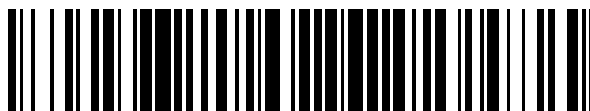


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 696 235**

51 Int. Cl.:

C12N 1/20 (2006.01)
C12N 1/21 (2006.01)
C12N 1/22 (2006.01)
C12N 9/02 (2006.01)
C12N 9/12 (2006.01)
C12N 9/90 (2006.01)
C12P 7/62 (2006.01)
C07K 14/245 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.07.2011 E 14185969 (4)**

97 Fecha y número de publicación de la concesión europea: **22.08.2018 EP 2845897**

54 Título: **Microorganismos con amplia gama de utilización de sustratos**

30 Prioridad:

16.07.2010 US 364898 P
16.07.2010 EP 10169808

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
14.01.2019

73 Titular/es:

NESTE OYJ (100.0%)
Keilaranta 21
02150 Espoo, FI

72 Inventor/es:

SICHWART, SHANNA;
HASCHENHERMES, BIRGIT;
BRÖKER, DANIEL;
STEINBÜCHEL, ALEXANDER;
HETZLER, STEFAN y
KOSKINEN, PERTTU

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 696 235 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Microorganismos con amplia gama de utilización de sustratos

5 Campo de la invención

La presente invención se refiere a nuevos microorganismos que tienen capacidad mejorada para utilizar azúcares monoméricos encontrados en recursos renovables, en particular lignocelulosa. Más específicamente, la invención se refiere a hospedadores *Ralstonia* y *Cupriavidus* genéticamente modificados y a métodos para utilizarlos.

10 Antecedentes

La lignocelulosa es el principal recurso renovable con varias aplicaciones, que incluye la materia prima para productos químicos a granel. Los principales componentes de la lignocelulosa son celulosa, hemicelulosa y lignina. La fracción de celulosa comprende polímeros del azúcar hexosa, glucosa. La fracción de hemicelulosa está compuesta principalmente por polímeros de los azúcares pentosa xilosa y arabinosa, y los azúcares hexosa manosa, glucosa y galactosa. Estos son abundantes en la biomasa lignocelulósica de maderas blandas, mientras que la xilosa se produce en maderas blandas y duras.

Existen numerosos ejemplos de esfuerzos para utilizar lignocelulosas más eficazmente por microorganismos y para extender la gama de utilización de sustratos de los microorganismos hacia fuentes de carbono abundantes y de bajo costo. Por ejemplo, se han hecho intentos para ampliar la utilización de pentosas no fermentables como la xilosa y la arabinosa que se producen en la biomasa lignocelulósica en muchos microorganismos. Muchas de estas investigaciones apuntaron a la extensión del metabolismo de los carbohidratos en los productores de etanol, como *Zymomonas mobilis* y *Saccharomyces cerevisiae* para bioprocesos de fermentación de etanol. El documento WO 2009/081941, por ejemplo, describe la modificación genética de la bacteria productora de etanol que pertenece al género *Zymomonas* para utilizar glucosa, manosa y xilosa.

Las bacterias que pertenecen al género *Ralstonia* y *Cupriavidus* son quimiolitotrofas que son capaces de acumular grandes cantidades de polihidroxialcanoatos (PHA) de recursos renovables o del dióxido de carbono. Se ha demostrado que las cepas del género *Ralstonia* y *Cupriavidus*, en particular, *R. eutropha* (nuevo nombre *Cupriavidus necator*) poseen un gran potencial biotecnológico, ya que estas cepas se pueden cultivar a altas densidades celulares a gran escala (Schlegel et al. 1961) y es adecuado para bioprocesos industriales a gran escala. Lamentablemente, *Cupriavidus* y *Ralstonia* tienen una gama de sustrato de carbohidratos muy estrecho limitado a fructosa, N-acetilglucosamina y gluconato. Los intentos de ampliar la gama de utilización de sustratos de H16 de *R. eutropha* se han descrito en la técnica anterior (Pries et al. 1990, Schlegel y Gottschalk, 1965). Se ha informado que se ha aislado un mutante espontáneo que utiliza glucosa de H16 de *R. eutropha* que supuestamente transportaba este azúcar a la célula mediante un transportador pasivo (Schlegel y Gottschalk 1965, König, et al. 1969). Además, Pries et al. 1990 han informado sobre la introducción del gen de la β -galactosidasa y el operón *gal* de *E. coli* en la cepa H16 de *R. eutropha* y la utilización de lactosa y galactosa mediante esta bacteria.

Además, Buchholz et al. 1994 han descrito la construcción de una biblioteca genómica de *Pseudomonas saccharophila* y su transferencia a *Alcaligenes eutrophus* (*Ralstonia eutropha*). Un inserto que contiene los genes *xyl* que codifican xilosa isomerasa y xiluloquinasa hizo que el hospedador *A. eutrophus* (pGN3) creciera en la xilosa.

A pesar de los intentos anteriores de extender la gama de utilización de sustratos de ciertos microorganismos hacia fuentes de carbono abundantes y de bajo costo, todavía existe una necesidad de nuevos métodos y microorganismos para una utilización más eficaz de los recursos renovables, en particular las lignocelulosas.

50 Sumario

Un objetivo de la presente invención es proporcionar un proceso para utilizar eficazmente lignocelulosas, en particular azúcares monoméricos encontrados en lignocelulosas.

Otro objetivo de la presente invención es proporcionar nuevos microorganismos capaces de utilizar lignocelulosas, en particular azúcares monoméricos encontrados en la lignocelulosa.

Para lograr estos objetivos, la invención se caracteriza por las características que se incluyen en las reivindicaciones independientes. Otras reivindicaciones representan las realizaciones preferidas de la invención.

La invención descrita en el presente documento proporciona, en un aspecto, un nuevo microorganismo, que se modifica genéticamente para expresar genes que amplían su capacidad de utilizar diferentes azúcares monoméricos que están presentes en la lignocelulosa.

Un primer aspecto de la presente invención proporciona un microorganismo del género *Cupriavidus* o *Ralstonia* modificado genéticamente para expresar:

xilosa isomerasa (EC 5.3.1.5), xiluloquinasa (EC 2.7.1.17) y Xyl G (EC 3.6.3.17), Xyl F y Xyl H, en el que dicho microorganismo es capaz de crecer en arabinosa y, opcionalmente, también en glucosa y/o galactosa como fuente de carbono.

- 5 En una realización, el microorganismo se modifica genéticamente para expresar fosfomanosa isomerasa (EC 5.3.1.8) y la proteína de difusión facilitada para la absorción de manosa (EC 1.3.1.74) y, opcionalmente, manofructoquinasa (EC 2.7.1.4). Normalmente, los microorganismos del género *Cupriavidus* o *Ralstonia* son incapaces de utilizar la manosa sin dicha modificación, mientras que los microorganismos modificados son capaces de utilizar la manosa. Los genes utilizados para dicha modificación genética pueden ser exógenos o endógenos, normalmente son exógenos.
- 10 En diversas realizaciones, cuando el microorganismo se modifica genéticamente para expresar un gen que codifica una proteína de difusión facilitada para la absorción de manosa (EC 1.3.1.74), en particular el gen *glf*, el microorganismo hospedador es sorprendentemente capaz de utilizar también glucosa.

- 15 El microorganismo de la presente invención se modifica genéticamente para expresar xilosa isomerasa (EC 5.3.1.5), xiluloquinasa (EC 2.7.1.17) y un transportador ABC de alta afinidad Xyl G (EC 3.6.3.17), Xyl F y Xyl H. Normalmente, los microorganismos del género *Cupriavidus* o *Ralstonia* son incapaces de utilizar la xilosa sin dicha modificación, mientras que los microorganismos modificados son capaces de utilizar la xilosa. Los genes utilizados para dicha modificación genética pueden ser exógenos o endógenos, normalmente son exógenos. Normalmente, dichos microorganismos son incapaces de utilizar arabinosa sin dichos genes, mientras que en diversas realizaciones, los
- 20 microorganismos modificados son sorprendentemente capaces de utilizar arabinosa como fuente de carbono. Además, cuando los microorganismos se modifican genéticamente para expresar el transportador ABC de alta afinidad (en particular los genes Xyl G (EC 3.6.3.17), Xyl F y Xyl H), los microorganismos hospedadores son sorprendentemente capaces de utilizar glucosa y galactosa.

- 25 En una realización adicional de la invención, los microorganismos hospedadores están modificados genéticamente para expresar glucoquinasa (EC 2.7.1.2). Esta modificación hace que el microorganismo modificado genéticamente también sea capaz de utilizar glucosa.

- 30 Por lo tanto, los microorganismos son capaces de utilizar glucosa como fuente de carbono además de manosa, o además de manosa, xilosa y arabinosa.

En una realización preferida de la invención, los microorganismos son capaces de utilizar manosa, glucosa, galactosa, xilosa y arabinosa como fuente de carbono.

- 35 En una realización adicional, los microorganismos se modifican genéticamente para expresar transaldolasa (EC 2.2.1.2) y/o transcetolasa (EC 2.2.1.1). Los genes que codifican transaldolasa y/o transcetolasa mejoran la resistencia del crecimiento del microorganismo modificado, en particular en pentosas, tales como xilosa y arabinosa.

- 40 El microorganismo puede ser cualquier especie o cepa perteneciente al género *Cupriavidus* o *Ralstonia*. Preferentemente, el microorganismo pertenece a la especie *C. necator* (anteriormente *R. eutropha* o *Alkaligenes eutrophus*).

- 45 En otro aspecto, la invención proporciona un vector que comprende un gen que codifica xilosa isomerasa (EC 5.3.1.5), xiluloquinasa (EC 2.7.1.17), un transportador ABC de alta afinidad Xyl G, (EC 3.6.3.17), Xyl F y Xyl y elementos reguladores para regular la expresión de dichos genes en *Cupriavidus* o *Ralstonia*. Dichos genes pueden estar en el mismo vector o en vectores separados.

- 50 En una realización, el vector puede comprender un gen que codifica transaldolasa (EC 2.2.1.2) y/o transcetolasa (EC 2.2.1.1). Dichos genes pueden estar en el mismo vector o en vectores separados.

- En una realización adicional, el vector puede comprender un gen que codifica glucoquinasa (EC 2.7.1.2). Dichos genes pueden estar en el mismo vector o en vectores separados como los genes enumerados anteriormente.

- 55 En una realización adicional más, un vector comprende uno cualquiera o cualquier combinación de los genes enumerados anteriormente.

Se puede utilizar en el vector cualquier promotor que pueda expresarse en cepas de los géneros *Ralstonia* o *Cupriavidus*.

- 60 Los promotores adecuados comprenden, por ejemplo, el promotor *lac*, el promotor de *neokanamicina*, el promotor de gliceraldehído-3-fosfato deshidrogenasa, el P2, el promotor hidrogenasa (SH) y el P_L.

- En una realización preferida, el vector es un plásmido de amplia gama de hospedadores que se puede utilizar para extender la gama de utilización de sustratos de varias cepas de hospedadores diferentes de *Cupriavidus* o *Ralstonia*.

- 65 Se divulga un método para modificar los microorganismos de los géneros *Ralstonia* o *Cupriavidus*. El método

comprende que los hospedadores se modifican genéticamente para expresar una o más secuencias de nucleótidos seleccionadas del grupo que comprende

una secuencia o secuencias de nucleótidos que codifican fosfomanosa isomerasa (EC 5.3.1.8) y la proteína de difusión facilitada para la absorción de manosa (EC 1.3.1.74) y/o cotransportador unidireccional E de protones de xilosa y,

opcionalmente, manofructoquinasa (EC 2.7.1.4), en el que dicho microorganismo es capaz de crecer en manosa y, opcionalmente, también en glucosa como fuente de carbono, y/o

una secuencia o secuencias de nucleótidos que codifican xilosa isomerasa (EC 5.3.1.5), xiluloquinasa (EC 2.7.1.17) y un transportador ABC de alta afinidad y/o cotransportador unidireccional E de protones de xilosa, en el que dicho microorganismo es capaz de crecer en xilosa o xilosa y arabinosa y, opcionalmente, también en glucosa y galactosa

como fuente de carbono.

El método puede comprender además que el microorganismo se modifique genéticamente para expresar transaldolasa (EC 2.2.1.2) y/o transcetolasa (EC 2.2.1.1).

El método puede comprender además que el microorganismo se modifique genéticamente para expresar glucoquinasa (EC 2.7.1.2).

En un aspecto adicional más, la invención proporciona un método para cultivar el microorganismo de acuerdo con la invención, o un microorganismo de *Cupriavidus* o *Ralstonia* que comprende el vector de acuerdo con la invención o el sistema de vectores de acuerdo con la invención en un medio de cultivo que comprende manosa, glucosa, galactosa, xilosa o arabinosa, cualquiera de ellos, cualquier combinación de ellos o todos ellos, como fuente de carbono.

En una realización, el medio de cultivo puede comprender manosa como fuente de carbono y, opcionalmente, glucosa. En otra realización, el medio de cultivo puede comprender o bien xilosa y arabinosa como fuente de carbono y, opcionalmente, también glucosa y/o galactosa.

En una realización adicional, el medio de cultivo puede comprender manosa y una o ambas xilosa y arabinosa como fuente de carbono y, opcionalmente, también glucosa y/o galactosa.

En un aspecto adicional, el método comprende que el microorganismo pueda sintetizar productos químicos a granel seleccionados de lípidos, polihidroxialcanoatos, ácidos grasos y derivados de ácidos grasos, y el producto químico se recupera de las células o del medio de cultivo.

Se obtienen ventajas considerables por medio de la presente invención. Por medio de la invención, se obtienen nuevos microorganismos que pertenecen al género *Cupriavidus* o *Ralstonia*, teniendo dichos microorganismos una amplia gama de utilización de sustrato. Los microorganismos modificados como se describen en el presente documento pueden utilizar eficazmente carbohidratos de bajo costo, en particular lignocelulosa que comprende azúcares monoméricos. La xilosa, la manosa y la arabinosa son azúcares de madera presentes en forma polimerizada en las hemicelulosas de las paredes celulares vegetales. Los microorganismos de la presente invención también son capaces de utilizar glucosa monomérica, que se encuentra en particular en sacarosa y almidón y también en fracciones de celulosa y hemicelulosa en materiales lignocelulósicos. Los microorganismos modificados de la presente invención son adecuados para producir muchos productos químicos a granel.

Descripción de las figuras

Figura 1. Rutas de utilización de manosa y glucosa establecidas en la cepa H16 de *R. eutropha*. GLF, transportador de difusión facilitada de glucosa; KDPG, ruta de Entner-Doudoroff; MAK, manofructoquinasa; PMI, fosfomanosa isomerasa, GLK, glucoquinasa; PGI, fosfoglucosa isomerasa.

Figura 2. Mapas físicos de los plásmidos construidos pBBR1MCS-3:: *glf*, pBBR1MCS-3:: *glf::mak*, pBBR1MCS-3:: *glf::pmi*, pBBR1MCS-3:: *glf::mak::pmi*, pBBR1MCS-2::P_{nk}:: *glf* y pBBR1MCS-2::P_{nk}:: *glk::glf*. Los genes *glf*, *glk*, *mak* y *pmi*, respectivamente, se cortaron de plásmidos híbridos de pJET1.2 o pCR2.1 después de la subclonación utilizando los sitios de restricción indicados y, luego, se ligaron en varias matrices a los vectores de expresión linealizados pBBR1MCS-2 y pBBR1MCS-3. Se indican los sitios de escisión relevantes y los genes estructurales (Km^R, casete de resistencia a kanamicina, Tc^R, casete de resistencia a tetraciclina; *mob*, sitio de movilización; *rep*, origen de replicación; *glf*, gen que codifica el transportador de difusión facilitada de glucosa (GLF) de *Z. mobilis*; *glk*, gen que codifica la glucoquinasa (GLK) de *E. coli*, *mak*, gen que codifica una manofructoquinasa (MAK) de *E. coli*, *pmi*, gen que codifica una fosfomanosa isomerasa (PMI) de *E. coli*).

Figura 3. Análisis de transcripción de *glf*, *mak* y *pmi* en la cepa H16 de *R. eutropha* que alberga pBBR1MCS-3:: *glf::mak::pmi* y en el control negativo que alberga solo el vector pBBR1MCS-3. La expresión de *glf*, *mak* y *pmi* se analizó mediante RT-PCR en muestras que contenían ARN total aislado de células de la fase de crecimiento exponencial de la cepa H16 de *R. eutropha* que alberga pBBR1MCS-3:: *glf::mak::pmi* y de la cepa H16 de *R. eutropha* que alberga el vector pBBR1MCS-3. Las células se cultivaron en MSM que contenía gluconato de sodio al 1 % (p/v) como única fuente de carbono. Los productos de la PCR resultantes se separaron mediante electroforesis en gel de agarosa y se tiñeron con bromuro de etidio. Los carriles 1, 3 y 5 representan el ensayo de RT-PCR para *glf*, *mak* y *pmi*, respectivamente, para la cepa H16 de *R. eutropha* que alberga pBBR1MCS-3:

glf::mak::pmi, mientras que los carriles 2, 4 y 6 representan el controles negativos para detectar si hay cualquier expresión de *glf*, *mak* y *pmi* en la cepa que alberga solo el vector pBBR1MCS-3. Un marcador de ADN *λ*PstI (carriles M, Fermentas, Alemania) sirvió para la comparación de tamaño.

Figura 4. Extractos crudos y proteínas de membrana solubilizadas de la cepa recombinante H16 de *R. eutropha* que alberga pBBR1MCS-2 o pBBR1MCS-2::P_{nk}::*glk*::*glf*. Se cultivaron las células durante 48 h en medio MSM que contenía glucosa sódica al 1 % (p/v) como fuente de carbono. Se separaron las proteínas en geles de SDS-poliacrilamida al 12,5 % (p/v) y se tiñeron con azul brillante de Coomassie. A, extractos crudos. Carril 1: cepa H16 de *R. eutropha* pBBR1MCS-2, carril 2: cepa H16 de *R. eutropha* pBBR1MCS-2::P_{nk}::*glk*::*glf*; B, proteínas de membrana solubilizadas. Carril 3: cepa H16 de *R. eutropha* pBBR1MCS-2, carril 4: cepa H16 de *R. eutropha* pBBR1MCS-2::P_{nk}::*glk*::*glf*. Est, peso molecular estándar. Las flechas indican transportador de glucosa (Glf) y glucoquinasa (GK) de proteínas recombinantes detectadas.

Figura 5. Crecimiento de la cepa recombinante H16 de *R. eutropha*. Cepas recombinantes de *R. eutropha* que albergan diferentes plásmidos. Las células se cultivaron en MSM líquido que contenía glucosa al 1 % (p/v) como única fuente de carbono. □ pBBR1MCS-3:::*glf*, *pBBR1MCS-2::P_{nk}::*glf*, ΔpBBR1MCS-2::P_{nk}::*glk*::*glf*, ▲ pBBR1MCS-3:::*glf*::*mak*::*pmf*, ♦ cepa H16 de *R. eutropha* G⁺1 control. Se realizaron los cultivos en matraces Erlenmeyer incubados en un agitador rotatorio horizontal a 110 rpm a 30 °C. La densidad óptica se midió a 600 nm en un fotómetro Klett.

