grado de bloqueo de los epítopos respectivos. El caso extremo es el del MAb 1H12 cuyo reconocimiento se ve igualmente bloqueado por uno u otro tipo de suero. Este hecho se puede explicar por la existencia de algún tipo de impedimento estérico producto de la unión de anticuerpos no anti-gE a epítopos próximos a los reconocidos por estos MAbs. A modo de ejemplo, y dado que en el virus la gE forma el complejo gE/I, cabe considerar que ciertos anticuerpos anti-gI pueden bloquear la unión de algunos MAbs anti-gE. Por otro lado, los MAbs 2C9 y 3F9 sí muestran una cierta diferencia -20 y 33 puntos, respectivamente- y por consiguiente un cierto grado de discriminación entre el bloqueo inducido por ambos tipos de suero. Es por ello que entre los MAbs anti-gE obtenidos, se ha seleccionado en primera instancia el 2C9 y el 3F9 como posibles candidatos para ser utilizados en un test ELISA gE de bloqueo.

Para optimizar el poder discriminatorio de los MAbs 2C9 y 3F9 se ha seguido dos estrategias distintas: i) el aumento de la cantidad de antígeno en placa, y ii) el tratamiento del antígeno con un detergente no iónico. El aumento de antígeno adsorbido en el pocillo conlleva, para ambos MAbs, un incremento insuficiente de la diferencia en el bloqueo entre sueros gE<sup>+</sup> y gE<sup>-</sup> (Tabla XXV). En el caso del MAb 3F9, el aumento de 10<sup>6</sup> a 10<sup>7</sup> pfu/pocillo conlleva un incremento de 35 a 49 puntos de la diferencia en el bloqueo. A pesar de la mejoría, el poder discriminatorio del MAb 3F9 sigue siendo pobre si se compara con el que muestran sistemas de diagnóstico parecidos de aplicación para otros herpesvirus. Por el contrario, el tratamiento del antígeno con el detergente no iónico Nonidet P40 sí parece mejorar sustancialmente el poder discriminatorio

del anticuerpo en el caso del MAb 3F9 (Tabla XXVI). La diferencia entre los sueros gE<sup>+</sup> y gE<sup>-</sup> en el porcentaje de bloqueo cuando se utiliza el MAb 3F9 se eleva a casi 80 puntos, mientras que para el MAb 2C9 permanece casi invariable. Este incremento de la diferencia en el bloqueo asociado a la utilización del MAb 3F9 se debe, esencialmente, a una reducción drástica en el porcentaje de bloqueo de los sueros gE<sup>-</sup>. Se ha observado, además, que el tratamiento con detergente comporta un incremento de 0.3-0.4 puntos en el registro de A<sub>405</sub> cuando se ensaya el bloqueo de sueros gE<sup>-</sup>. En principio, estos resultados apoyan el argumento de los impedimentos estéricos como responsables del bloqueo producido por los sueros gE<sup>-</sup>. Se ha descrito que el Nonidet P-40 solubiliza la envoltura vírica (van Drunen Littel-van den Hurk & Babiuk, 1986), si bien a la concentración utilizada no constituye un agente desnaturalizante. Cabe suponer que la disgregación que induce el detergente aumenta la accesibilidad del epítopo del MAb 3F9 a la vez que conlleva un alejamiento de epítopos próximos disminuyendo así las posibilidades de un impedimento estérico. Por otro lado, se ha descrito que el complejo gE/I se mantiene intacto en presencia de detergentes no iónicos (Whitbeck *et al.*, 1996), lo cual apunta a que el bloqueo de los sueros gE<sup>-</sup> en el caso del MAb 3F9 no viene dado por un epítopo residente en la gl. Éste, en cambio, sí podría ser el caso del MAb 2C9 para el que el tratamiento con Nonidet P40 no conlleva ninguna mejora significativa en su poder de discriminación. Finalmente, y en base a estos resultados, se ha seleccionado el MAb 3F9 como el mejor de los MAbs obtenidos procediendo así a la clonación del hibridoma que lo produce. Asimismo, las condiciones optimizadas, y desde entonces practicadas, del ELISA de bloqueo determinan la dosificación de 10<sup>7</sup> pfu/pocillo en tampón carbonato suplementado con Nonidet P-40 al 0.1%.