Figura 6. Cepas recombinantes de *R. eutropha* que albergan el plásmido pBBR1MCS-3:::*glf*::*mak*::*pmi*. Se cultivaron células en MSM líquido que contenía tetraciclina (12,5 µg/ml) y Δ gluconato de sodio al 0,5 % (p/v), ▲ manosa al 0,5 % (p/v), □ manosa al 1 % (p/v), ♦ manosa al 2 % (p/v), o manosa al 4 % (p/v) como única fuente de carbono. Se realizaron los cultivos en matraces Erlenmeyer incubados en un agitador rotatorio horizontal a 130 rpm a 30 °C. La densidad óptica se midió a 600 nm en un fotómetro Klett.

Figura 7. La ruta de utilización de la D-xilosa en bacterias (XI, xilosa isomerasa; XK, xiluloquinasa; PPP, ruta pentosa fosfato)

Figura 8. Mapas físicos de los plásmidos construidos pBBR1MCS-3:: *xyIA*B, pBBR1MCS-3::*xyIA*BE y pBBR1MCS-3::*xyIA*BFGH. Los genes *xyIA*, *xyIB*, *xyIE* y *xyIFGH*, respectivamente, se cortaron de plásmidos híbridos de pJET1.2 después de la subclonación utilizando los sitios de restricción indicados y, luego, se ligaron al vector de expresión linealizado pBBR1MCS-3. Se indican los sitios de escisión relevantes y los genes estructurales (Tc^R, casete de resistencia a tetraciclina; *mob*, sitio de movilización; *rep*, origen de replicación; *xyIA*, gen que codifica la xilosa-isomerasa (XylA); *xyIB*, gen que codifica la xiluloquinasa (XylB); *xyIE*, gen que codifica el transportador de baja afinidad (XylE); y *xyIFGH*, genes que codifican el transportador ABC (XylFGH) de *E. coli*, respectivamente).

Figura 9. Crecimiento de la cepa recombinante H16 de *R. eutropha* que alberga el plásmido pBBR1MCS-3::*xyIA*BFGH, y la cepa H16 de *R. eutropha*, que se cultivaron en un fluido MSM que contenía diferentes concentraciones (0,5 %, 1 %, 2 % [p/v]) de D-xilosa, glucosa al 1 % (p/v) y con gluconato de Na al 1 % (p/v) como única fuente de carbono, respectivamente. □ pBBR1MCS-3::*xyIA*BFGH, D-xilosa al 0,5 % (p/v); ■ pBBR1MCS-3::*xyIA*BFGH, D-xilosa al 1,0 % (p/v); Δ pBBR1MCS-3w*xyIA*BFGH, D-xilosa al 2,0 % (p/v); ▲ pBBR1MCS-3::*xyIA*BFGH, D-xilosa al 4,0 % (p/v); ♦ pBBR1MCS-3::*xyIA*BFGH, Na-gluconato al 1,0 % (p/v); ♦ pBBR1MCS-3::*xyIA*BFGH, glucosa al 1,0 % (p/v); + control H16 de *R. eutropha*, gluconato de Na al 1,0 % (p/v). Se realizaron los cultivos en matraces Erlenmeyer incubados en un agitador rotatorio horizontal a 110 rpm a 30 °C. La densidad óptica se midió a 600 nm en un fotómetro Klett.

Descripción detallada de la invención

"Un microorganismo modificado genéticamente" se refiere aquí a un microorganismo cuyo material genético ha sido alterado utilizando técnicas de ingeniería genética. Estas técnicas implican la utilización de moléculas de ácido nucleico de diferentes fuentes, por ejemplo, de otro microorganismo o pueden ser al menos parcialmente sintéticas. Las moléculas de ácido nucleico de diferentes fuentes pueden combinarse en una o más moléculas. Estas construcciones pueden transferirse a un microorganismo, dándole construcciones genéticas modificadas o novedosas.

"Un gen" se refiere aquí normalmente a una secuencia de nucleótidos que codifica un producto específico, aquí, generalmente, una enzima. La región codificante está normalmente unida operativamente a regiones reguladoras, en particular a un promotor, que ser capaz de funcionar en la célula hospedadora. Aquí se introduce, normalmente, en una célula hospedadora una construcción genética, que consiste esencialmente en una región codificante de un gen y una región reguladora, en particular, un promotor, funcional en la célula hospedadora. Puede utilizarse un vector adecuado para introducir la construcción genética deseada en una célula hospedadora. "Un gen" puede referirse a uno o más genes.

La modificación genética puede significar que uno o más genes exógenos se introducen y se expresan en un microorganismo hospedador, o que uno o más de los genes expresados son endógenos y uno o más son exógenos, o que la expresión de uno o más genes endógenos aumenta en un hospedador en comparación con el hospedador

principal.

Los microorganismos hospedadores en diversas realizaciones de la invención pertenecen al género *Cupriavidus* o *Ralstonia*.

5 Un microorganismo se refiere aquí, normalmente, a un microorganismo aislado, es decir, un microorganismo se utiliza como un cultivo puro.

10 En una realización, la especie puede ser cualquier especie de *Ralstonia* que incluye, pero no se limita a *Ralstonia insidiosa*, *R. mannitolytica*, *R. pickettii*, *R. solanacearum* y *R. syzygii*. En otra realización, la especie puede ser cualquier especie de *Cupriavidus* que incluye, pero no se limita a *Cupriavidus basilensis*, *C. campinensis*, *C. gilardii*, *C. laharis*, *C. metallidurans*, *C. necator*, *C. oxalaticus*, *C. pauculus*, *C. pinatubonensis*, *C. respiraculi* y *C. taiwanensis* (*Nomenclatura bacteriana actualizada (listas aprobadas, lista de validación)*, Mayo de 2010, recopilada por DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH). Las especies preferidas son *R. eutropha* y *C. metallidurans*. La especie más preferida es *R. eutropha* (nombre actual *C. necator*, nombres anteriores *Alcaligenes eutrophus* e *Hydrogenomonas eutropha*). Las cepas de *Cupriavidus* y *Ralstonia* están disponibles para el público a partir de colecciones de cultivos microbianos, como ATCC, DSMZ y NCIB.

20 La invención se ha ejemplificado mediante la utilización de la cepa H 16 de *R. eutropha* (nombre actual *C. necator*). Esta cepa es una cepa de tipo silvestre, que está disponible en Deutsche Sammlung von Mikroorganismen und Zellkulturen, DSMZ, con el número de referencia DSM428, en la Colección Americana de Cultivos Tipo con el número de referencia ATCC17699, y en la Colección Nacional de Bacterias Industriales y Marinas, NCIBM, con el número de referencia NCIB10442. Todas estas cepas están disponibles para el público.

25 Las cepas H16 de *R. eutropha* depositadas, que se pueden utilizar en la invención son

H16 de *Ralstonia eutropha*, N.º de referencia de DSMZ: DSM428, DSM 15443 (mutante de DSM 428)
ID taxonomía: 381666; Nombre heredado del BLAST: b-proteobacteria
Otros nombres:

30 sinónimos: H16 de *Cupriavidus necator*, H16 de *Alcaligenes eutropha*, H16 de *Wautersia eutropha*, ATCC 17699 de *Cupriavidus necator*, ATCC 17699 de *Ralstonia eutropha*;

35 nombres equivalentes: cepa H16 de *Ralstonia eutropha*, cp. H16 de *Ralstonia eutropha*.

Otras cepas de *R. eutropha* que se pueden utilizar en la invención son:

H850 de *Ralstonia eutropha*

40 N.º de referencia de DSMZ: DSM 13439
ID taxonomía: 53482; Nombre heredado del BLAST: b-proteobacteria

Otros nombres:

45 sinónimos: H850 de *Wautersia eutropha*, H850 de *Cupriavidus necator*, H850 de *Alcaligenes eutropha*

JMP134 de *Ralstonia eutropha*

50 N.º de referencia de DSMZ: DSM 4058
ID taxonomía: 264198; Nombre heredado del BLAST: b-proteobacteria

Otros nombres:

55 sinónimos: JMP134 de *Cupriavidus pinatubonensis*, JMP134 de *Cupriavidus necator*, JMP134 de *Wautersia eutropha*, JMP134 de sp. *Ralstonia*;

nombres equivalentes: cp. JMP134 de *Ralstonia eutropha* y cepa JMP134 de *Ralstonia eutropha*.

Además, también las siguientes otras cepas de *Ralstonia eutropha* se pueden utilizar en la invención:

60 335 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 531, ATCC 17697)
N9A de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 518;
B19 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 515):
G27 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 516):
G29 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 517):
65 336 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 529, ATCC 17698)
338 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 530, ATCC 17700);

TA06 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 4182);
H850 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 5536; NRRL B-15940)
KTO₂ de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 6519);
Tfa17 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 11098)

5 N-1 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 13513; ATCC 43291)
NRRL B-2804 de *Ralstonia eutropha* (N.º de referencia de DSMZ: DSM 30029, ATCC 25207, NRRL B-2804);

10 "Biomasa lignocelulósica" se refiere aquí a la biomasa vegetal que comprende principalmente celulosa, hemicelulosas y lignina. La biomasa lignocelulósica es, por ejemplo, residuos agrícolas (por ejemplo, paja de trigo, cebada o arroz, cáscara (paja), rastrojos de maíz, bagazo de caña de azúcar), cultivos energéticos, madera o materiales o residuos vegetales y desechos municipales, por ejemplo, desechos de papel.

15 "Producto químico a granel" se refiere a cualquier producto químico útil, que puede producirse a partir de materias primas económicas en grandes cantidades. Ejemplos de productos químicos a granel útiles son lípidos, por ejemplo, polihidroxialcanoatos, en particular, polihidroxibutirato (PHB), u otros lípidos, tales como ácidos grasos y/o derivados de ácidos grasos, por ejemplo, acilglicerol.

20 El término "lípidos" se refiere a una sustancia grasa, cuya molécula generalmente contiene, como una parte, una cadena de hidrocarburo alifática, que se disuelve en disolventes orgánicos no polares pero es poco soluble en agua. Los lípidos son un grupo esencial de moléculas grandes en células vivas. Los lípidos son, por ejemplo, grasas, aceites, ceras, ésteres de cera, esteroides, terpenoides, isoprenoides, carotenoides, polihidroxialcanoatos, ácidos nucleicos, ácidos grasos, alcoholes grasos, ésteres de ácidos grasos, fosfolípidos, glucolípidos, esfingolípidos y acilglicerol. Los términos lípidos y aceites se utilizan en esta descripción de forma sinónima.

25 El término "acilglicerol" se refiere a un éster de glicerol y ácidos grasos. Los acilglicerol se producen naturalmente como grasas y aceites grasos. Los ejemplos de acilglicerol incluyen triacilglicerol (TAG, triglicéridos), diacilglicerol (diglicéridos) y monoacilglicerol (monoglicéridos).

30 Las "materias primas" o "materiales de partida" son materiales que se forman o se incluyen en el medio de cultivo de un hospedador *Cupriavidus* o *Ralstonia* y que el microorganismo puede utilizar para su crecimiento y/o producción de los productos químicos deseados. Ventajosamente, los materiales son económicos y están disponibles en grandes cantidades. Los materiales de partida adecuados son, por ejemplo, biomasa lignocelulósica (por ejemplo, madera o materiales vegetales, sus fracciones o residuos) y desecho(s) o residuo(s) agrícolas, micro o macroalgas y cualquier material, que comprenda carbohidratos, tal como glucosa, galactosa, manosa, xilosa o arabinosa o combinaciones de los mismos. Algunos materiales de partida pueden comprender principalmente xilosa (hemicelulosas), algunos pueden comprender principalmente glucosa (celulosa o almidón) y algunos pueden comprender principalmente manosa (hemicelulosas basadas en galactoglucomanano).

40 Los azúcares, manosa, glucosa, xilosa, arabinosa y galactosa deben proporcionarse a los microorganismos como azúcares monoméricos. Las materias primas que contienen estos azúcares en forma polimérica, por ejemplo, lignocelulosas, se pueden tratar con cualquier método conocido para hidrolizar los polímeros de azúcar en azúcares monoméricos que sean utilizables por los microorganismos de acuerdo con la presente invención.

45 En el cultivo de *Cupriavidus* o *Ralstonia* se puede utilizar cualquier medio de cultivo que comprenda componentes adecuados para el crecimiento de *Cupriavidus* o *Ralstonia*.

50 En aplicaciones industriales, el medio de cultivo comprende normalmente azúcares monoméricos procedentes de materiales lignocelulósicos como fuente de carbono, y además, habitualmente, nutrientes y sales o componentes de los mismos como micro y macronutrientes, y agua. Un medio de cultivo puede comprender manosa, glucosa, galactosa, xilosa o arabinosa o cualquier combinación de las mismas como fuente de carbono, normalmente como la principal fuente de carbono, en algunas realizaciones como la única fuente de carbono.

55 En el cultivo de *Cupriavidus* o *Ralstonia* se pueden utilizar condiciones de cultivo adecuadas para el cultivo de *Cupriavidus* o *Ralstonia*. Tales condiciones de cultivo son bien conocidas por un experto en la técnica. En el cultivo se pueden utilizar medios líquidos o sólidos. Los cultivos líquidos pueden agitarse, agitación adecuada de 50 a 1000 rpm, normalmente de 100 a 600 rpm. La temperatura de cultivo puede ser de 15 a 45 °C, preferentemente de 25 a 37 °C, normalmente 30 °C. El tiempo de incubación puede ser de 1 día a 14 días, normalmente de 2 a 7 días, preferentemente de 24 a 72 horas. El cultivo puede ser un cultivo discontinuo, semicontinuo o continuo.

60 El término "gen endógeno" se refiere aquí a un gen que es natural para el hospedador *Cupriavidus* o *Ralstonia*.

El término "gen exógeno" se refiere aquí a un gen que no es natural para el hospedador *Cupriavidus* o *Ralstonia*.

65 Los genes introducidos en el hospedador *Ralstonia* o *Cupriavidus* pueden ser endógenos o exógenos, normalmente son exógenos, si/mientras el hospedador no posea los genes de forma natural. En algunas realizaciones, los genes endógenos producidos de forma natural por el hospedador pueden sobreexpresarse en el hospedador, por ejemplo,

mediante la expresión de la región codificante en un potente promotor.

Un gen que codifica "fosfomanosa isomerasa" puede ser cualquier gen que codifique una enzima que tenga actividad fosfomanosa isomerasa. De acuerdo con la Clasificación de Enzimas, la enzima tiene el número EC 5.3.1.8. En una
5 realización específica de la invención, el gen puede ser un gen que codifica fosfomanosa isomerasa (PMI, de sus siglas en inglés) o manosa-6-fosfato isomerasa de la cepa K12 de *Escherichia coli* (EC 5.3.1.8) (N.º DE REFERENCIA P00946) (nombre del gen: *pmi*, origen: cepa K12 de *Escherichia coli*) (SEQ ID NO:1).

Un gen que codifica una "proteína de difusión facilitada para la absorción de manosa" puede ser cualquier gen que
10 codifique la proteína de difusión facilitada para la actividad de absorción de manosa (EC 1.3.1.74) o un gen que codifica el cotransportador unidireccional E de protones de xilosa. En una realización específica de la invención, el gen puede ser un gen que codifica el transportador de difusión facilitada de glucosa (GLF) de *Zymomonas mobilis* (EC 1.3.1.74) (N.º DE REFERENCIA P21906) (nombre del gen: *glf*, origen: *Zymomonas mobilis*) (SEQ ID NO:2). Sorprendentemente, un hospedador *Ralstonia* o *Cupriavidus* introducido para expresar un gen que codifica una
15 proteína de difusión facilitada para la actividad de absorción de manosa, en particular *glf*, puede utilizar glucosa como fuente de carbono.

Un gen que codifica "manofructoquinasa" puede ser cualquier gen que codifique la actividad de la manofructoquinasa (EC 2.7.1.4). En una realización específica de la invención, el gen puede ser un gen que codifica la manofructoquinasa (MAK) (EC 2.7.1.4) de *Escherichia coli*, en particular de la cepa K12 (N.º DE REFERENCIA P23917) (nombre del gen: *mak*, origen: cepa K12 de *Escherichia coli*) (SEQ ID NO:3).

Un gen que codifica "xilosa isomerasa" puede ser cualquier gen, que codifique la actividad de la xilosa isomerasa (EC 5.3.1.5). En una realización específica de la invención, el gen puede ser un gen que codifica la D-xilosa isomerasa (XylA) de la cepa K12 de *Escherichia coli* (N.º DE REFERENCIA P00944) (EC 5.3.1.5) (nombre del gen: *xylA*, origen: cepa K12 de *Escherichia coli*) (SEQ ID NO:4).

Un gen que codifica "xiluloquinasa" puede ser cualquier gen, que codifica la actividad xiluloquinasa (EC 2.7.1.17). En una realización específica de la invención, el gen puede ser un gen que codifica la xiluloquinasa (XylB) de la cepa K12 de *Escherichia coli* (N.º DE REFERENCIA P09099) (EC 2.7.1.17) (nombre del gen: *xylB*, origen: cepa K12 de *Escherichia coli*) (SEQ ID NO:5).

Un gen que codifica "cotransportador unidireccional E de protones de xilosa" puede ser cualquier gen, que codifica el cotransportador unidireccional E de protones de xilosa. El gen puede ser un gen que codifica el cotransportador unidireccional de protones de xilosa (XylE) de la cepa K12 de *Escherichia coli* (nombre del gen: *xylE*, origen: cepa K12 de *Escherichia coli* (N.º DE REFERENCIA P0AGF4) (SEQ ID NO:6).

En una realización específica de la invención, el gen que codifica "Xyl G, Xyl F y Xyl H" puede ser un gen que codifica la proteína periplásmica de unión a D-xilosa (XylF) de la cepa K12 de *Escherichia coli* (nombre del gen: *xylF*, origen: cepa K12 de *Escherichia coli*) (N.º de REFERENCIA P37387) (SEQ ID NO: 7), un gen que codifica la proteína de unión a ATP (XylG) de importación de xilosa de la cepa K12 de *Escherichia coli* (nombre del gen: *xylG*, origen: cepa K12 de *Escherichia coli*) (N.º de REFERENCIA P37388) (EC 3.6.3.17) (SEQ ID NO: 8), y un gen que codifica la proteína permeasa del sistema de transporte de xilosa (XylH) de la cepa K12 de *Escherichia coli* (nombre del gen: *xylH*, origen: cepa K12 de *Escherichia coli* (N.º DE REFERENCIA P0AGF4) (SEQ ID NO:9). Sorprendentemente, un hospedador *Ralstonia* o *Cupriavidus* introducido para expresar genes, que codifican el transportador ABC de alta afinidad, en particular los genes que codifican Xyl F, Xyl G y Xyl H, pueden utilizar glucosa y galactosa como fuente de carbono.

Un gen que codifica "transaldolasa" puede ser cualquier gen, que codifica la transaldolasa (EC 2.2.1.2). En una realización específica de la invención, el gen puede ser un gen que codifica transaldolasa B (TalB) de la cepa K12 de *Escherichia coli* (nombre del gen: *talB*, origen: cepa K12 de *Escherichia coli* (N.º DE REFERENCIA BAB96586) (EC 2.2.1.2) (SEQ ID NO: 10).

Un gen que codifica "cetolasa" puede ser cualquier gen, que codifica la transcetolasa (EC 2.2.1.1). En una realización específica de la invención, el gen puede ser un gen que codifica transcetolasa 1 (TktA) de la cepa K12 de *Escherichia coli* (nombre del gen: *tktA*, origen: cepa K12 de *Escherichia coli* (N.º DE REFERENCIA P27302) (EC 2.2.1.1) (SEQ ID NO: 11).

Un gen que codifica "glucoquinasa " puede ser cualquier gen, que codifica la glucoquinasa (EC 2.7.1.2). En una realización específica de la invención, el gen puede ser un gen que codifica glucoquinasa (*glk*) de la cepa K12 de *Escherichia coli* (nombre del gen: *glk*, origen: cepa K12 de *Escherichia coli*) (N.º DE REFERENCIA P0A6V8.1) (EC 2.7.1.2) GLK_ECOLI RecName: Completo = Glucoquinasa; AltName: Completo = Glucosa quinasa [K-12 de *Escherichia coli*] (SEQ ID NO: 12).

Un hospedador *Ralstonia* o *Cupriavidus* introducido para expresar glucoquinasa es capaz de utilizar glucosa. Se divulgan productos genéticos que se sabe que catalizan la reacción de interés. Se pueden utilizar genes, que codifican la misma función o equivalente y son altamente homólogos a los genes. Esto significa, por ejemplo, que las proteínas codificadas por los genes tienen un % de identidad alto cuando se comparan entre sí.

Por ejemplo, se pueden utilizar los siguientes genes:

Los genes respectivos de diferentes cepas de *Escherichia* (por ejemplo, *Escherichia albertii* [fosfomanosa isomerasa ZP_02903372.1, manosa-6-fosfato isomerasa, clase I [Escherichia albertii TW07627] (SEQ ID NO: 13);
 5 manofructoquinasa ZP_02902328.1, mano quinasa [Escherichia albertii TW07627] (SEQ ID NO: 14); *Escherichia fergusonii* [fosfomanosa isomerasa YP_002382576.1, manosa-6-fosfato isomerasa [Escherichia fergusonii ATCC 35469] (SEQ ID NO: 15); xilosa isomerasa YP_002384641.1]) xilosa isomerasa [Escherichia fergusonii ATCC 35469] (SEQ ID NO: 16); cepas de *Shigella* (por ejemplo, *Shigella flexneri* [fosfomanosa isomerasa NP_837299.1], manosa-6-fosfato isomerasa [Shigella flexneri 2a cp. 2457T] (SEQ ID NO: 17), *Shigella sonnei* [fosfomanosa isomerasa
 10 YP_310481.1, manosa-6-fosfato isomerasa [Shigella sonnei Ss046] (SEQ ID NO: 18); xilosa isomerasa YP_312597.1], xilosa isomerasa [Shigella sonnei Ss046] (SEQ ID NO: 19); *Shigella dysenteriae* [fosfomanosa isomerasa ZP_03064125.1, manosa-6-fosfato isomerasa, clase I [Shigella dysenteriae 1012] (SEQ ID NO: 20); manofructoquinasa ZP_03064822.1, mano(fructo)quinasa [Shigella dysenteriae 1012] (SEQ ID NO: 21), xilosa isomerasa ZP_03063610.1]), xilosa isomerasa [Shigella dysenteriae 1012] (SEQ ID NO: 22); cepas de *Citrobacter* (por ejemplo, *Citrobacter koseri* [fosfomanosa isomerasa P_001453192.1, proteína hipotética CKO_01624 [Citrobacter koseri ATCC BAA-895] (SEQ ID NO: 23); xilosa isomerasa YP_001456500.1], xilosa isomerasa [Citrobacter koseri ATCC BAA-895] (SEQ ID NO: 24); *Citrobacter rodentium* [fosfomanosa isomerasa YP_003365034, manosa-6-fosfato isomerasa [Citrobacter rodentium ICC168] (SEQ ID NO: 25); manofructoquinasa YP_003364074.1, probable mano(fructo)quinasa [Citrobacter rodentium ICC168] (SEQ ID NO: 26); xilosa isomerasa YP_003367655.1], D-xilosa isomerasa [Citrobacter rodentium ICC168] (SEQ ID NO: 27); *Citrobacter youngae* [fosfomanosa isomerasa ZP_06352970.2, manosa-6-fosfato isomerasa, clase I [Citrobacter youngae ATCC 29220] (SEQ ID NO: 28); xilosa isomerasa ZP_06571492.1]), xilosa isomerasa [Citrobacter youngae ATCC 29220] (SEQ ID NO: 29); cepas de *Salmonella* (por ejemplo, *Salmonella enterica* [fosfomanosa isomerasa YP_002146578.1, manosa-6-fosfato isomerasa, clase I [Salmonella enterica subsp. enterica serovar agona cp. SL483] (SEQ ID NO: 30); xilosa isomerasa ZP_02342859.1], xilosa isomerasa [Salmonella enterica subsp. enterica serovar Saintpaul cp. SARA29] (SEQ ID NO: 31); *Salmonella typhimurium* [fosfomanosa isomerasa CAA40399.1]), fosfomanosa isomerasa [Salmonella typhimurium] (SEQ ID NO: 32); cepas de *Enterobacter* (por ejemplo, *Enterobacter cloacae* [phosphomanosa isomerasa ABF71065.1, manosa-6-fosfato isomerasa [Enterobacter cloacae] (SEQ ID NO: 33); xilosa isomerasa YP_003610718.1], xilosa isomerasa [Enterobacter cloacae subsp. cloacae ATCC 13047] (SEQ ID NO: 34);
 30 *Enterobacter cancerogenus* [phosphomanosa isomerasa ZP_05967481.1]), manosa-6-fosfato isomerasa, clase I [Enterobacter cancerogenus ATCC 35316] (SEQ ID NO: 35); cepas de *Cronobacter* (por ejemplo, *Cronobacter turicensis* [fosfomanosa isomerasa YP_003210382.1], manosa-6-fosfato isomerasa [Cronobacter turicensis z3032] (SEQ ID NO: 36); *Cronobacter sakazakii* [fosfomanosa isomerasa YP_001438057.1]), proteína hipotética ESA_01968 [Cronobacter sakazakii ATCC BAA-894] (SEQ ID NO: 37); cepas de *Klebsiella* (por ejemplo, *Klebsiella pneumoniae* [fosfomanosa isomerasa YP_002238770.1, manosa-6-fosfato isomerasa, clase I [Klebsiella pneumoniae 342] (SEQ ID NO: 38); manofructoquinasa YP_002240149.1]), mano(fructo)quinasa [Klebsiella pneumoniae 342] (SEQ ID NO: 39); cepas de *Serratia* (por ejemplo, *Serratia odorifera* [fosfomanosa isomerasa ZP_06640499.1], manosa-6-fosfato isomerasa [Serratia odorifera DSM 4582] (SEQ ID NO: 40).

Además, se pueden utilizar los siguientes genes:

transportador de D-xilosa XylE [Escherichia fergusonii ATCC 35469] YP_002385136.1 (SEQ ID NO: 41);
 Transportador de D-xilosa XylE [Shigella flexneri 2a cp. 301] NP_709888.1 (SEQ ID NO: 42);
 45 cotransportador unidireccional de protones de D-xilosa (transportador D-xylose) [Salmonella enterica subsp. enterica serovar Schwarzengrund cp. SL480], Xyl E, ZP_02664392.1 (SEQ ID NO: 43);
 proteína de la familia de transportadores de azúcar [Citrobacter youngae ATCC 29220], Xyl E, ZP_06356420.1 (SEQ ID NO: 44);
 cotransportador unidireccional de protones de D-xilosa [Acidobacterium capsulatum ATCC 51196], Xyl E, YP_002753901.1 (SEQ ID NO: 45);
 50 xiluloquinasa [Shigella flexneri 2a cp. 301] NP_709351.1 (SEQ ID NO: 46);
 xilulosa quinasa [Citrobacter rodentium ICC168] YP_003367656.1 (SEQ ID NO: 47);
 Subunidad XylF del transportador de D-xilosa [Shigella flexneri 5 cp. 8401] YP_691285.1 (SEQ ID NO: 48);
 Subunidad XylF del transportador de D-xilosa [Citrobacter koseri ATCC BAA-895] YP_001456501.1 (SEQ ID NO: 49);
 55 Transportador ABC de la D-xilosa, Proteína de unión a ATP [Shigella boydii CDC 3083-94] YP_001882237.1, Xyl G, (SEQ ID NO: 50);
 Transportador ABC de la D-xilosa, proteína de unión a la D-xilosa periplásmica [Klebsiella pneumoniae 342], Xyl G, YP_002236058.1 (SEQ ID NO: 51);
 subunidad de unión a ATP del transportador de xilosa [Citrobacter koseri ATCC BAA-895] YP_001456502.1, Xyl G, (SEQ ID NO: 52);
 60 Transportador ABC de la D-xilosa, proteína permeasa [Shigella dysenteriae 1012] ZP_03063638.1, Xyl H, (SEQ ID NO: 53);
 translocador de membrana interna [Enterobacter cloacae subsp. cloacae ATCC13047] YP_003610715.1, Xyl H, (SEQ ID NO: 54);
 65 glucoquinasa [Shigella dysenteriae 1012] ZP_03065183.1 (SEQ ID NO: 55);
 glucoquinasa [Escherichia fergusonii ATCC 35469] YP_002381954.1 (SEQ ID NO: 56);

transaldolasa [Shigella dysenteriae 1012] ZP_03066451.1 (SEQ ID NO: 57);
 transaldolasa B [Escherichia fergusonii ATCC 35469] YP_002381239.1 (SEQ ID NO: 58);
 transcetolasa [Shigella dysenteriae 1012] ZP_03063878.1 (SEQ ID NO: 59);
 transcetolasa [Escherichia fergusonii ATCC 35469] YP_002383968.1 (SEQ ID NO: 60)

se pueden utilizar en la invención.

Las fuentes alternativas para varios genes que codifican las actividades enzimáticas deseadas son, por ejemplo:

Fuentes alternativas para las glucoquinasas:

glucoquinasa [Shigella boydii Sb227] YP_408803.1
 glucoquinasa [Escherichia albertii TW07627] ZP_02901664.1
 glucoquinasa [Salmonella enterica subsp. enterica serovar Typhimurium cp. LT2] NP_461344.1
 glucoquinasa [Citrobacter youngae ATCC 29220] ZP_06351950.1
 Glucoquinasa [Citrobacter rodentium ICC168] YP_003366335.1
 glucoquinasa [Enterobacter cloacae subsp. cloacae NCTC 9394] CBK87115.1
 glucoquinasa [Citrobacter koseri ATCC BAA-895] YP_001452005.1
 glucoquinasa [Cronobacter sakazakii ATCC BAA-894] YP_001436959.1
 glucoquinasa [Enterobacter cancerogenus ATCC 35316] ZP_05968594.1
 glucoquinasa [Klebsiella pneumoniae subsp. pneumoniae MGH 78578] YP_001336384.1
 glucoquinasa [Serratia odorifera DSM 4582] ZP_06639362.1
 Glucoquinasa [Erwinia billingiae Eb661] CAX60782.1
 Glucoquinasa [Erwinia amylovora CFBP1430] YP_003531826.1

Fuentes alternativas para las manofructoquinasas son, por ejemplo:

fructoquinasa [Shigella sonnei Ss046] YP_309383.2
 proteína hipotética CKO_02777 [Citrobacter koseri ATCC BAA-895] YP_001454319.1
 fructoquinasa [Enterobacter sp. 638] YP_001175599.1
 fructoquinasa [Escherichia fergusonii ATCC 35469] YP_002383736.1
 fructoquinasa [Enterobacter cloacae subsp. cloacae ATCC 13047] YP_003611661.1

Fuentes alternativas para el transportador de difusión facilitado, GLF y XylE son, por ejemplo:

transportador XylE de D-xilosa [Shigella sonnei Ss046] YP_312943.1
 cotransportador unidireccional de protones de xilosa [Citrobacter sp. 30_2] ZP_04558597.1
 proteína de la familia de transportadores de azúcar [Citrobacter youngae ATCC 29220] ZP_06356420.1
 Transportador de D-xilosa XylE [Salmonella enterica subsp. arizonae serovar 62: z4, z23: ~] YP_001572426.1
 cotransportador unidireccional de protones de D-xilosa (transportador de D-xilosa) [Citrobacter rodentium ICC168] YP_003367154.1
 transportador XylE de D-xilosa [Mannheimia succiniciproducens MBEL55E] YP_089566.1
 transportador XylE de D-xilosa [Actinobacillus succinogenes 130Z] YP_001343806.1
 transportador XylE de D-xilosa [Bacillus cereus ATCC 10987] NP_978526.1
 permeasa de xilosa [Bacillus megaterium DSM319] YP_003597063.1

Debido a la alta homología entre la proteína de difusión facilitada para la absorción de manosa (EC1.3.1.74), en particular GLF, y el cotransportador unidireccional E de protones de xilosa, xyl E, los genes que codifican el cotransportador unidireccional E de protones de xilosa, xyl E, pueden utilizarse en lugar de los genes que codifican la proteína de difusión facilitada para la absorción de manosa, y viceversa.

Fuentes alternativas para las xiluloquinasas son, por ejemplo:

xiluloquinasa [Shigella sp. D9] ZP_05434777.1
 xiluloquinasa [Escherichia fergusonii ATCC 35469] YP_002384640.1
 xiluloquinasa [Citrobacter sp. 30_2] ZP_04559359.1
 xiluloquinasa [Shigella boydii Sb227] YP_409882.1
 xilulosa quinasa [Salmonella enterica subsp. enterica serovar Paratyphi C, cepa RKS4594] YP_002639258.1
 proteína hipotética CKO_05020 [Citrobacter koseri ATCC BAA-895] YP_001456499.1
 xiluloquinasa [Shigella sonnei Ss046] YP_312598.1
 xiluloquinasa [Citrobacter youngae ATCC 29220] ZP_06355683.1
 xiluloquinasa [Klebsiella sp. 1_1_55] ZP_06551221.1
 proteína hipotética ECL_00202 [Enterobacter cloacae subsp. cloacae ATCC 13047] YP_003610719.1
 xiluloquinasa [Klebsiella pneumoniae 342] YP_002236061.1
 xiluloquinasa [Klebsiella variicola At-22] YP_003437118.1
 Xilulosa quinasa [Cronobacter turicensis z3032] YP_003212435.1

Xilulosa quinasa [*Yersinia intermedia* ATCC 29909] ZP_04636833.1
 xiluloquinasa [*Erwinia billingiae* Eb661] CAX60830.1
 xiluloquinasa [*Serratia odorifera* DSM 4582] ZP_06638097.1

5 **Fuentes alternativas para las XylF son, por ejemplo:**

Subunidad XylF del transportador de D-xilosa [*Shigella boydii* Sb227] YP_409884.1
 Subunidad XylF del transportador de D-xilosa [*Enterobacter* sp. 638] YP_001174896.1
 Subunidad XylF del transportador de D-xilosa [*Cronobacter sakazakii* ATCC BAA-894] YP_001440178.1
 10 Proteína periplásmica de unión a D-xilosa [*Cronobacter turicensis* z3032] YP_003212438.1
 Subunidad XylF del transportador de D-xilosa [*Enterobacter cloacae* subsp. *cloacae* ATCC 13047] YP_003610717.1
 proteína hipotética ENTCAN_08293 [*Enterobacter cancerogenus* ATCC 35316] ZP_05969667.2
 Transportador ABC de la D-xilosa, proteína de unión al sustrato periplásmico [*Erwinia billingiae* Eb661] CAX60828.1
 15 transportador de casete de unión a ATP de la superfamilia ABC de la xilosa, proteína de unión [*Serratia odorifera* DSM 4582] ZP_06638100.1
 subunidad XylF del transportador de d-xilosa [*Actinobacillus succinogenes* 130Z] YP_001343809.1

20 **Fuentes alternativas para las XylG son, por ejemplo:**

subunidad de unión a ATP del transportador de xilosa [*Shigella flexneri* 5 cp. 8401] YP_691284.1
 proteína de unión a ATP del transportador ABC de la D-xilosa [*Shigella dysenteriae* 1012] ZP_03063359.1
 subunidad de unión a ATP del transportador de xilosa [*Enterobacter cloacae* subsp. *cloacae* ATCC 13047] YP_003610716.1
 25 proteína de unión a ATP del transportador ABC de la D-xilosa [*Klebsiella pneumoniae* 342] YP_002236058.1
 proteína hipotética ENTCAN_08292 [*Enterobacter cancerogenus* ATCC 35316] ZP_05969666.1
 subunidad de unión a ATP del transportador de xilosa [*Cronobacter sakazakii* ATCC BAA-894] YP_001440177.1
 30 proteína xylG de unión a ATP de importación de xilosa [*Cronobacter turicensis* z3032] YP_003212439.1
 subunidad ATPasa del transportador ABC de la D-xilosa [*Erwinia billingiae* Eb661] CAX60827.1 proteína de unión del transportador de casete de unión a la superfamilia ATP de ABC de la xilosa [*Serratia odorifera* DSM 4582] ZP_06638101.1
 COG1129: componente ATPasa del sistema de transporte de azúcar de tipo ABC [*Actinobacillus pleuropneumoniae* serovar 1 cp. 4074] ZP_00134918.2
 35

Fuentes alternativas para las XylH son, por ejemplo:

proteína hipotética CKO_05024 [*Citrobacter koseri* ATCC BAA-895] YP_001456503.1
 40 proteína de membrana del transportador ABC de xilosa [*Enterobacter cloacae* subsp. *cloacae* ATCC 9394] CBK87790.1
 proteína hipotética ENTCAN_08291 [*Enterobacter cancerogenus* ATCC 35316] ZP_05969665.1
 translocador de membrana interna [*Enterobacter* sp. 638] YP_001174894.1 proteína hipotética ESA_04159 [*Cronobacter sakazakii* ATCC BAA-894] YP_001440176.1
 45 proteína permeasa del transportador ABC de la D-xilosa [*Klebsiella pneumoniae* 342] YP_002236057.1
 proteína xylH permeasa del sistema de transporte de xilosa [*Cronobacter turicensis* z3032] YP_003212440.1
 proteína permeasa del sistema de transporte de azúcar [*Yersinia enterocolitica* subsp. *enterocolitica* 8081] YP_001008391.1
 transportador ABC de la D-xilosa, proteína permeasa [*Erwinia billingiae* Eb661] CAX60826.1
 50 ATPasa transportadora de monosacáridos [*Serratia proteamaculans* 568] YP_001476339.1
 transportador de casete de unión a ATP de la superfamilia ABC de la xilosa, proteína de membrana [*Serratia odorifera* DSM 4582] ZP_06638102.1
 ATPasa transportadora de monosacáridos [*Actinobacillus succinogenes* 130Z] YP_001343807.1