TABLA XXV. Optimización de la cantidad de antígeno en el ELISA gE de bloqueo

Hibridoma	Antígeno BoHV-1 FM (pfu/pocillo)	Bloqueo suero gE <sup>+</sup> (%)	Bloqueo suero gE <sup>-</sup> (%)	Diferencia
2C9	10 <sup>6</sup>	72	58	14
	3 x 10 <sup>6</sup>	70	52	18
	10 <sup>7</sup>	69	42	27
3F9	10 <sup>6</sup>	82	47	35
	3 x 10 <sup>6</sup>	81	40	41
	10 <sup>7</sup>	80	31	49

Se ha adsorbido cantidades crecientes de BoHV-1 FM para maximizar la diferencia entre los porcentajes de bloqueo de los MAbs 2C9 y 3F9 frente a sueros gE<sup>+</sup> y sueros gE<sup>-</sup>. La adsorción del antígeno se ha desarrollado en tampón carbonato. El ELISA gE de bloqueo se ha llevado a cabo como se describe en la tabla XXIV.

TABLA XXVI. Optimización de la formulación del antígeno en el ELISA gE de bloqueo

Hibridoma	Antígeno BoHV-1 FM (pfu/pocillo)	Bloqueo suero gE <sup>+</sup> (%)	Bloqueo suero gE <sup>-</sup> (%)	Diferencia
2C9	3 x 10 <sup>6</sup>	79	58	21
3F9	10 <sup>6</sup>	83	10	73
	3 x 10 <sup>6</sup>	85	8	77
	10 <sup>7</sup>	88	9	79

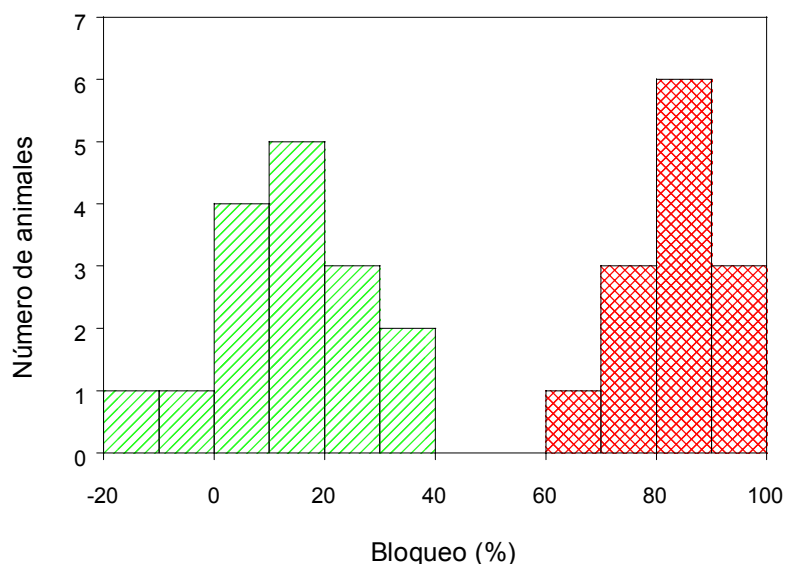
Se ha adsorbido cantidades crecientes de BoHV-1 FM para maximizar la diferencia entre los porcentajes de bloqueo de los MAbs 2C9 y 3F9 frente a sueros gE<sup>+</sup> y sueros gE<sup>-</sup>. La adsorción del antígeno se ha desarrollado en tampón carbonato suplementado con detergente no iónico Nonidet P-40 0.1%. El ELISA gE de bloqueo se ha llevado a cabo como se describe en la tabla XXIV.

### III.E.2. Establecimiento de los umbrales y puntos de corte

Para establecer los valores umbrales y puntos de corte del ELISA gE de bloqueo que utiliza el MAb 3F9, se ha analizado una muestra de 29 sueros de campo. Entre éstos, 16 ya habían sido caracterizados como negativos para el BoHV-1 (BoHV-1<sup>-</sup>, procedentes de terneras no infectadas), mientras que los 13 restantes se han caracterizado como positivos mediante ELISA indirecto (BoHV-1<sup>+</sup>, procedentes de terneras infectadas).