55 **Fuentes alternativas para transaldolasa son, por ejemplo**

transaldolasa B [*Shigella boydii* Sb227] YP_406573.1
 transaldolasa [*Citrobacter youngae* ATCC 29220] ZP_06354813.1
 transaldolasa B [*Citrobacter koseri* ATCC BAA-895] YP_001454897.1
 60 transaldolasa B [*Salmonella enterica* subsp. *arizonae* serovar 62:z4,z23:-- YP_001571972.1
 transaldolasa B [*Citrobacter rodentium* ICC168] YP_003363651.1
 transaldolasa B [*Enterobacter* sp. 638] YP_001175308.1
 proteína hipotética ENTCAN_05194 [*Enterobacter cancerogenus* ATCC 35316] ZP_05966848.1
 transaldolasa B [*Citrobacter* sp. 30_2] ZP_04559788.1
 65 transaldolasa [*Enterobacter cloacae* subsp. *cloacae* NCTC 9394] CBK87626.1
 transaldolasa B [*Klebsiella pneumoniae* 342] YP_002240539.1

transaldolasa B [*Enterobacter cloacae* subsp. *cloacae* ATCC 13047] YP_003611333.1
transaldolasa B [*Cronobacter sakazakii* ATCC BAA-894] YP_001439387.1
transaldolasa B [*Cronobacter turicensis* z3032] YP_003208998.1
transaldolasa [*Escherichia albertii* TW07627] ZP_02902842.1
transaldolasa [*Yersinia kristensenii* ATCC 33638] ZP_04625574.1
transaldolasa B [*Serratia odorifera* DSM 4582] ZP_06637249.1
transaldolasa B [*Erwinia amylovora* CFBP1430] YP_003532302.1
transaldolasa B [*Erwinia amylovora* ATCC 49946] YP_003537736.1
transaldolasa B [*Erwinia pyrifoliae* DSM 12163] CAY73102.1

Transcetolasas alternativas son, por ejemplo:

transcetolasa [*Shigella flexneri* 2a cp. 2457T] NP_838419.1
transcetolasa [*Shigella boydii* CDC 3083-94] YP_001881708.1
transcetolasa [*Escherichia albertii* TW07627] ZP_02900973.1
transcetolasa [*Citrobacter koseri* ATCC BAA-895] YP_001455802.1
transcetolasa [*Citrobacter* sp. 30_2] ZP_04560349.1
transcetolasa [*Citrobacter youngae* ATCC 29220] ZP_06355044.1
transcetolasa [*Enterobacter cancerogenus* ATCC 35316] ZP_05970189.1
transcetolasa [*Enterobacter cloacae* subsp. *cloacae* NCTC 9394] CBK86553.1
transcetolasa [*Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578] YP_001336984.1
transcetolasa 1 [*Citrobacter rodentium* ICC168] YP_003368408.1
transcetolasa [*Klebsiella variicola* At-22] YP_003440166.1
transcetolasa [*Klebsiella pneumoniae* 342] YP_002239249.1
transcetolasa 1 [*Cronobacter turicensis* z3032] YP_003211835.1
transcetolasa [*Cronobacter sakazakii* ATCC BAA-894] YP_001436540.1
transcetolasa [*Enterobacter cloacae* subsp. *cloacae* ATCC 13047] YP_003614730.1
transcetolasa [*Erwinia tasmaniensis* Et1/99] YP_001908733.1
transcetolasa [*Serratia odorifera* DSM 4582] ZP_06641778.1

Los números de referencia se derivan de la base de datos NCBI (Centro Nacional de Información Biotecnológica, Biblioteca Nacional de Medicina de los Estados Unidos, 8600 Rockville Pike, Bethesda MD, 20894 USA). La fecha de entrega de los números de referencia fue el 2 de junio de 2010.

Se desvelan proteínas codificadas por cualquiera de los genes enumerados anteriormente, en particular, aquellos que comprenden las secuencias de aminoácidos SEQ ID NO: 1 a 60, preferentemente las secuencias SEQ ID NO: 1 a 12.

Es evidente que pequeñas variaciones en la secuencia de nucleótidos de un gen no cambian significativamente las propiedades catalíticas de la proteína codificada. Por ejemplo, muchos cambios en la secuencia de nucleótidos no cambian la secuencia de aminoácidos de la proteína codificada. Además, una secuencia de aminoácidos puede tener variaciones, que no cambian las propiedades funcionales de una proteína, en particular, no impiden que una enzima lleve a cabo su función catalítica. Como es conocido por un experto en la técnica, pueden existir variaciones o pueden realizarse en regiones que están fuera de la región crítica para la función de la enzima y todavía dan como resultado un polipéptido activo. Dichas variaciones en la secuencia de nucleótidos o las moléculas de ADN o en una secuencia de aminoácidos se conocen como "equivalentes funcionales", porque no cambian significativamente la función del gen para codificar una proteína con una función particular, por ejemplo, catalizar una reacción particular o, respectivamente, cambiar la función particular de la proteína. Dentro del alcance de la presente invención están los equivalentes funcionales, que incluyen fragmentos o derivados, u homólogos más cercanos de las proteínas enumeradas, en particular, los equivalentes funcionales de las secuencias de SEQ ID NO: 1 a 60, preferentemente las secuencias de SEQ ID NO: 1 a 12.

Se divulgan genes que codifican secuencias de aminoácidos que muestran al menos el 50 %, preferentemente al menos el 60 % de identidad, preferentemente al menos el 65 %, preferentemente al menos el 70 %, preferentemente al menos el 75 %, preferentemente al menos el 80 %, preferentemente al menos el 85 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 %, más y más preferiblemente al menos el 98 % de identidad con cualquiera de las secuencias de aminoácidos de las proteínas enumeradas anteriormente, en particular las secuencias de SEQ ID NO: 1 a 60, preferentemente las secuencias de SEQ ID NO: 1 a 12.

Se divulgan genes que codifican proteínas que comprenden cualquiera de las secuencias de aminoácidos de SEQ ID NO: 1 a 60, preferentemente las secuencias de SEQ ID NO: 1 a 12.

Se divulgan también genes que codifican secuencias de aminoácidos que muestran al menos el 50 %, preferentemente al menos el 60 % de identidad, preferentemente al menos el 65 %, preferentemente al menos el 70 %, preferentemente al menos el 75 %, preferentemente al menos el 80 %, preferentemente al menos el 85 %, más preferentemente al menos el 90 %, aún más preferentemente al menos el 95 %, más y más preferiblemente al menos el 98 % de identidad con cualquiera de las secuencias de aminoácidos de SEQ ID NO: 1 a 60, preferentemente las secuencias de SEQ ID

NO: 1 a 12. El término "identidad" se refiere a la identidad entre dos secuencias de ácido nucleico o aminoácidos, respectivamente comparadas entre sí desde el primer ácido nucleico hasta el último ácido nucleico o desde el primer aminoácido codificado por el gen correspondiente hasta el último aminoácido. La identidad de las secuencias de longitud completa se puede medir utilizando el programa BLAST (Altschul, S. F., T.L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller y D. J. Lipman 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402). En la comparación se utilizan preferentemente las secuencias maduras de las proteínas.

Las bacterias que pertenecen al género *Cupriavidus* o *Ralstonia* y que son capaces de expresar genes que codifican la fosfomanosa isomerasa y la proteína de difusión facilitada para la absorción de manosa o el cotransportador unidireccional E de protones de xilosa y, opcionalmente, la manofructoquinasa son capaces de crecer sobre la manosa como fuente de carbono. Un hospedador *Cupriavidus* introducido para expresar un gen que codifica una proteína de difusión facilitada para la actividad de absorción de manosa, en particular *glf*, también puede utilizar glucosa como fuente de carbono.

La manosa se puede encontrar en la madera y en los materiales vegetales, especialmente en hemicelulosas de maderas blandas (por ejemplo, pino o abeto) que consiste en polímero de galactoglucomanano, en algunas frutas o bayas (por ejemplo, arándano), y en estructuras celulares de diversos microorganismos como hongos (levaduras y hongos filamentosos) y microalgas (por ejemplo, *Chlorella*, *Nannochloropsis* y *Phaeodactylum*).

Las bacterias que pertenecen al género *Cupriavidus* o *Ralstonia* y que son capaces de expresar genes que codifican xilosa isomerasa y xiluloquinasa y un transportador ABC de alta afinidad o cotransportador unidireccional E de protones de xilosa, son capaces de crecer en xilosa o en xilosa y arabinosa como fuente de carbono. Un hospedador *Cupriavidus* introducido para expresar genes, que codifican el transportador ABC de alta afinidad, en particular los genes que codifican Xyl F, Xyl G y Xyl H, es capaz de utilizar también glucosa y galactosa como fuente de carbono.

Además, el microorganismo puede modificarse genéticamente para expresar transaldolasa y/o transcetolasa, lo que acelera el crecimiento del microorganismo.

Como se describe en el presente documento, un microorganismo introducido para expresar un gen que codifica una proteína de difusión facilitada para la absorción de manosa es capaz de utilizar también glucosa. Esto se logra en particular por el gen *glf*.

El microorganismo hospedador también es capaz de utilizar glucosa, si se introduce al hospedador para expresar un gen que codifica glucoquinasa (EC 2.7.1.2).

El microorganismo de acuerdo con esta divulgación es capaz de utilizar xilosa y arabinosa y, opcionalmente, glucosa y/o galactosa como fuente de carbono. En una realización, un microorganismo de acuerdo con esta divulgación es capaz de utilizar manosa y, opcionalmente, glucosa como fuente de carbono.

Un microorganismo introducido para expresar xilosa isomerasa, xiluloquinasa y un transportador ABC de alta afinidad o cotransportador unidireccional E de la proteína de la xilosa, es capaz de utilizar xilosa. Sorprendentemente, el microorganismo hospedador es capaz de utilizar también arabinosa. Para el mejor conocimiento de los presentes inventores, esta es la primera vez que un hospedador *Ralstonia* o *Cupriavidus* obtiene la capacidad de utilizar arabinosa. En la técnica anterior, Buchholz et al. (1994) transfirieron un hospedador *Alcaligenes eutrophus* mediante un inserto de 28 kb, de los cuales 16,4 comprendían la información mínima requerida para la utilización de xilosa por *A. eutrophus*. Buchholz et al. asumieron que los genes *xyl* que codifican la xilosa isomerasa y la xiluloquinasa estaban localizados dentro de esta región, ya que se indujeron durante el crecimiento en la xilosa. Buchholz et al. obtuvieron tres transconjugantes que crecieron en xilosa como única fuente de carbono y energía, pero no se obtuvieron cepas que utilizan ribosa o arabinosa. Como se describe en el presente documento, el microorganismo hospedador se ha introducido para expresar insertos que consisten esencialmente en secuencias de nucleótidos necesarias para la utilización de xilosa. Sorprendentemente, las construcciones hacen que el hospedador sea capaz de utilizar también arabinosa.

Un microorganismo introducido para expresar un transportador ABC de alta afinidad, en particular los genes que codifican Xyl F, Xyl G y Xyl H, es sorprendentemente capaz de utilizar también glucosa y galactosa como fuente de carbono.

La actividad de MAK (manofructoquinasa), PMI (fosfomanosa isomerasa), GLK (glucoquinasa) se puede analizar mediante métodos bien conocidos en la técnica. Las fracciones de proteínas solubles de la cepa hospedadora recombinante se pueden aplicar para las mediciones de la actividad de MAK, PMI y GLK a 30 °C utilizando un espectrofotómetro Nicolet Evolution 100 UV/VIS (Thermo Electron Corporation, Cambridge, UK). La actividad de MAK se puede analizar con un sistema acoplado a NADH como lo describen Sebastian y Asensio (1972) con algunas modificaciones como se describe en los ejemplos.

La actividad de PMI se puede medir con un ensayo acoplado a NADP (Kang y Markovitz 1967) con algunas

modificaciones como se describe en los ejemplos.

La actividad de GLK se puede medir con un sistema acoplado a NADPH de acuerdo con el método de Gottschalk (Gottschalk et al. 1964) con modificaciones como se describe en los ejemplos.

La actividad de XylA y XylB se puede ensayar mediante métodos bien conocidos en la técnica. Las fracciones de proteínas solubles de la cepa hospedadora recombinante se pueden aplicar para las mediciones de la actividad de XylA y XylB a 30 °C utilizando un espectrofotómetro Nicolet Evolution 100 UV/VIS (Thermo Electron Corporation, Cambridge, UK). La actividad de XylA (Gao et al. 2002, Mejnen et al. 2008) o XylB (Shamanna y Sangerson 1979, Eliasson et al. 2000) se puede analizar con un sistema acoplado a NADH con algunas modificaciones, utilizando sorbitol deshidrogenasa o piruvato quinasa y lactato deshidrogenasa, respectivamente, como enzimas auxiliares como se describe en los ejemplos

Los "elementos reguladores" se refieren aquí a elementos reguladores que pueden regular la expresión de un gen extraño introducido en una célula hospedadora, en particular en *Cupriavidus* o *Ralstonia*. Los elementos reguladores incluyen promotores, terminadores, potenciadores y secuencias señal.

"Vector" se refiere aquí a una molécula de ácido nucleico, normalmente una molécula de ADN, que se puede utilizar para transferir material genético extraño a una célula hospedadora. Un vector puede ser un plásmido, bacteriófago, virus o cósmido, aquí, normalmente, es un plásmido. Preferentemente, el vector es un vector de amplia gama de hospedadores. En una realización, el vector puede ser un vector episomal no incorporado al genoma de la célula hospedadora. En otra realización, el vector puede integrarse en el genoma de la célula hospedadora. Los plásmidos de amplia gama de hospedadores permiten la replicación en diversas especies bacterianas y no están restringidos a bacterias de un determinado género o especie.

Un vector de la presente invención comprende una secuencia de nucleótidos que codifica xilosa isomerasa (EC 5.3.1.5), una secuencia de nucleótidos que codifica xiluloquinasa (EC 2.7.1.17) y una secuencia de nucleótidos que codifica Xyl G (EC 3.6.3.17), Xyl F y Xyl H, y elementos reguladores para regular la expresión de dichos genes en *Cupriavidus* o *Ralstonia*.

En una realización de la invención, un vector puede comprender un gen que codifica fosfomanosa isomerasa, un gen que codifica la proteína de difusión facilitada para la absorción de manosa, un gen que codifica el cotransportador unidireccional E de protones de xilosa y/o un gen que codifica la manofructoquinasa. Los genes pueden estar en el mismo vector o en vectores separados.

En una realización de la invención, el vector puede ser un vector de amplia gama de hospedadores. Esto proporciona la ventaja de que la extensión de la gama de sustrato se puede lograr con los vectores de amplia gama de hospedadores en cualquier especie o cepa de *Cupriavidus* o *Ralstonia*.

Preferentemente, el vector es un vector de alto número de copias.

Además, el vector comprende elementos reguladores, en particular, un promotor, para regular la expresión de dichos genes en *Cupriavidus* o *Ralstonia*. Preferentemente, el vector comprende un potente promotor para la expresión de los genes deseados en *Cupriavidus* o *Ralstonia*.

En una realización, los elementos reguladores en el vector pueden comprender el promotor de *neokanamicina*, el promotor *lac* o el promotor de gliceraldehído-3-fosfato deshidrogenasa. Además, se pueden utilizar los siguientes promotores: promotor P2, normalmente de *Pseudomonas syringae* (Solaiman y Swingle 2010), en particular del plásmido pBS29-P2 (Solaiman et al. 2010), el promotor de la hidrogenasa (SH), normalmente de *R. eutropha* (Porthun et al. 2002, Schwartz et al. 1998), y el promotor P_L, normalmente del operón *cbb* (Jeffke et al. 1999, Schaferjohann et al. 1996, Windhövel and Bowien 1990). Preferentemente, los promotores son constitutivos.

Los promotores adecuados son aquí en particular un promotor *lac* constitutivo (P_{lac}) (Siedow et al. 1999), *promotor de neokanamicina* (P_{nk}), promotor de gliceraldehído-3-fosfato deshidrogenasa (GAPDH) (P_{GAPDH}).

Los vectores, adecuados para utilizar en la invención son, por ejemplo, pBBR1MCS y sus derivados pBBR1MCS-2, pBBR1MCS-3, pBBR1MCS-4 y pBBR1MCS-5 (Kovach et al. 1995).

Además, se puede utilizar el plásmido lanzadera pBHR1 (MoBiTec GmbH, Göttingen, Alemania) (Aneja et al. 2009) en la invención.

El plásmido pKT230 (Park et al. 1995), que procede del plásmido de amplia gama de hospedadores RSF1010 (Franklin et al. 1981) también se puede utilizar en la invención.

También se puede utilizar en esta invención el plásmido pBS29 (Swingle et al. 2008) o pBS29-P2 (Solaiman et al. 2010), que proceden del plásmido pUCP26 capaz de replicarse en especies de *E. coli* y *Pseudomonas* (West et al.

1994).

Dentro del alcance de la presente invención también está la producción de productos químicos deseados seleccionados de lípidos, polihidroxialcanoatos, ácidos grasos y derivados de ácidos grasos que utilizan los microorganismos modificados de acuerdo con la invención. Los productos químicos deseados pueden ser productos producidos de forma natural por el hospedador, tales como polihidroxialcanoatos (Reinecke y Steinbüchel 2009 y Verlinden et al. 2007) o el hospedador puede introducirse para expresar los productos químicos deseados, por ejemplo, los lípidos. En el documento WO 2007136762 se han descrito los genes que se pueden utilizar para modificar genéticamente los microorganismos hospedadores para expresar diversos lípidos. La capacidad de un hospedador *Cupriavidus* o *Ralstonia* para producir lípidos se ejemplificó mediante la introducción de un hospedador *Cupriavidus* o *Ralstonia* para expresar ésteres etílicos de ácidos grasos (FAEE, de sus siglas en inglés). Esto se llevó a cabo mediante la expresión de en un mutante negativo a polihidroxibutirato (PHB) de genes del hospedador H16 de *R. eutropha* que codifican la cera éster sintasa/diacilglicerol aciltransferasa (*atfA*), la piruvato descarboxilasa (*pdc*) y el alcohol deshidrogenasa B (*adhB*).

El beneficio de esta invención es una amplia gama de utilización de sustrato de *Cupriavidus* o *Ralstonia* hacia carbohidratos abundantes y de bajo costo. La extensión de la gama de sustrato se logró mediante la utilización de plásmidos de amplia gama de hospedadores, lo que hace posible extender la gama de sustrato en cualquier especie o cepa del hospedador *Cupriavidus* o *Ralstonia*. Mediante métodos químicos es posible introducir una mutación solo en una cepa específica. Además, en mutantes inducidos químicamente existe el riesgo de mutaciones adicionales que pueden afectar la productividad y/o pueden llevar a efectos secundarios desfavorables.

La presente invención se ejemplificó mediante la construcción de dos cepas recombinantes diferentes de *Ralstonia*, más específicamente dos cepas H16 de *R. eutropha*. Las cepas fueron desarrolladas para ser capaces de utilizar manosa/glucosa y xilosa/arabinosa, respectivamente, mediante ingeniería metabólica. La inducción de diferentes genes en *R. eutropha*, necesaria para la absorción de glucosa/manosa y la utilización de manosa, o para la absorción y utilización de xilosa, respectivamente, complementó las partes carentes de catabolismo de glucosa/manosa o xilosa/arabinosa, respectivamente. Las cepas recombinantes de *R. eutropha* con amplia gama de utilización del sustrato sintetizaron poli (3-hidroxibutirato) PHB a partir de glucosa, manosa, xilosa y arabinosa en cantidades similares a las de gluconato.

En una realización de la invención, en *Cupriavidus* o *Ralstonia* pueden expresarse genes, que responden a la absorción y conversión de azúcar. Las hexoquinazas de otros procariotas, como *Leuconostoc mesenteroides* o *Streptomyces* sp., que también muestran una actividad para la manosa (Anderson y Sapico 1975, Coulombel et al. 1982), pueden expresarse en *Cupriavidus* o *Ralstonia* para mejorar la conversión de azúcar en el hospedador. Se puede utilizar una sobreexpresión de la hexoquinasa codificada cromosómicamente con su actividad secundaria para la manosa para mejorar la utilización de la manosa en *Ralstonia*. Los plásmidos pBBR utilizados en los ejemplos se pueden reemplazar, por ejemplo, con vectores de alto número de copias y/o con promotores más fuertes (Kovach et al. 1995, Siedow et al. 1999), u otros vectores más adecuados para la expresión de los genes deseados para la utilización de la manosa. El promotor *lac* puede reemplazarse con un promotor constitutivo más fuerte, por ejemplo, el promotor de *neokanamicina* como se describe en el presente documento y que da niveles de expresión elevados. Los genes deseados también pueden integrarse en el genoma de *Cupriavidus* o *Ralstonia* lo que da ventajas, como la estabilidad. Además, los mutantes que contienen integraciones libres de marcadores en sus genomas no están sujetos a ninguna selección, lo que ahorra costos especialmente para los cultivos a gran escala.

De acuerdo con una realización adicional, la presente invención proporciona un método, que comprende que el microorganismo puede sintetizar productos químicos a granel seleccionados de lípidos, polihidroxialcanoatos, ácidos grasos y derivados de ácidos grasos, y el producto químico se recupera del medio de cultivo o de las células.

En el cultivo de *Cupriavidus* o *Ralstonia* se puede utilizar cualquier medio de cultivo que comprenda componentes adecuados para el crecimiento de *Cupriavidus* o *Ralstonia*. Normalmente, incluyen una fuente de carbono y/o de nutrientes, sales, normalmente sales minerales y agua. Los marcadores de selección, como los genes de resistencia a antibióticos, pueden ser necesarios para confirmar la presencia de los genes deseados en las células hospedadoras.

La utilización del sustrato de *Ralstonia* se ha ejemplificado aquí mediante el cultivo de *Cupriavidus* o *Ralstonia* en un medio de sales minerales (MSM, de sus siglas en inglés) descrito por Schlegel et al. (1961). En aplicaciones industriales, el medio de cultivo comprende normalmente azúcares monoméricos procedentes de materiales lignocelulósicos, como fuente de carbono, y además posiblemente nutrientes y sales o componentes de los mismos como micro y macronutrientes, y agua.

En el cultivo de *Cupriavidus* o *Ralstonia* se pueden utilizar condiciones de cultivo adecuadas para el cultivo de *Cupriavidus* o *Ralstonia*. Tales condiciones de cultivo son bien conocidas por un experto en la técnica. En el cultivo se pueden utilizar medios líquidos o sólidos. Los cultivos líquidos pueden agitarse, agitación adecuada de 50 a 1000 rpm, normalmente de 100 a 600 rpm. La temperatura de cultivo es normalmente de 15 a 45, preferentemente de 25 a 37 °C, normalmente 30 °C. El tiempo de incubación puede ser de 1 día a 14 días, normalmente de 2 a 7 días, preferentemente de 24 a 72 horas. El cultivo puede ser un cultivo discontinuo, semicontinuo o continuo.

Se cultivaron *Escherichia coli* y *Zymomonas mobilis* mediante métodos bien conocidos por los expertos en la técnica y se describen en los ejemplos.

Como se describe en el presente documento, las cepas de *Ralstonia* recombinantes y las cepas de tipo silvestre se cultivaron en diferentes fuentes de carbono. Se incubaron células en placas de agar MSM que contenían el 1% (p/v) de la fuente de carbono respectiva durante 5 días a 30 °C. Se examinó fenotípicamente el crecimiento de colonias visibles. Caracteres: (++) crecimiento después de un tiempo de incubación de 24-72 h, (+) crecimiento después de más de 72 h de incubación, (-) no se detecta crecimiento visible. Se examinó la utilización de las siguientes fuentes de carbono: L-Arabinosa, D-fructosa, D-galactosa, D-gluconato, D-glucosa, D-glicerol, D-maltosa-manitol, D-manosa, N-acetil D-glucosamina, piruvato, D, ribosa, D-trehalosa y D-xilosa.

La actividad de MAK (manofructoquinasa), PMI (fosfomanosa isomerasa), GLK (glucoquinasa) se puede analizar mediante métodos bien conocidos en la técnica. Las fracciones de proteínas solubles de la cepa hospedadora recombinante se pueden aplicar para mediciones de actividad de MAK, PMI y GLK a 30 °C utilizando un espectrofotómetro Nicolet Evolution 100 UV/VIS (Thermo Electron Corporation, Cambridge, UK). La actividad de MAK se puede analizar con un sistema acoplado a NADH como lo describen Sebastian y Asensio (1972) con algunas modificaciones como se describe en los ejemplos.

La actividad de PMI se puede medir con un ensayo acoplado a NADP (Kang y Markovitz 1967) con algunas modificaciones como se describe en los ejemplos.

La actividad de GLK se puede medir con un sistema acoplado a NADPH de acuerdo con el método de Gottschalk (Gottschalk et al. 1964) con modificaciones como se describe en los ejemplos.

La presente invención se ha ejemplificado mediante la generación de recombinantes de la cepa H16 de *R. eutropha* que, en contraste con el tipo silvestre, fueron capaces de utilizar manosa y/o glucosa como fuentes únicas de carbono. Hasta cuatro genes (*glf*, *glk*, *mak*, *pmi*) se introdujeron episomalmente en la cepa H16 de *R. eutropha* bajo el control de un promotor constitutivo de *neokanamicina* o *lac*, lo que permite la absorción y conversión de glucosa y manosa en intermedios de la ruta KDPG.

Las bacterias que pertenecen a los géneros *Ralstonia* y *Cupriavidus* normalmente tienen un espectro de utilización de sustratos muy estrecho y son solo una cantidad limitada de azúcares. Por ejemplo, la capacidad de H16 de *Ralstonia eutropha* para utilizar carbohidratos y compuestos relacionados se limita a la fructosa, N-acetilglucosamina y gluconato. Para ampliar la gama de utilización de sustrato de las cepas *Ralstonia* y *Cupriavidus*, las cepas fueron diseñadas para utilizar la manosa como única fuente de carbono para el crecimiento. Los genes para una proteína de difusión facilitada (*glf*) de *Zymomonas mobilis* y para manofructoquinasa (*mak*), fosfomanosa isomerasa (*pmi*) y glucoquinasa (*glk*) de *E. coli*, ya sea solo o en combinación, se expresan de forma constitutiva en *R. eutropha* bajo el control del promotor *lac*, utilizando un vector de amplia gama de hospedadores episomal. Las cepas recombinantes que albergan pBBR1MCS-3::*glf::mak::pmi* o pBBR1MCS-3::*glf::pmi* crecieron en manosa, mientras que pBBR1MCS-3::*glf::mak* y pBBR1MCS-3::*glf* no confirieron la capacidad de utilizar manosa como fuente de carbono para *R. eutropha*. Las cepas recombinantes que albergan pBBR1MCS-3::*glf::mak::pmi* mostraron un crecimiento más rápido que pBBR1MCS-3::*glf::pmi* en la manosa. Estos datos indicaron que la fosfomanosa isomerasa es necesaria para convertir la manosa-6-fosfato en fructosa-6-fosfato para el catabolismo posterior a través de la ruta Entner-Doudoroff. Además, sorprendentemente, todos los plásmidos confirieron también a *Ralstonia* la capacidad de crecer en presencia de glucosa.

Además, la expresión de las enzimas respectivas se demostró a nivel transcripcional y mediante la medición de las actividades de manofructoquinasa (0,14 U mg⁻¹ proteína) y fosfomanosa isomerasa (0,25 U mg⁻¹ proteína).

Para identificar genes en el genoma de H16 de *R. eutropha* que codifican enzimas funcionales requeridas para la absorción y el catabolismo de la D-xilosa, se realizó un análisis BLAST proteína-proteína (Altschul et al. 1990) utilizando una base de datos de secuencias de aminoácidos de H16 de *R. eutropha* (Pohlmann et al. 2006). No se pudieron identificar genes que codifiquen un transportador específico para D-xilosa (Pohlmann et al. 2006). Además, no hay genes que codifiquen enzimas que catabolicen la D-xilosa, por ejemplo, se identificó una xilosa isomerasa o xilulosa quinasa. De hecho, H16 de *R. eutropha* posee genes para una fructoquinasa y una glucoquinasa (Pohlmann et al. 2006). Todos los otros genes, que codifican enzimas de la ruta de fosfato de pentosa no oxidativa para un catabolismo adicional de D-xilulosa-5-fosfato, están presentes en la cepa H16 de *R. eutropha*.

La construcción de los hospedadores *Ralstonia* para utilizar la xilosa como única fuente de carbono para el crecimiento se ejemplificó mediante la utilización de las cepas de *R. eutropha*. Los genes para un cotransportador unidireccional E de protones de D-xilosa (*xylE*) de baja afinidad, un transportador ABC de alta afinidad (*xylFGH*), una xilosa isomerasa (*xylA*) y una xiluloquinasa (*xylB*), de *E. coli* se combinaron expresados constitutivamente en hospedadores *Ralstonia* bajo el control de un promotor *lac* y un promotor gliceraldehído-3-fosfato deshidrogenasa (GADPH), respectivamente, utilizando un vector de amplia gama de hospedadores episomal. Las cepas recombinantes que albergan pBBR1MCS-3::*xylABFGH* crecieron en xilosa, lo que confiere la capacidad de utilizar la xilosa como fuente de carbono para *Ralstonia*. Además, sorprendentemente, los plásmidos que confieren la utilización de xilosa a *Ralstonia* confirieron

también la capacidad de crecer en presencia de arabinosa.

Para ejemplificar la capacidad de las cepas de *R. eutropha* recombinantes para producir PHB a partir de manosa o glucosa, se cultivaron células de *Ralstonia* bajo condiciones permisivas para la acumulación de PHB en MSM que contiene manosa al 1 % (p/v), glucosa o gluconato de sodio como única fuente de carbono que contiene tetraciclina o kanamicina, respectivamente, para la selección y NH_4Cl al 0,05 % (p/v). Se analizaron los contenidos de PHB de las células mediante cromatografía de gases como se describe en la sección de materiales y métodos. Los cultivos que contienen gluconato como única fuente de carbono sirvieron como referencia para la cepa recombinante, y se tomaron muestras en la fase de crecimiento estacionario temprana. El análisis de GC de las células reveló un contenido de PHB del 63 % (p/p) de masa seca celular. Las células cultivadas en presencia de glucosa o manosa al 1 % (p/v) también se examinaron en la fase de crecimiento estacionario, y se determinaron los contenidos de PHB de 65 y 67 % (p/p) de masa seca celular, respectivamente.

Para ejemplificar la capacidad de las cepas recombinantes de *R. eutropha* para producir PHB a partir de xilosa y arabinosa, se cultivaron células bajo condiciones permisivas para la acumulación de PHB en MSM que contiene D-xilosa, L-arabinosa o gluconato de sodio al 1 % (p/v) como carbono fuente que contiene tetraciclina, respectivamente, para la selección y NH_4Cl al 0,05 % (p/v). Se analizaron los contenidos de PHB de las células mediante cromatografía de gases como se describe en los ejemplos en la sección de materiales y métodos. Los cultivos que contienen gluconato como única fuente de carbono sirvieron como referencia para la cepa recombinante, y se tomaron muestras en la fase de crecimiento estacionario temprana. El análisis de GC de las células reveló que las células eran capaces de producir PHB. Las células cultivadas en presencia de L-arabinosa o D-xilosa al 1 % (p/v) también se examinaron en la fase de crecimiento estacionario, y se mostró la producción de PHB.

Las cepas recombinantes de *R. eutropha* con amplia gama de utilización del sustrato fueron capaces así de sintetizar poli (3-hidroxibutirato) PHB a partir de glucosa, manosa, xilosa y arabinosa en cantidades similares a las de gluconato.

Se desvelan otros artículos.

Artículo 1. Un microorganismo del género *Cupriavidus* o *Ralstonia* modificado genéticamente para expresar:

a) fosfomanosa isomerasa (EC5.3.1.8) y proteína de difusión facilitada para la absorción de manosa (EC1.3.1.74) o fosfomanosa isomerasa (EC5.3.1.8) y cotransportador unidireccional E de protones de xilosa, y opcionalmente manofructoquinasa (EC2.7.1.4), en el que dicho microorganismo es capaz de crecer en manosa y opcionalmente también en glucosa como fuente de carbono, y/o

b) xilosa isomerasa (EC 5.3.1.5), xiluloquinasa (E 2.7.1.17) y un transportador ABC de alta afinidad o xilosa isomerasa (EC 5.3.1.5), xiluloquinasa (E 2.7.1.17) y cotransportador unidireccional E de protones de xilosa, en el que dicho microorganismo es capaz de crecer en xilosa o xilosa y arabinosa y, opcionalmente, también en glucosa y/o galactosa como fuente de carbono.

Artículo 2. El microorganismo de acuerdo con el artículo 1, en el que el microorganismo se modifica genéticamente para expresar transaldolasa (EC 2.2.1.2) y/o transcetolasa (EC 2.2.1.1).

Artículo 3. El microorganismo de acuerdo con el artículo 1 o 2, en el que el microorganismo se modifica genéticamente para expresar glucoquinasa (EC 2.7.1.2).

Artículo 4. El microorganismo de acuerdo con uno cualquiera de los artículos 1 a 3, en el que el microorganismo es capaz de utilizar manosa y opcionalmente glucosa y/o xilosa y arabinosa y opcionalmente glucosa y galactosa, o manosa, xilosa, arabinosa, glucosa y galactosa como fuente de carbono.

Artículo 5. El microorganismo de acuerdo con uno cualquiera de los artículos 1 a 4, en el que el microorganismo pertenece a la especie *R. eutropha* (*C. necator*).

Artículo 6. Un vector que comprende una secuencia de nucleótidos que codifica la fosfomanosa isomerasa (EC 5.3.1.8) y una secuencia de nucleótidos que codifica la proteína de difusión facilitada para la absorción de manosa (EC1.3.1.74) o una secuencia de nucleótidos que codifica la fosfomanosa isomerasa (EC 5.3.1.8) y una secuencia de nucleótidos que codifica el cotransportador unidireccional E de protones de xilosa y, opcionalmente, una secuencia de nucleótidos que codifica manofructoquinasa (EC 2.7.1.4), y elementos reguladores para regular la expresión de dichos genes en *Cupriavidus* o *Ralstonia*.

Artículo 7. Un vector que comprende una secuencia de nucleótidos que codifica la xilosa isomerasa (EC 5.3.1.5), una secuencia de nucleótidos que codifica la xiluloquinasa (EC2.7.1.17) y una secuencia de nucleótidos que codifica un transportador ABC de alta afinidad o una secuencia de nucleótidos que codifica la xilosa isomerasa (EC 5.3.1.5), una secuencia de nucleótidos que codifica la xiluloquinasa (EC2.7.1.17) y una secuencia de nucleótidos que codifica el cotransportador unidireccional E de protones de xilosa, y elementos reguladores para regular la

expresión de dichos genes en *Cupriavidus* o *Ralstonia*.

Artículo 8. Un vector de acuerdo con el artículo 6 o 7 que comprende una secuencia de nucleótidos que codifica transaldolasa (EC 2.2.1.2) y/o transcetolasa (EC 2.2.1.1).

Artículo 9. Un vector de acuerdo con uno cualquiera de los artículos 6, 7 u 8 que comprende un gen que codifica glucoquinasa (EC2.7.1.2).

Artículo 10. El vector de acuerdo con uno cualquiera de los artículos 6 a 9, en el que el elemento regulador comprende un promotor constitutivo fuerte.

Artículo 11. El vector de acuerdo con uno cualquiera de los artículos 6 a 10, en el que el vector es un vector de amplia gama de hospedadores.

Artículo 12. El vector de acuerdo con uno cualquiera de los artículos 6 a 11, en el que el promotor es el promotor de neokanamicina, el promotor lac, el promotor de gliceraldehído-3-fosfato deshidrogenasa, P2, el promotor hidrogenasa (SH) o el PL.

Artículo 13. Un sistema vectorial que comprende el vector de los artículos 6 y 7 y opcionalmente 8 y/o 9.

Artículo 14. Un método para cultivar el microorganismo de acuerdo con uno cualquiera de los artículos 1 a 5, o un microorganismo *Cupriavidus* o *Ralstonia* que comprende el vector de acuerdo con uno cualquiera de los artículos 6 a 13, que comprende que el medio de cultivo comprende manosa, glucosa, xilosa, arabinosa o galactosa o una combinación de las mismas como fuente de carbono.

Artículo 15. El método de acuerdo con el artículo 14, que comprende que el microorganismo puede sintetizar, opcionalmente después de la modificación genética, un producto químico, como un lípido, y el producto químico se recupera del medio de cultivo o de las células.

Artículo 16. El método de acuerdo con el artículo 14 o 15, que comprende que el microorganismo puede sintetizar polihidroxicanoatos, tales como PHB o ácidos grasos y/o derivados de ácidos grasos.