Todos los sueros probados BoHV-1<sup>-</sup> han resultado ser negativos también para el ELISA gE de bloqueo (sueros gE<sup>-</sup>), y producen una media de bloqueo del 16.8% (DE=13.9%) con respecto al registro de A<sub>405</sub> cuando no hay bloqueo alguno. Los sueros BoHV-1<sup>+</sup>, por su parte, han resultado ser sueros gE<sup>+</sup> y producen un bloqueo del 81.7% (DE=9.3%) (Fig. 58). Si consideramos que la distribución del porcentaje de bloqueo de uno y otro tipo de suero siguen una distribución normal, el valor de la media  $\pm 3x$ DE define el intervalo que comprende el 99% de los valores de cada una de las poblaciones. Así, tomando el conjunto de sueros gE<sup>-</sup>, calculamos que éstos pueden producir un bloqueo de hasta el 58.5%, mientras que los gE<sup>+</sup> serían aquéllos que producen un bloqueo superior al 58.5%. Igualmente, y tomando el conjunto de sueros gE<sup>+</sup>, calculamos que éstos son los que producen un bloqueo superior al 53.8%, mientras que los gE<sup>-</sup> serían los que producen un bloqueo inferior al 53.8%. Estos resultados indican que el ELISA gE de bloqueo que utiliza el MAb 3F9 registra una región de solapamiento entre 53.8% y 58.5% de bloqueo en la que cabe la duda de si el suero problema es realmente positivo (infectado, gE<sup>+</sup>) o negativo (no infectado, gE<sup>-</sup>). Consideramos que la existencia de esta región puede deberse, en parte, al escaso número de sueros (29) con los que se ha realizado esta prueba. Cabe suponer, que un ensayo con un número mayor de muestras permitiría establecer unos intervalos de bloqueo por suero positivo y negativo más precisos, y que incluso podría reducirse la amplitud de la región dudosa o de solapamiento entre ambos tipos de bloqueo.

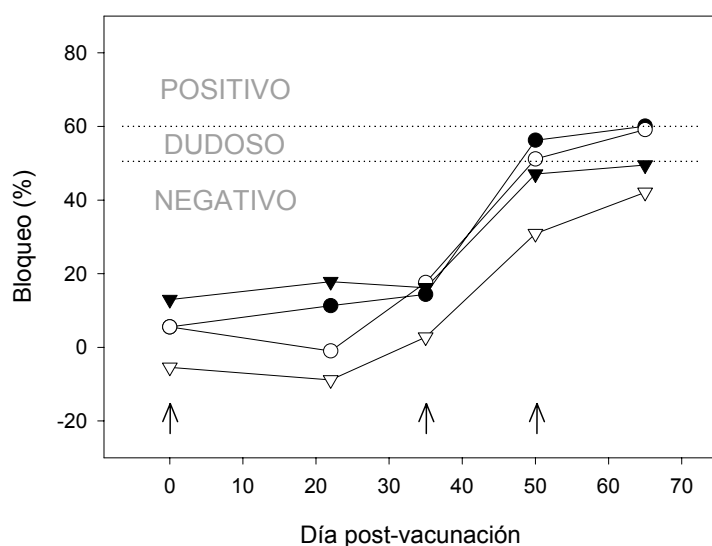
En base a los resultados obtenidos, y reconociendo que el ELISA gE de bloqueo que utiliza el MAb 3F9 está todavía en una fase experimental, concluimos de manera provisional que el test define como negativo aquel suero que produce un bloqueo inferior al 50%, dudoso (positivo o negativo) el suero que produce un bloqueo entre el 50-60%, y positivo el que produce un bloqueo superior al 60%.



**Figura 58.** Valores de corte del ELISA gE de bloqueo. La figura representa la distribución de frecuencias del % de bloqueo del MAb 3F9 obtenido con 29 sueros de campo: 16 sueros negativos para el BoHV-1 o sueros gE<sup>-</sup> (en verde) y 13 sueros positivos para el BoHV-1 o sueros gE<sup>+</sup> (en rojo).

### III.E.3. Especificidad del ELISA gE de bloqueo

Una de las pruebas a realizar para valorar la especificidad del ELISA de bloqueo consiste en analizar el suero procedente de animales a los que se ha administrado la vacuna marcadora en repetidas ocasiones. En este caso, la vacuna administrada ha sido la cepa defectiva BoHV-1 gE<sup>-</sup>. Cabe esperar que cuanto mayor es el número de revacunaciones, mayor es el título de anticuerpos circulantes anti-BoHV-1 en el suero del animal. Por otro lado, y dado que la vacuna no contiene la gE, no debería aumentar la capacidad del suero para interferir en la unión del MAb 3F9 sobre la gE. No obstante, y en alusión a los impedimentos estéricos, cabe suponer que un título muy elevado de anticuerpos contra otras glicoproteínas del BoHV-1 sí pueda bloquear de manera inespecífica el reconocimiento del MAb 3F9 en el ELISA gE de bloqueo, provocando así la aparición de falsos positivos. Para comprobar esta hipótesis, se ha administrado tres dosis de la vacuna BoHV-1 gE<sup>-</sup> a cuatro terneras en los días 0, 35 y 50, y se ha analizado la capacidad de interferencia del suero en el ELISA gE de bloqueo (Fig. 59).



**Figura 59.** Especificidad del ELISA gE de bloqueo. Valoración del % de bloqueo del MAb 3F9 con el suero de 4 terneras vacunadas repetidamente con el BoHV-1 gE<sup>-</sup>. Las flechas señalan los días 0, 35 y 50 correspondientes a la 1<sup>a</sup>, 2<sup>a</sup> y 3<sup>a</sup> vacunación. Se representa también las tres zonas a las que se adscribe el resultado del ELISA gE de bloqueo y que caracterizan el suero valorado como positivo, dudoso o negativo de infección por BoHV-1.

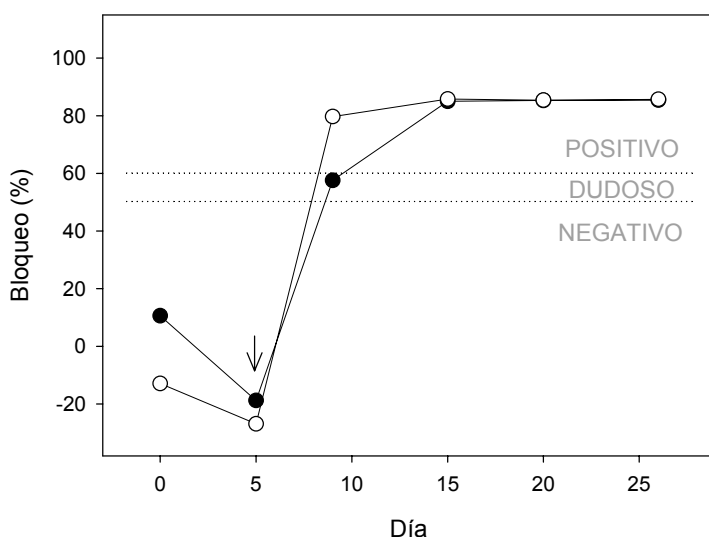
Los resultados obtenidos indican que el porcentaje de bloqueo de los cuatro sueros se mantiene bajo (<20%) durante los 35 días siguientes a la primera vacunación. A los 15 días de la segunda vacunación, se detecta un incremento importante de la capacidad de interferencia de los antisueros producidos (35-55% de bloqueo). A los 15 días de la tercera vacunación, esta capacidad de bloqueo aumenta todavía más y se sitúa en la zona de no discriminación entre sueros positivos y negativos del MAb 3F9. El incremento de la capacidad de bloqueo de los antisueros de animales vacunados parece no ser un hecho infrecuente. El incremento del bloqueo posterior a la vacunación repetida, se observa también en al menos uno de los test ELISA para la detección de anticuerpos contra la gE del SuHV-1 (van Oirschot *et al.*, 1988; van Oirschot *et al.*, 1990). En estos ensayos se obtienen resultados de bloqueo incluso muy superiores a los aquí descritos. Cabe señalar, no obstante, que estos niveles tan elevados de bloqueo tienden a disminuir si se sigue el régimen de administración de la vacuna gE<sup>-</sup>. Estos incrementos son, por consiguiente, transitorios y aparecen generalmente tras las primeras vacunaciones (van Oirschot *et al.*, 1990; van Oirschot *et al.*, 1997).