Artículo 17 Uso del microorganismo de acuerdo con uno cualquiera de los artículos 1 a 5, posible después de una modificación genética, para producir un producto químico deseado

Ejemplos

Ejemplo 1

Materiales y métodos

Cepas bacterianas, plásmidos, oligonucleótidos y condiciones de cultivo. Todas las bacterias, plásmidos y cebadores utilizados en esta aplicación se enumeran en la Tabla 1. Se cultivaron células de la cepa H16 de *R. eutropha* en medio de sales minerales (MSM) como lo describen Schlegel et al. (1961). Se añadieron fuentes de carbono al MSM líquido como se indica en el texto. Se incubaron los cultivos líquidos en matraces Erlenmeyer en un agitador rotatorio horizontal a una agitación de 110 rpm. Se prepararon medios sólidos mediante la adición de agar-agar al 1,5 % (p/v). Se cultivaron células de *Escherichia coli* a 37 °C en caldo de lisogenia (LB, de sus siglas en inglés) [Sambrook et al.1989]. Se cultivaron células de *Zymomonas mobilis* en un medio que contenía Bacto peptona 10 g/l, extracto de levadura 10 g/l y glucosa 20 g/l. Se aplicaron antibióticos de acuerdo con Sambrook et al. (1989) y como se indica en el texto.

Tabla 1. Cepas bacterianas, plásmidos y oligonucleótidos

Cepa, plásmido o cebador	Características relevantes	Fuente o referencia
Cepas		
K-12 de <i>E. coli</i>	Tipo silvestre	DSM 426
TOP 10 de <i>E. coli</i>	F ⁻ <i>mcrA</i> , Δ (<i>mrr</i> - <i>hsdRMS-mcrBC</i>) <i>f80lacZ</i> Δ M15, Δ <i>lacX74</i> , <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , Δ (<i>ara-leu</i>)7697, <i>galU</i> , <i>galK</i> , <i>galT</i> , <i>rpsL</i> , <i>endA1</i> , <i>nupG</i>	Invitrogen
S17-1 de <i>E. coli</i>	<i>recA1</i> , <i>thi1</i> , <i>hsdR17</i> (<i>r_K</i> , <i>m_K</i> ⁺), <i>proA</i> , genes tra del plásmido RP4 cromosómicamente integrados (cepa de movilización)	Simon et al. 1983
JE5511 de <i>E. coli</i>	Hfr, <i>manA4</i> , <i>lpp-1</i> , <i>pps-6</i>	<i>E. coli</i> Genetic Stock Center (Hirota et al. 1977)
cepa H16 de <i>R. eutrophia</i>	Tipo silvestre	DSM428
<i>Z. mobilis</i> subsp. <i>mobilis</i>	Tipo silvestre	DSM424 ATCC 10988
<i>Z. mobilis</i> subsp. <i>mobilis</i>	Tipo silvestre	DSM3580 ATCC 9191
Plásmidos		
pJET1.2/blunt	Ap ^r	Fermentas
pCR2.1	Ap ^r , Km ^r , <i>LacZ</i> - α	Invitrogen
pJET1.2:: <i>glf</i>	con <i>glf</i> como producto de PCR de extremo romo	en los ejemplos
pJET1.2:: <i>mak</i>	con <i>mak</i> como producto de PCR de extremo romo	en los ejemplos
pJET1.2:: <i>pml</i>	con <i>pml</i> como producto de PCR de extremo romo	en los ejemplos

pCR2.1::P	con promotor de neokanamicina como fragmento <i>XhoI/NdeI</i>	en los ejemplos
pCR2.1::P _{nik::gfk}	con promotor de neokanamicina y <i>gfk</i> como fragmento <i>XhoI/EcoRI</i>	en los ejemplos
pBBR1MCS-3	Tc ^r , LacZ- α , <i>mob</i> , <i>rep</i>	Kovach et al. 1995
pBBR1MCS-2::P _{nik::gfk}	con promotor de neokanamicina y <i>gfk</i> como fragmento <i>XhoI/SacI</i>	en los ejemplos
pBBR1MCS-2::P _{nik::gfk::gfk}	con promotor de neokanamicina, <i>gfk</i> y <i>gfk</i> como fragmento <i>XhoI/SacI</i>	en los ejemplos
pBBR1MCS-3::gfk	con <i>gfk</i> como fragmento <i>KpnI/PstI</i>	en los ejemplos
pBBR1MCS-3::gfk::mak	con <i>gfk</i> y <i>mak</i> como fragmento <i>KpnI/XbaI</i>	en los ejemplos
pBBR1MCS-3::gfk::pmi	con <i>gfk</i> como fragmento <i>KpnI/PstI</i> y fragmento <i>pmi</i> <i>XbaI/SacI</i>	en los ejemplos
pBBR1MCS-3::gfk::mak::pmi	con <i>gfk</i> , <i>mak</i> y <i>pmi</i> como fragmento <i>KpnI/SacI</i>	en los ejemplos
Oligonucleótidos		
<i>Pgfk_KpnI_f</i>	5' <u>G</u> GTACCAAGGAAGGACTGATCATG AGTTCTGAAAGTAGTCAGGGTCTA GTC3' (SEQ ID NO:61)	
<i>Pgfk_PstI_r</i>	5' <u>CTG</u> CAGCTACTTCTGGGAGCGCCACATC3' (SEQ ID NO: 62)	
<i>Pmak_PstI_f</i>	5' <u>CTG</u> CAGAAGGAAGGACTGATC GTGCGTATAGGTATCGATTAGGCG3' (SEQ ID NO: 63)	
<i>Pmak_XbaI_r</i>	5' <u>TCT</u> AGATTACTTCTGTGGCCATAACCCAGC3' (SEQ ID NO: 64)	
<i>Ppmi_XbaI_f</i>	5' <u>TCT</u> AGAAAGGAAGGTGACTC ATGCAAAAACCTCAATTAACCTCAGTG CAAAAC3' (SEQ ID NO:65)	
<i>Ppmi_SacI_r</i>	5' <u>GAG</u> CTCTTACAGCTTGTGTAAACACGCGCTAAAC3' (SEQ ID NO: 66)	
<i>gfk_RT_f</i>	5' <u>GTT</u> CTATCGATTGGGTTAATGCCAGTGG3' (SEQ ID NO: 67)	
<i>gfk_RT_r</i>	5' <u>GGA</u> ACATCTCGGTGCATAATACAGC3' (SEQ ID NO: 68)	

<i>mak_RT_f</i>	5'GCGAGGTTGCAGCGGGAAGTG3' (SEQ ID NO: 69)	
<i>mak_RT_r</i>	5'AATTTCACTGCGCTTTTCAGCGCATGTCC3' (SEQ ID NO: 70)	
<i>pmi_RT_f</i>	5'CGCTGAGCGCTTTCTTGGCGAT3' (SEQ ID NO: 71)	
<i>pmi_RT_r</i>	5'GCGGTGTTTCAGCGAACAGGAACA3' (SEQ ID NO: 72)	
cebador de secuenciación directo pJET1.2, 23 meros	5'CGACTCACTATAGGGAGAGCGGC3' (SEQ ID NO: 73)	Fermentas
cebador de secuenciación inverso pJET1.2, 24 meros	5'AAGAACATCGATTTCCATGGCAG3' (SEQ ID NO: 74)	Fermentas
M13/pUC-directo, 17 meros	5'GTAAAAACGACGGCCAGT3' (SEQ ID NO: 75)	Jena Bioscience
<i>F_Pnk_XhoI</i>	5'CTCGAGCGGGAATTGCCAGCTGGGG3' (SEQ ID NO: 76)	
<i>R_Pnk_NdeI</i>	5'CATAATGAAACGATCCTCATCCTCTCTTG3' (SEQ ID NO: 77)	
<i>P_glk_NdeI</i>	5'CATATGACAAAGTATGCAATTAGTCGGTGATGTGGG3' (SEQ ID NO: 78)	
<i>R_glk_SalI</i>	5'GTCGACTTACAGAAATGTGACCTAAGGTCTGGCGTAAATGTGC3' (SEQ ID NO: 79)	
<i>P_glf_BamHI</i>	5'GGATCCATGAGGATCGTTTCGCATGAGTTCTGAAAGTAGTCAGGGTCTAGTC3' (SEQ ID NO: 80)	
<i>R_glf_SacI</i>	5'GAGCTCCTACTTCTGGGAGCGCCACATCTCCTCG3' (SEQ ID NO: 81)	

Tabla 2. Crecimiento de la cepa H16 recombinante de *R. eutropha* y del tipo silvestre (control negativo) en diferentes fuentes de carbono. Se incubaron células en placas de agar MSM que contenían un 1 % (p/v) de la fuente de carbono respectiva durante 5 días a 30 °C. Se examinó fenotípicamente el crecimiento de colonias visibles.

Caracteres: (++) crecimiento después de un tiempo de incubación de 24-72 h, (+) crecimiento después de más de 72 h de incubación, (-) no se detecta crecimiento visible.

Fuente de carbono	cepa H16 de <i>R. eutropha</i> de tipo silvestre	<i>R. eutropha</i> que alberga pBBR1MCS-3::glf::mak::pmi	<i>R. eutropha</i> que alberga pBBR1MCS-2::P _{nk} ::glk::glf
L-Arabinosa	-	-	-
D-fructosa	++	++	++
D-galactosa	-	-	-
D-gluconato	++	++	++
D-glucosa	-	+	++
D-glicerol	+	+	+
D-maltosa	-	-	-
D-manitol	-	-	-
D-manosa	-	+	-
N-acetil-D-glucosamina	++	++	++
piruvato	++	++	++
D-ribosa	-	-	-
D-trehalosa	-	-	-
D-xilosa	-	-	-

Aislamiento, análisis y modificación de ADN. Se preparó ADN plasmídico a partir de lisados crudos por el método de extracción alcalina (Birnboim y Doly 1979). Se preparó el ADN total de DSM3580 o DSM 424 de *Z. mobilis* y de la cepa K-12 de *E. coli* utilizando el kit de tejido y sangre Qiagen DNeasy (Qiagen, Hilden, Alemania) de acuerdo con el protocolo del fabricante. Se restringió el ADN con endonucleasas de restricción (Gibco/BRL, Gaithersburg, EE. UU.) bajo las condiciones recomendadas por el fabricante. Todas las demás manipulaciones y procedimientos genéticos se realizaron según lo descrito por Sambrook et al. (1989).

Construcciones de plásmidos y su transferencia a la cepa H16 de *R. eutropha*. Las regiones codificantes de *glf* de DSM424 y DSM3580 de *Z. mobilis mobilis*, y de *glk*, *mak* y *pmi* de *E. coli*, respectivamente, se amplificaron mediante PCR utilizando oligonucleótidos *Pglf_KpnI_f*, *Pglf_PstI_r*, *F_glf_BamHI* y *R_glf_SacI* para la amplificación de *glf*, *F_Pnk_XhoI* y *R_Pnk_NdeI* para *P_{nk}*, *F_glk_NdeI* y *R_glk_Sall* para *glk*, *Pmak_PstI_f* y *Pmak_XbaI_r* para la amplificación de *mak*, o *Ppmi_XbaI_f* y *Ppmi_SacI_r* para la amplificación de *pmi*, respectivamente (tabla 1). Para la PCR, se utilizó ADN polimerasa de inicio en caliente KOD (Merck, Darmstadt, Alemania) de acuerdo con las instrucciones del fabricante. Luego se clonaron los productos de la PCR en el vector de clonación pJET1.2/blunt (Fermentas, Alemania) utilizando la ADN ligasa T4 (Gibco BRL, Gaithersburg, EE. UU.) o en el vector de clonación pCR2.1 (Invitrogen, Karlsruhe, Alemania) y se transfirieron a la cepa TOP10 de *E. Coli*. Se aislaron los plásmidos de clones resistentes a la ampicilina y se cortaron los fragmentos clonados mediante restricción con las respectivas enzimas de restricción adecuadas para la clonación adicional, se extrajeron del gel después de la separación utilizando

el kit de extracción de gel E.Z.N.A (Omega Bio-tec, Bangalore, India). Para los experimentos de expresión en la cepa H16 de *R. eutropha*, se utilizaron los vectores de amplia gama de hospedadores pBBR1MCS-2 y pBBR1MCS-3 para la clonación de *glf*, *glk*, *mak* y *pmi*. El vector pBBR1MCS-2 confirió resistencia a la kanamicina, mientras que pBBR1MCS-3 confirió resistencia a la tetraciclina para la selección de la cepa H16 de *E. coli* y *R. eutropha*. Se cortó la región codificante de *glf* mediante restricción con *KpnI* y *PstI* y se ligó al plásmido pBBR1MCS-3 linealizado con *KpnI/PstI*, que produce los plásmidos pBBR1MCS-3::*glf* (Fig. 2). La secuencia codificante del promotor de neokanamicina *P_{nk}* se amplificó utilizando los cebadores *F_Pnk_XhoI* y *R_Pnk_NdeI* que utilizan el vector aislado pBBR1MCS-2 y se ligó adicionalmente en el vector pCR2.1 (Invitrogen, Carlsbad, EE. UU.) de acuerdo con las instrucciones del fabricante, que proporcionaron pCR2.1::*P_{nk}*. Se amplificó el gen *glk* mediante PCR adaptada a partir del ADN total de K12 de *E. coli* utilizando los oligonucleótidos *F_glk_NdeI* y *R_glk_SalI* como cebadores. El producto de PCR resultante se clonó como fragmento *NdeI/SalI* en pCR2.1::*P_{nk}*, produciendo pCR2.1::*P_{nk}::glk*. Se amplificó el gen *glf* mediante PCR adaptada a partir del ADN total de DSM3580 de *Z. mobilis* utilizando los oligonucleótidos *F_glf_BamHI* y *R_glf_SacI* como cebadores. El producto de PCR resultante se clonó como fragmento *BamHI/SalI* en pBBR1MCS-2, que produce pBBR1MCS-2::*glf*. Para obtener el plásmido pBBR1MCS-2::*P_{nk}::glk::glf*, la región codificante de *P_{nk}::glk* se cortó mediante restricción con *XhoI* y *EcoRI* del plásmido pCR2.1::*P_{nk}::glk* y se ligó en *XhoI/EcoRI* linealizado con pBBR1MCS-2::*glf*. La secuencia codificante de *glk* se escindió adicionalmente mediante restricción con *NdeI/SalI* del plásmido pBBR1MCS-2::*P_{nk}::glk::glf*, el plásmido linealizado se embotó utilizando polimerasa T4 (Gibco BRL, Gaithersburg, EE. UU.) y se volvió a utilizando ADN ligasa T4 (Gibco BRL, Gaithersburg, EE. UU.) que produce el plásmido pBBR1MCS-2::*P_{nk}::glf*. Para obtener los plásmidos pBBR1MCS-3::*glf::mak* y pBBR1MCS-3::*glf::pmi*, las regiones codificantes de *mak* y *pmi* se eliminaron mediante restricción con *PstI* y *XbaI* o *XbaI* y *SacI*, respectivamente, y se ligaron al plásmido pBBR1MCS-3::*glf* linealizado con *PstI/XbaI* y *XbaI/SacI*, respectivamente (Fig. 2). Las regiones codificantes escindidas de *mak* y *pmi* también se ligaron juntas al plásmido pBBR1MCS-3::*glf* linealizado con *PstI/SacI*, produciendo pBBR1MCS-3::*glf::mak::pmi* (Fig. 2). Todos los plásmidos se transfirieron a la cepa S17-1 de *E. coli* mediante transformación (Hanahan 1983).

Transferencia de ADN mediante conjugación. La transferencia de los plásmidos pBBR1MCS-2, pBBR1MCS-2::*P_{nk}::glf*, pBBR1MCS-2::*P_{nk}::glk::glf*, pBBR1MCS-3, pBBR1MCS-3::*glf*, pBBR1MCS-3::*glf::mak*, pBBR1MCS-3::*glf::pmi* y pBBR1MCS-3::*glf::mak::pmi* (Tabla 1) se realizó mediante conjugación aplicando un protocolo descrito previamente (Friedrich et al. 1981), que utiliza S17-1 de *E. coli* como donante y la cepa H16 de *R. eutropha* como receptor.

Análisis de secuencia de ADN. El ADN se secuenció en un sistema de electroforesis del analizador de ADN 3730 de 48 capilares (Applied Biosystems, Foster City, CA). Para la secuenciación de insertos en el vector de clonación pJET1.2/Blunt (Fermentas, Alemania), se utilizaron el cebador de secuenciación directo pJET1.2, de 23 meros y el cebador de secuenciación inversa pJET1.2, de 24 meros. El cebador directo M13/pUC y el cebador inverso M13/pUC universales se utilizaron para secuenciar los insertos en los plásmidos pBBR1MCS-2 y pBBR1MCS-3 (Kovach et al. 1995).

Análisis mediante RT-PCR de ARN total aislado de la cepa H16 de *R. eutropha*. El ARN total libre de ADN de la cepa H16 de *R. eutropha* se preparó utilizando el kit de purificación de ARN RNeasy (Qiagen, Hilden, Alemania) de acuerdo con el protocolo del fabricante. Para la identificación del ARNm procedente de *glf*, *mak* y *pmi*, se aplicó RT-PCR utilizando los oligonucleótidos *glf_RT_f*, *glf_RT_r*, *mak_RTJ*, *mak_RT_r*, *pmi_RT_f* y *pmi_RT_r* (Tabla 1). La RT-PCR se llevó a cabo utilizando un kit de RT-PCR de una fase de Qiagen (Qiagen, Hilden, Alemania) de acuerdo con el protocolo del fabricante y 0,5 ng de ARN como plantilla. Para excluir cualquier contaminación de ADN que pueda servir como plantilla para la PCR, se añadió ARN de plantilla en un experimento de control, después de la inactivación de la transcriptasa inversa durante 15 minutos a 95 °C en presencia de Tag-polimerasa. La ausencia de productos de PCR indicó que los productos de RT-PCR no procedieron de ADN contaminante.

Preparación de fracciones de proteínas solubles de la cepa H16 de *R. eutropha*. Las células de cepas recombinantes de la cepa H16 de *R. eutropha* se cultivaron como se describió anteriormente en presencia de gluconato de sodio al 1 % (p/v) como única fuente de carbono. Se recogieron las células mediante centrifugación durante 15 minutos a 4 °C y 3.500 x g, se lavaron con tampón MOPS 100 mM (pH 7,0) y se resuspendieron en dos volúmenes del mismo tampón. La interrupción se realizó mediante la exposición a ultrasonidos utilizando un Sonifer 250 (Branson Sonic Power Company) con una amplitud de 16 µm (1 min/ml; control de salida del 50 %). Durante la exposición a ultrasonidos, se enfriaron las muestras con una mezcla de hielo-NaCl. Se prepararon fracciones de proteínas libres de membrana solubles mediante ultracentrifugación de 60 min de los extractos brutos a 100.000 xg y 4 °C.

Solubilización de proteínas de membrana. Se cultivaron células recombinantes de H16 de *R. eutropha* como se describe anteriormente en presencia de glucosa al 1 % (p/v) como única fuente de carbono. Después de 48 h, se recogieron las células mediante centrifugación de 15 min a 3.500 x g y 4 °C, se lavaron una vez con NaCl al 0,9 % (p/v) y se resuspendieron en un volumen de tampón Tris/HCl 100 mM (pH 7,0). Se llevó a cabo la interrupción celular mediante un triple paso a través de una celda de presión francesa preenfriada a 1.000 MPa. Se centrifugaron los lisados obtenidos como antes para eliminar las células residuales. Se preparó la fracción de membrana mediante centrifugación de 1 h del sobrenadante a 100.000 x g y 4 °C, se resuspendió en 2 ml de tampón Tris/HCl 20 mM (pH 7,4) que contenía NaCl 200 mM y TritonX-114 al 2 % (v/v) y se incubó durante 12 h en hielo en un agitador rotatorio (W3, VWR international GmbH, Darmstadt, Alemania). Se centrifugó el lisado durante 20 minutos a 3.500 xg y 4 °C.