Como ya se ha indicado, el bloqueo al reconocimiento de la gE puede venir dado por el impedimento estérico que conlleva la unión de otros anticuerpos a epítomos próximos a la gE. Por otro lado, la transitoriedad del incremento del bloqueo descarta que dicha interferencia sea causada por IgGs ya que el título de éstas no disminuye tras una cuarta, quinta o sexta vacunación. Los resultados obtenidos son más acordes con un impedimento estérico producido p.e. por IgMs. Éstas son moléculas pentaméricas de elevada masa molecular que podrían inducir un bloqueo secundario en epítomos relativamente alejados de aquél que en realidad reconocen. Además, la producción temporal de IgMs correlaciona bastante bien con los períodos de tiempo en los que se produce este bloqueo inespecífico, y con el hecho que éste es transitorio. Finalmente, cabe señalar que si el bloqueo a la unión del MAb 3F9 se debe realmente a una interferencia producida por IgMs, un régimen de vacunación más espaciado —que es el que se aplica en realidad en explotaciones ganaderas— contribuiría a reducir la aparición de falsos positivos.

#### III.E.4. Sensibilidad del ELISA gE de bloqueo

Uno de los criterios que permite valorar la sensibilidad de un ensayo ELISA es la rapidez con la que se detecta la respuesta inmune después de una infección experimental. Para valorar la sensibilidad del ELISA gE de bloqueo, se ha analizado la capacidad de bloqueo del antisuero producido por dos terneras infectadas experimentalmente con virus BoHV-1 FM (Fig. 60). Los resultados obtenidos reflejan que el ELISA gE de bloqueo detecta la seroconversión de ambas terneras a los cuatro días post-infección, si bien una de ellas tendría todavía un diagnóstico dudoso. Diez días después de la infección, los valores de bloqueo de ambos antisueros alcanzan un valor máximo y se mantienen estables hasta el final del experimento. Estos resultados contrastan p.e. con los que se obtienen habitualmente mediante ELISA indirecto que utiliza viriones completos adsorbidos, y que detectan la seroconversión aproximadamente a los diez días post-infección (E. España, comunicación personal). La sensibilidad del ELISA gE de bloqueo basado en un MAb resulta pues igual, si no superior, a la de un ELISA indirecto que aplica viriones completos como antígeno a detectar. Por otro lado, la elevada sensibilidad del ELISA gE de bloqueo que utiliza el MAb 3F9 sugiere que el epítipo reconocido es inmunodominante, no sólo en la gE, sino en el virión en su conjunto.



**Figura 60.** Sensibilidad del ELISA gE de bloqueo. Valoración del % de bloqueo del MAb 3F9 con el suero de 2 terneras infectadas experimentalmente con BoHV-1 FM. La flecha señala un único día de infección. Se representa también las tres zonas a las que se adscribe el resultado del ELISA gE de bloqueo y que caracterizan el suero valorado como positivo, dudoso o negativo de infección por BoHV-1.

En conclusión, la obtención del MAb 3F9 ha permitido desarrollar un modelo de lo que podría ser un ELISA de bloqueo sensible y específico para la detección de anticuerpos contra la gE del BoHV-1. El futuro desarrollo del test deberá contemplar, no obstante, la utilización de un mayor número de sueros procedentes de animales infectados y de animales vacunados que permita establecer unos márgenes de confianza más fiables para el diagnóstico. Convendría, asimismo, que estos sueros tuvieran procedencias geográficas diversas para valorar así la variabilidad del epítipo que reconoce el MAb 3F9 entre diferentes poblaciones del BoHV-1. Cabe señalar, no obstante, que los epítipos de la gE y del complejo gE/I hasta ahora estudiados permanecen conservados en todos los aislados analizados del BoHV-1 (Rijsewijk *et al.*, 2000), y que por consiguiente no se espera tener problemas derivados de la aparición de falsos negativos durante el desarrollo de un futuro test de diagnóstico basado en el MAb 3F9.