Se incubó el sobrenadante durante 5 minutos a 37 °C y se centrifugó durante 20 minutos a 3.500 xg y 37 °C para eliminar TritonX-114 de acuerdo con Bordier (Bordier, 1981). Se diluyó adicionalmente el sobrenadante con 2 volúmenes del mismo tampón sin TritonX-114 para disminuir la concentración de detergente.

- 5 **PAGE unidimensional.** Se resuspendieron las muestras de proteína en tampón de carga de gel (SDS al 0,6 % (p/v); β-mercaptoetanol al 1,25 % (v/v); EDTA 0,25 mM; glicerol al 10 % (v/v); azul de bromofenol al 0,001 % (p/v); Tris/HCl 12,5 mM, pH 6,8) y se separaron en geles de SDS-poliacrilamida al 12,5% (p/v) como describe Laemmli (1970). Se tiñeron las proteínas con azul brillante de Coomassie R-250 (Weber y Osborn, 1969). Las muestras de extractos crudos y proteínas de membrana solubilizadas fueron analizadas por este método.

- 10 **Ensayos de actividad de MAK, PMI y GLK.** Las fracciones de proteínas solubles de la cepa H16 recombinante de *R. eutropha* se aplicaron para las mediciones de actividad de MAK, PMI y GLK a 30 °C utilizando un espectrofotómetro Nicolet Evolution 100 UV/VIS (Thermo Electron Corporation, Cambridge, UK). La actividad de MAK se analizó con un sistema acoplado a NADH como lo describen Sebastian y Asensio (1972) con algunas modificaciones, utilizando
15 piruvato quinasa y lactato deshidrogenasa como enzimas auxiliares. La mezcla de reacción tamponada (MOPS 100 mM, pH 7,0) contenía MgCl₂ 2 mM, ATP 2 mM, 0,2 NADH, 5 mM de manosa, PEP 0,2 mM, piruvato quinasa 5 U ml⁻¹, lactato deshidrogenasa 5 U ml⁻¹ y de 5 a 100 µl de extracto soluble. La actividad de PMI se midió con un ensayo acoplado a NADP (Kang y Markovitz 1967) con algunas modificaciones, utilizando fosfoglucoisomerasa y glucosa-6-fosfato deshidrogenasa como enzimas auxiliares. La mezcla de reacción tamponada (MOPS 100 mM, pH 7,0)
20 contenía MgCl₂ 5 mM, NADP 1 mM, manosa-6-fosfato 3 mM, fosfoglucoisomerasa 1 U ml⁻¹, glucosa-6-fosfato deshidrogenasa 1 U ml⁻¹ y de 10 a 100 µl de extracto soluble.

- La actividad de GLK se midió con un sistema acoplado a NADPH de acuerdo con el método de Gottschalk (Gottschalk et al. 1964) con modificaciones relacionadas con el tampón. La mezcla de tampón (Tris/HCl 100 mM, pH 7,6) contenía
25 MgCl₂ 7 mM, NADP 0,9 mM, D-glucosa 460 mM, ATP 0,7 mM, glucosa-6-fosfato deshidrogenasa 0,35 U ml⁻¹ y de 10 a 50 µl de preparación enzimática.

- Análisis del contenido de PHB de células recombinantes de *R. eutropha* por GC.** El material celular liofilizado se sometió a metanolisis en presencia de metanol y ácido sulfúrico (MeOH al 85 %, v/v y H₂SO₄ al 15 %, v/v) durante 4
30 horas a 100 °C, y los ésteres metílicos resultantes de los constituyentes de PHA fueron caracterizados mediante cromatografía de gases utilizando un Agilent 6850 GC (Agilent Technologies, Waldbronn, Alemania) equipado con una columna capilar BP21 (50 m por 0,22 mm; espesor de la película, 250 nm; SGE, Darmstadt, Alemania) y un detector de ionización de llama (Agilent Technologies). Se analizó una porción de 2 µl de la fase orgánica después de la inyección dividida (relación de división, 1:5), y se utilizó un flujo constante de hidrógeno de 0,6 ml min⁻¹ como gas
35 portador. Las temperaturas del inyector fueron de 250 °C y 220 °C, respectivamente. Se aplicó el siguiente programa de temperatura: 120 °C durante 5 min, aumento de 3 °C min⁻¹ hasta 180 °C y aumento de 10 °C min⁻¹ hasta 220 °C y 220 °C durante 31 min. Se identificaron las sustancias mediante la comparación de sus tiempos de retención con los del éster metílico de ácido graso estándar.

- 40 El gen *glf* de *Z. mobilis* más los genes *mak* y *pmi* de *E. coli* se introdujeron episomalmente y se expresaron en *R. eutropha* y se caracterizaron las cepas recombinantes. Las rutas de ingeniería del catabolismo de la glucosa y la manosa en la cepa recombinante de *R. eutropha* se muestran en la Fig. 1.

- Construcción de diferentes vectores de expresión pBBR1 para la cepa H16 de *R. eutropha*.** Las regiones codificantes del gen *glf* de *Z. mobilis* (1.449 pb) y los genes *mak* (936 pb), *pmi* (1.203 pb) y *glk* (966 pb) de *E. coli* se amplificaron mediante PCR a partir de ADN genómico de DSM 424 de *Z. mobilis* y la cepa K-12 de *E. coli*, respectivamente. Los cuatro productos de la PCR comprendidos junto a las regiones codificantes también son sitios de unión al ribosoma adecuados para *R. eutropha* cadena arriba del codón de inicio respectivo para permitir la expresión en este hospedador. Estos fragmentos se clonaron a través del vector pJET1,2/blunt o pCR2.1 en los
50 vectores de amplia gama de hospedadores pBBR1MCS-2 y pBBR1MCS-3 bajo el control del promotor de *neokanamycin* o *lac*, respectivamente, que permiten la expresión constitutiva de genes clonados en H16 de *R. eutropha*, que producen los plásmidos pBBR1MCS-2::P_{nk}::*glf*, pBBR1MCS-2::P_{nk}::*glk*::*glf*, pBBR1MCS-3::*glf*, pBBR1MCS-3::*glf*::*mak*, pBBR1MCS-3::*glf*::*pmi* y pBBR1MCS-3::*glf*::*mak*::*pmi* (Fig. 2). La construcción precisa de estos plásmidos se confirmó mediante secuenciación. Adicionalmente, la funcionalidad del gen *pmi* en pBBR1MCS-3::*glf*::*pmi* y pBBR1MCS-3::*glf*::*mak*::*pmi* se confirmó mediante la complementación del mutante JE5511 de *E. coli* deficiente en *pmi* (*manA*) (Hirota et al. 1977). Los mutantes *manA* de *E. coli* recombinantes mostraron un buen crecimiento en medio M9 sólido que contenía manosa al 0,5 % (p/v) como única fuente de carbono después de dos días de incubación a 37 °C. En contraste, el vector pBBR1MCS-3 no confirió crecimiento al mutante *manA*.

- 60 **Transferencia de plásmidos a la cepa H16 de *R. eutropha* y establecimiento de la utilización de glucosa y manosa.** Ambos plásmidos pBBR1MCS-2::P_{nk}::*glf* y pBBR1MCS-2::P_{nk}::*glk*::*glf* se transfirieron a través de la conjugación de S17-1 de *E. coli* a la cepa H16 *R. eutropha*. Los transconjugantes se seleccionaron en placas de agar MSM que contenían gluconato de sodio al 1 % (p/v) y kanamicina (300 µg/ml), y se transfirieron de allí a placas MSM que contenían glucosa al 1 % (p/v) como única fuente de carbono. Solo las cepas recombinantes que albergan
65 pBBR1MCS-2::P_{nk}::*glf* o pBBR1MCS-2::P_{nk}::*glk*::*glf* mostraron un crecimiento significativo después de 5 días de incubación.

Para investigar si los tres genes son necesarios para la utilización de manosa, todos los plásmidos contruidos y el plásmido pBBR1MCS-3 se movilizaron de la cepa S17-1 de *E. coli* a la cepa H16 de *R. eutropha*. Los transconjugantes se seleccionaron en placas de agar MSM que contenían gluconato de sodio al 1 % (p/v) como única fuente de carbono y tetraciclina (12,5 µg/ml). Luego se transfirieron a placas de agar MSM que contenían manosa como única fuente de carbono a diversas concentraciones (0,2, 0,5, 1,0, 2,0, 4,0 % [p/v]). Solo las cepas recombinantes que albergan pBBR1MCS-3::glf::pmi o pBBR1MCS-3::glf::mak::pmi mostraron un crecimiento significativo en presencia de cualquier concentración de manosa probada. Mientras que las células que albergan el plásmido pBBR1MCS-3::glf::mak::pmi mostraron un crecimiento distinto en manosa después de 3-4 días de incubación, las células con pBBR1MCS-3::glf::pmi solo mostraron un crecimiento después de un tiempo de incubación de 5-6 días. Por el contrario, las cepas recombinantes que albergan el plásmido pBBR1MCS-3::glf o pBBR1MCS-3::glf::mak, respectivamente, apenas mostraron crecimiento incluso después de un período de incubación de 30 días. Solo después de la exposición previa de las últimas cepas recombinantes en placas MSM que contenían fructosa en lugar de gluconato, se produjo un ligero crecimiento después de un tiempo prolongado de incubación.

Control de expresión de glf, glk, mak y pmi. Para confirmar la presencia de ARNm procedente de *glf*, *mak* y *pmi*, se realizaron análisis transcripcionales de las cepas recombinantes. Para ello, las células de las cepas recombinantes de *R. eutropha* que albergan pBBR1MCS-3::glf::mak::pmi o pBBR1MCS-3 se cultivaron en MSM que contenía gluconato de sodio al 1 % (p/v) como única fuente de carbono. Dado que P_{lac} es un promotor constitutivo en la cepa H16 de *R. eutropha* (Siedow et al. 1999), los tres genes deben expresarse de forma constitutiva sin inducción adicional. Los análisis transcripcionales se realizaron mediante RT-PCR de una fase. El ARN total de las células de la fase de crecimiento logarítmico tardío sirvió para RT-PCR utilizando los cebadores *glf_RT_f*, *glf_RT_r*, *mak_RT_f*, *mak_RT_r*, *pmi_RT_f* y *pmi_RT_r* específicos para *glf*, *mak* y *pmi* (Tabla 1). Se obtuvieron productos de PCR del tamaño esperado de 351 pb para *glf*, *mak* y *pmi*, respectivamente, si se analizó el ARN aislado de pBBR1MCS-3::glf::mak::pmi recombinante que alberga células de la cepa H16. Esto demostró claramente que *glf*, *mak* y *pmi* solo se transcribieron en células de la cepa recombinante de la cepa H16 de *R. eutropha* que alberga pBBR1MCS-3::glf::mak::pmi pero no en el control negativo que alberga solo el vector pBBR1MCS-3. La ausencia de productos de PCR en el control sin ninguna actividad de la transcriptasa inversa indicó que el producto de RT-PCR no procedió de ADN contaminante (Fig. 3).

Para confirmar aún más la transcripción de *glk* y *glf* en cepas recombinantes que albergan el plásmido pBBR1MCS-2::P_{nk}::glk::glf a nivel de proteína, se realizó un análisis mediante SDS-PAGE. Como Glf es una proteína transmembrana (Parker et al. 1995), las proteínas de membrana de la cepa H16 de *R. eutropha* que alberga el vector pBBR1MCS-2 y la cepa H16 de *R. eutropha* que albergan el plásmido pBBR1MCS-2::P_{nk}::glk::glf se solubilizaron y se separaron por PAGE. El electroferograma mostró una banda de proteína distinta con una masa aparente de 50 kDa que representa a Glf, que estaba ausente en el control. De manera similar, el análisis de la fracción de células solubles mostró una banda de proteína distinta a 35 kDa, que representa a Glk (Fig. 4).

La presencia de MAK, PMI y GLK activas funcionales en las cepas recombinantes se confirmó mediante análisis enzimáticos. Ensayos enzimáticos acoplados que emplean las fracciones de proteínas solubles obtenidas de células de la cepa recombinante de H16 que albergan pBBR1MCS-3::glf::mak::pmi, o pBBR1MCS-2::P_{nk}::glk::glf cultivadas en MSM que contienen gluconato de sodio al 1 % (p/v) como única fuente de carbono, demostraron la presencia de PMI activo (0,25 U mg⁻¹), MAK (0,14 U mg⁻¹) y también GLK (0,52 U mg⁻¹), respectivamente, mientras que estas actividades, excepto GLK (0,02 U mg⁻¹), estuvieron ausentes en las fracciones de proteína soluble del control negativo de la cepa H16 que albergan los plásmidos pBBR1MCS-2 o pBBR1MCS-3, respectivamente, lo que indica la ausencia de PMI y MAK en la cepa de control negativo.

Investigaciones sobre la utilización de varias fuentes de carbono por las cepas recombinantes de H16 de *R. eutropha*. En comparación con las otras cepas recombinantes, la cepa H16 de *R. eutropha* que alberga los plásmidos pBBR1MCS-2::P_{nk}::glk::glf o pBBR1MCS-3::glf::mak::pmi exhibió el crecimiento más rápido en glucosa o manosa de todos los sustratos probados, respectivamente, como única fuente de carbono. Por lo tanto, estas cepas se investigaron más con respecto a la utilización de varias otras fuentes de carbono que emplean placas de agar MSM que contienen cada una un 1 % (p/v) de L-arabinosa, D-fructosa, D-galactosa, D-gluconato, D-glucosa, glicerina, maltosa, D-manitol, D-manosa, N-acetil D-glucosamina, piruvato, D-ribosa, trehalosa o D-xilosa como única fuente de carbono. El tipo silvestre de *R. eutropha* sirvió como control negativo. Los resultados se resumen en la tabla 2. Tanto la cepa de tipo silvestre como las cepas recombinantes, fueron capaces de crecer en presencia de D-fructosa, D-gluconato, glicerol, N-acetilglucosamina y piruvato (Tabla 2). Todas las demás fuentes de carbono mencionadas anteriormente no se utilizaron para el crecimiento, excepto la manosa por el recombinante que alberga pBBR1MCS-3::glf::mak::pmi y la glucosa por las cepas recombinantes que albergan pBBR1MCS-2::P_{nk}::glk::glf o pBBR1MCS-3::glf::mak::pmi (Tabla 2).

Crecimiento de la *R. eutropha* recombinante con glucosa como única fuente de carbono. Las cepas recombinantes de la cepa H16 de *R. eutropha* se cultivaron en MSM líquido que contenía glucosa al 1 % (p/v). Aunque todas las cepas mostraron crecimiento en glucosa, la cepa recombinante que alberga pBBR1MCS-2::P_{nk}::glk::glf mostró el crecimiento más rápido ($\mu = 0,32 \text{ h}^{-1}$) en comparación con la cepa recombinante que alberga pBBR1MCS-3::glf::mak::pmi ($\mu = 0,12 \text{ h}^{-1}$) (Fig. 5). Esto fue en el mismo orden que la cepa de control positivo G⁺1 ($\mu = 0,30 \text{ h}^{-1}$).

Curiosamente, las células recombinantes que albergan plásmidos que llevan solo el gen *glf* ya sea bajo el control del promotor de *neokanamicina* o *lac* exhibieron un crecimiento fuertemente disminuido ($\mu = 0,085$ y $0,073$, respectivamente) cuando se compararon con células recombinantes que albergan plásmidos con los otros genes (Fig. 5).

Crecimiento de *R. eutropha* recombinante con manosa o gluconato como única fuente de carbono. Las células recombinantes de la cepa H16 de *R. eutropha* que albergan el plásmido pBBR1MCS-3::*glf::mak::pmi* se cultivaron en MSM líquido que contenía diferentes concentraciones de manosa como única fuente de carbono para evaluar la concentración óptima para el crecimiento máximo. Además, se examinó si los genes introducidos afectan el crecimiento de la cepa recombinante en comparación con la cepa H16 de tipo silvestre en MSM que contiene gluconato de sodio al 0,5 % (p/v) como única fuente de carbono. Según las expectativas, la cepa recombinante H16 que alberga pBBR1MCS-3::*glf::mak::pmi* mostró un crecimiento más rápido en gluconato ($\mu = 0,33 \text{ h}^{-1}$) que en manosa ($\mu = 0,078 \text{ h}^{-1}$). El crecimiento máximo en manosa se produjo al 1 % (p/v) de manosa, mientras que las concentraciones más bajas y más altas de este azúcar produjeron un crecimiento ligeramente más lento (Fig. 6). Con el gluconato como única fuente de carbono, el crecimiento de la cepa recombinante solo disminuyó ligeramente en comparación con el tipo silvestre.

Ejemplo 2

Cepas bacterianas, plásmidos, oligonucleótidos y condiciones de cultivo. Todas las bacterias, plásmidos y cebadores utilizados en este estudio se enumeran en la tabla 3 y 4, respectivamente. Se cultivaron células de la cepa H16 de *R. eutropha* en caldo de cultivo o medio de sales minerales (MSM) como lo describen Schlegel et al. (1961). Se añadieron fuentes de carbono al MSM líquido como se indica en el texto. Se incubaron los cultivos líquidos en matraces Erlenmeyer en un agitador rotatorio horizontal a una agitación de 110 rpm. Se prepararon medios sólidos mediante la adición de agar-agar al 1,5 % (p/v). Se cultivaron células de *Escherichia coli* a 37 °C en caldo de lisogenia (LB, [Sambrook et al. 1989]), o medio de sales minerales M9 (Sambrook et al. 1989) con D-xilosa como única fuente de carbono. Se aplicaron antibióticos de acuerdo con Sambrook et al. (1989) y como se indica en el texto.

Aislamiento, análisis y modificación de ADN. Se preparó ADN plasmídico a partir de lisados crudos por el método de extracción alcalina (Birnboim y Doly 1979). Se preparó el ADN total de la cepa K-12 de *E. coli* utilizando el kit de tejido y sangre Qiagen DNeasy (Qiagen, Hilden, Alemania) de acuerdo con el protocolo del fabricante. Se restringió el ADN con endonucleasas de restricción (Gibco/BRL, Gaithersburg, EE. UU.) bajo las condiciones recomendadas por el fabricante. Todos las demás manipulaciones y procedimientos genéticos se realizaron según lo descrito por Sambrook et al. (1989).

Tabla 3. Cepas bacterianas, plásmidos y oligonucleótidos utilizados en este estudio

Cepa, plásmido o cebador	Características relevantes	Fuente o referencia
Cepas		
K-12 de <i>E. coli</i>	Tipo silvestre	DSM 426
TOP10 de <i>E. coli</i>	F ⁻ <i>mcrA</i> , $\Delta(mrr\text{-}hsdRMS\text{-}mcrBC)$ <i>f80lacZ</i> $\Delta M15$, $\Delta lacX74$, <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , $\Delta(ara\text{-}leu)7697$, <i>galU</i> , <i>galK</i> , <i>rpsL</i> , <i>endA1</i> , <i>nupG</i>	Invitrogen
S17-1 de <i>E. coli</i>	<i>recA1</i> , <i>thi1</i> , <i>hsdR17</i> (<i>r_k</i> ⁻ , <i>m_k</i> ⁺), <i>prok</i> , genes <i>tra</i> del plásmido RP4 cromosómicamente integrados (cepa de movilización)	Simon et al. 1983
HB101 de <i>E. coli</i>	<i>hsdS</i> , <i>recA</i> , <i>proA</i> , <i>lacy</i> , <i>9aK</i> , <i>xyl-5</i> , <i>mtl</i>	<i>E. coli</i> Genetic Stock Center (Hirota et al. 1977)
JWL255 de <i>E. coli</i>	<i>xy191pR</i>	<i>E. coli</i> Genetic Stock Center (Hirota et al. 1977)
LR2-168 de <i>E. coli</i>	<i>thi</i> , <i>arg</i> , <i>met</i> , <i>his</i> , <i>gal</i> , <i>xyl-7</i>	<i>E. coli</i> Genetic Stock Center (Hirota et al. 1977)
cepa H16 de <i>R. eutropha</i>	Tipo silvestre	DSM 428

Plásmidos		
pJET1.2/blunt	Ap ^r	Fermentas
pJET1.2::xyIA	con xyIA como producto de PCR de extremo romo	en los ejemplos
pJET1.2::xyIB	con xyIB como producto de PCR de extremo romo	en los ejemplos
pJET1.2::xyIE	con xyIE como producto de PCR de extremo romo	en los ejemplos
pBBR1MCS-3	Tc ^r , LacZ-a, mob, rep	Kovach et al. 1995
pBBR1MCS-3:: xyIFGH	con xyIFGH como fragmento SpeI/SacII	en los ejemplos
pBBR1MCS-3:: xyIAB	con xyIAB como fragmento ApaI/SpeI	en los ejemplos
pBBR1MCS-3:: xyIABE	con xyIABE como fragmento ApaI/SacII	en los ejemplos
pBBR1MCS-3:: xyIAB-FGH	con xyIABE como fragmento ApaI/SacII	en los ejemplos

Tabla 4. Cebadores de secuenciación y PCR utilizados en este estudio

Oligonucleótidos	Secuencia (5'→ 3')	Función
<i>P_xyIA_ApaI_f</i>	ATT <u>AGG GCC CAA</u> GGA GGT TAC AGC ATG CAA GCC TAT TTT GAC CAG (SEQ ID NO: 82)	cebador en el 5' para la amplificación de la región codificante de xyIA. El cebador contiene el sitio de restricción para ApaI (subrayado).
<i>P_xyIA_SmaI_r</i>	ATT GCC CGG <u>GTT ATT TGT</u> CGA ACA GAT AAT G (SEQ ID NO: 83)	cebador en el 3' para la amplificación de la región codificante de xyIA. El cebador contiene el sitio de restricción para SmaI (subrayado).
<i>P_xyIB_SmaI_f</i>	ATT <u>GCC CGG GAA</u> GGA GGT TAC AGC ATG TAT ATC GGG ATA GAT CTT (SEQ ID NO: 84)	cebador en el 5' para la amplificación de la región codificante de xyIB. El cebador contiene el sitio de restricción para SmaI (subrayado).
<i>P_xyIB_SpeI_r</i>	CCC GAC TAG <u>TTT ACG CCA</u> TTA ATG GCA GAA G (SEQ ID NO: 85)	cebador en el 3' para la amplificación de la región codificante de xyIB. El cebador contiene el sitio de restricción para SpeI (subrayado).

P _{xy/E_SpeI_f}	CCC GAC TAG TAA GGA GGT TAC AGC ATG AAT ACC CAG TAT AAT TC (SEQ ID NO: 86)	cebador en el 5' para la amplificación de la región codificante de <i>xy/E</i> . El cebador contiene el sitio de restricción para <i>SpeI</i> (subrayado).
P _{xy/E_SacII_r}	ATT GCC GCG GTT ACA GCG TAG CAG TTT GTT G (SEQ ID NO: 87)	cebador en el 3' para la amplificación de la región codificante de <i>xy/E</i> . El cebador contiene el sitio de restricción para <i>SacII</i> (subrayado).
P _{xy/FGH_SpeI_f}	CCC GAC TAG TAA GGA GGC TAC AGC ATG AAA ATA AAG AAC ATT CTA C (SEQ ID NO: 88)	cebador en el 5' para la amplificación de la región codificante de <i>xy/FGH</i> : El cebador contiene el sitio de restricción para <i>SpeI</i> (subrayado).
P _{xy/FGH_SacII_r}	ATT GCC GCG GTC AAG AAC GGC GTT TGG TTG (SEQ ID NO: 89)	cebador en el 3' para la amplificación de la región codificante de <i>xy/FGH</i> . El cebador contiene el sitio de restricción para <i>SacII</i> (subrayado).
	AAAACCGCGGCGGCTCGGA AGTCG (SEQ ID NO: 90)	cebador en el 5' para la amplificación de la región promotora P _{GAPDH} : El cebador contiene el sitio de restricción para <i>SacII</i> (subrayado).
pJET1.2 directo	CGACTCACTATAGGGAGAG CGGC (SEQ ID NO: 91)	Fermentas (cebador de secuenciación, 23 meros)
pJET1.2 inverso	AAGAACATCGATTTTCCATG GCAG (SEQ ID NO: 92)	Fermentas (cebador de secuenciación, 24 meros)
M13/pUC-directo	GTAAAACGACGGCCAGT (SEQ ID NO: 93)	Jena Bioscience (17 meros)

Construcciones de plásmidos y su transferencia a la cepa H16 de *R. eutropha*. Para la amplificación de las regiones codificantes de *xyIA*, *xyIB*, *xyIE*, *xyIFGH*, *talB* y *tktA* de la cepa K12 de *E. coli* mediante PCR, se utilizaron los oligonucleótidos enumerados en la Tabla 4. Los oligonucleótidos P_{xyIA_ApaI_f} and P_{xyIA_SmaI_r} se utilizaron para la amplificación de *xyIA*, P_{xyIB_SmaI_f} y P_{xyIB_SpeI_r} para *xyIB*, P_{xyIE_SpeI_f} y P_{xyIE_SacII_r} para *xyIE*, y P_{xyIFGH_SpeI_f} y P_{xyIFGH_SacII_r} para *xyIFGH*, respectivamente. Los cebadores contenían un sitio de unión ribosomal para permitir la traducción del ARN en la bacteria diana H16 de *R. eutropha*. Se utilizó ADN polimerasa de inicio en caliente KOD (Merck, Darmstadt, Alemania) para la amplificación por PCR de acuerdo con las instrucciones del fabricante. Luego se ligaron los productos de la PCR al vector de clonación pJET1,2/blunt (Fermentas, Alemania) utilizando la ADN ligasa T4 (Gibco BRL, Gaithersburg, EE. UU.) o al plásmido pBBR1MCS-3 (Kovach et al. 1995), y se transfirió a la cepa TOP10 de *E. coli*. Los plásmidos se aislaron a partir de ampicilina o clones resistentes a la tetraciclina, respectivamente, y se cortaron los fragmentos clonados mediante la restricción con las respectivas enzimas de restricción adecuadas para la clonación adicional, se extrajeron del gel después de la separación utilizando el kit de extracción de gel E.Z.N.A (Omega Biotec, Bangalore, India). Para los experimentos de expresión en la cepa H16 de *R. eutropha*, se utilizaron el vector de amplia gama de hospedadores pBBR1MCS-3 para la clonación de *xyIA*, *xyIB*, *xyIE*, y *xyIFGH*. El vector pBBR1MCS-3 confirió resistencia a la tetraciclina para la selección en la cepa H16 de *E. coli* y *R. eutropha*. Las regiones codificantes de *xyIA* o *xyIB* se cortaron por restricción con *ApaI* y *SmaI* o *SmaI* y

Spel, respectivamente, ligadas al plásmido pBBR1MCS-3 linealizado con *Apal* y *Spel*, que produce los plásmidos pBBR1MCS-3::xylAB (Fig. 8). Para obtener los plásmidos pBBR1MCS-3::xylABE and pBBR1MCS-3::xylABFGH, las regiones codificantes de *xylE* y *xylFGH* se cortaron por restricción con *Spel* y *SacII*, respectivamente, y se ligaron al plásmido pBBR1MCS-3::xylAB linealizado con *Spel* y *SacII* que produce los plásmidos pBBR1MCS-3::xylABE y pBBR1MCS-3::xylABFGH, respectivamente (Fig. 8). Todos los plásmidos se transfirieron a la cepa S17-1 de *E. coli* mediante transformación (Hanahan 1983).

Transferencia de ADN mediante conjugación. La transferencia de los plásmidos pBBR1MCS-3, pBBR1MCS-3::xylABE, pBBR1MCS-3::xylABFGH (Tabla 3) se realizó mediante conjugación aplicando un protocolo descrito previamente (Friedrich et al. 1981), que utiliza S17-1 de *E. coli* como donante y la cepa H16 de *R. eutropha* como receptor.

Análisis de secuencia de ADN. El ADN se secuenció en un sistema de electroforesis del analizador de ADN 3730 de 48 capilares (Applied Biosystems, Foster City, CA). Para la secuenciación de insertos en el vector de clonación pJET1.2/Blunt (Fermentas, Alemania), se utilizaron el cebador de secuenciación directo pJET1.2, de 23 meros y el cebador de secuenciación inversa pJET1.2, de 24 meros. El cebador directo M13/pUC y el cebador inverso M13/pUC universales se utilizaron para secuenciar los insertos en el plásmido pBBR1MCS-3 (Kovach et al. 1995).

Preparación de fracciones de proteínas solubles de la cepa H16 de *R. eutropha*. Se cultivaron células de cepas recombinantes de la cepa H16 de *R. eutropha* en medio de caldo de cultivo. Se recogieron las células mediante centrifugación durante 15 minutos a 4 °C y 3.500 x g, se lavaron con tampón Tris/HCl 50 mM (pH 7,5) y se resuspendieron en dos volúmenes del mismo tampón. La interrupción se realizó mediante la exposición a ultrasonidos utilizando un Sonifer 250 (Branson Sonic Power Company) con una amplitud de 16 µm (1 min/ml; control de salida del 50 %). Durante la exposición a ultrasonidos, se enfriaron las muestras con una mezcla de hielo-NaCl. Se prepararon fracciones de proteínas libres de membrana solubles mediante ultracentrifugación de 60 min de los extractos brutos a 100.000 xg y 4 °C.

Ensayos de actividad de XylA y XylB. Las fracciones de proteínas solubles de la cepa H16 de *R. eutropha* recombinante se aplicaron para las mediciones de actividad de XylA y XylB a 30 °C utilizando un espectrofotómetro Nicolet Evolution 100 UV/VIS (Thermo Electron Corporation, Cambridge, UK). La actividad de XylA (Gao et al. 2002, Mejnen et al. 2008) o XylB (Shamanna y Sangerson 1979, Eliasson et al. 2000) se analizó con un sistema acoplado a NADH con algunas modificaciones, utilizando sorbitol deshidrogenasa o piruvato quinasa y lactato deshidrogenasa, respectivamente, como enzimas auxiliares. La mezcla de reacción tamponada (Tris/HCl 50 mM, pH 7,5) para la medición de la actividad de XylA contenía MgSO₄ 10 mM, trietanolamina 1 mM, NaDH 0,2 mM, D-xilosa 50 mM, sorbitol deshidrogenasa 0,5 U ml⁻¹ y de 5 a 100 µl de extracto soluble. La mezcla de reacción tamponada (Tris/HCl 50 mM, pH 7,5) para la medición de la actividad de xylB contenía MgCl₂ 2 mM, ATP 2 mM, NaDH 0,2 mM, xilulosa 10 mM, PEP 0,2 mM, piruvato quinasa 10 U ml⁻¹, lactato deshidrogenasa 10 U ml⁻¹ y de 5 a 100 µl de extracto soluble.

Utilización de D-xilosa en *R. eutropha*. Después de que la D-xilosa es absorbida por la célula, el gen *xylA* que codifica la enzima D-xilosa isomerasa (XylI o XylA; EC 5.3.1.5) cataliza la conversión reversible de la aldosa D-xilosa en su forma ceto D-xilulosa. Posteriormente, la D-xilulosa se fosforila a D-xilulosa-5-fosfato, utilizando ATP como fuente de energía y fosfato. Este paso es catalizado por una xilulosa quinasa (XylK o XylB; EC 2.7.1.17). La D-xilulosa-5-fosfato se puede utilizar adicionalmente como un intermedio de la ruta de fosfato de pentosa no oxidativa. Como la cepa K12 de *E. coli* es uno de los organismos más ampliamente estudiados que utilizan D-xilosa, los genes *xylA*, que codifican la xilosa-isomerasa, *xylB* que codifica la xiluloquinasa, *xylE* que codifica el transportador de baja afinidad y *xylFGH* que codifica el transportador ABC se amplificaron utilizando ADN genómico de K12 de *E. coli*, introducida episomalmente y expresada en H16 de *R. eutropha*. Las rutas de ingeniería del catabolismo de la xilosa en la cepa recombinante de *R. eutropha* se muestran en la Fig. 7.

Construcción de diferentes vectores de expresión pBBR1 para la cepa H16 de *R. eutropha*. Las regiones codificantes de los genes *xylA* (1.323 pb), *xylB* (1.455 pb), *xylE* (1.476 pb) y *xylFGH* (3.771 pb) se amplificaron mediante PCR a partir del ADN genómico de la cepa K12 *E. coli*. Los tres productos de la PCR comprendidos junto a las regiones codificantes también son sitios de unión al ribosoma adecuados para *R. eutropha* cadena arriba del codón de inicio respectivo para permitir la expresión en este hospedador. Estos fragmentos se clonaron a través del vector pJET1.2/blunt en el vector de amplia gama de hospedadores pBBR1MCS-3 bajo el control del promotor *lac*, que permite la expresión constitutiva de los genes clonados en H16 de *R. eutropha* (Siedow et al. 1999), que produce los plásmidos pBBR1MCS-3::xylAB, pBBR1MCS-3::xylABE y pBBR1MCS-3::xylABFGH (Fig. 8). La construcción precisa de estos plásmidos se confirmó mediante secuenciación. Adicionalmente, la funcionalidad de los genes *xylA* o *xylB* genes en pBBR1MCS-3::xylAB, pBBR1MCS-3::xylABE y pBBR1MCS-3::xylABFGH se confirmó mediante la complementación del mutante HB101 de *E. coli* deficiente en *xylA*, el mutante JW255 de *E. coli* deficiente en *xylB* o el mutante LR2-168 de *E. coli* deficiente en *xylAB*, respectivamente (Hirota et al. 1977). Los mutantes recombinantes de *E. coli* HB101, JW255 o LR2-168 que albergan pBBR1MCS-3::xylAB, pBBR1MCS-3::xylABE o pBBR1MCS-3::xylABFGH, respectivamente, mostraron un buen crecimiento en medio M9 sólido que contenía xilosa al 0,5 % (p/v) como única fuente de carbono después de dos días de incubación a 37 °C. En contraste, el vector pBBR1MCS-3 no confirió crecimiento a los mutantes HB101, JW255 o LR2-168, respectivamente.

Transferencia de plásmidos a la cepa H16 de *R. eutropha* y establecimiento de la utilización de D-xilosa. Los plásmidos pBBR1MCS-3::xyIAB, pBBR1MCS-3::xyIABE y pBBR1MCS-3::xyIABFGH se transfirieron a través de la conjugación de S17-1 de *E. coli* la cepa H16 de *R. eutropha*. Los transconjugantes se seleccionaron en placas de agar MSM que contenían gluconato de sodio al 1 % (p/v) y tetraciclina (12,5 µg/ml), y se transfirieron posteriormente a placas MSM que contenían D-xilosa al 1 % (p/v) como única fuente de carbono. Las cepas recombinantes que albergan el plásmido pBBR1MCS-3::xyIABFGH mostraron un crecimiento débil después de 5 días de incubación.

Optimización evolutiva de la cepa de H16 de *R. eutropha* que alberga el plásmido pBBR1MCS-3::xyIABFGH Aunque la cepa recombinante de H16 de *R. eutropha* que alberga el plásmido pBBR1MCS-3::xyIABFGH se diseñó mediante ingeniería genética para la utilización de D-xilosa, esta cepa mostró solo un crecimiento débil después de 5 días de incubación. Se cultivó una única colonia de esta cepa recombinante en MSM fluido que contenía D-xilosa al 1 % (p/v) como única fuente de carbono durante 2 días a 30 °C. Se transfirieron las diluciones a MSM sólido que contenía D-xilosa al 1 % (p/v) como única fuente de carbono y se incubaron a 30 °C. Se cultivó nuevamente una única colonia que mostró un crecimiento más rápido en un MSM fluido que contenía D-xilosa al 1 % (p/v) como única fuente de carbono, y se transfirió nuevamente una dilución a un MSM sólido que contenía D-xilosa al 1 % (p/v) como única fuente de carbono. Después de la repetición de este procedimiento varias veces, se obtuvo una cepa de H16 de *R. eutropha* que alberga el plásmido pBBR1MCS-3::xyIABFGH que exhibe un crecimiento comparativamente bueno en MSM sólido que contiene D-xilosa al 1 % (p/v) como única fuente de carbono después de 2 días de incubación.

Investigaciones sobre la utilización de varias fuentes de carbono por la cepa recombinante de H16 de *R. eutropha* que alberga el plásmido pBBR1MCS-3::xyIABFGH y de la respectiva cepa optimizada. Células de la cepa recombinante de H16 que alberga pBBR1MCS-3::xyIABFGH y de la respectiva cepa recombinante optimizada de H16, de la cepa G+1 de *R. eutropha*, un mutante que utiliza glucosa de la cepa H16 de *R. eutropha* (Schlegel y Gottschalk 1965) y de la cepa H16 de *R. eutropha*, se cultivaron en placas de agar MSM sólido que contenían diferentes fuentes de carbono (1 % [p/v] de L-arabinosa, D-fructosa, D-galactosa, D-gluconato, D-glucosa, D-maltosa, D-manosa, N-acetilglucosamina, D-ribosa, D-trehalosa, D-xilosa) para comparar el crecimiento de las cepas en varias fuentes de carbono. El tipo silvestre de H16 de *R. eutropha* sirvió como control negativo. Todas las cepas fueron capaces de utilizar D-fructosa, D-gluconato y N-acetilglucosamina como fuentes únicas de carbono. El plásmido pBBR1MCS-3::xyIABFGH que contiene la cepa de la cepa H16 de *R. eutropha* también fue capaz de utilizar glucosa como única fuente de carbono que mostró un crecimiento similar al de la cepa G+1 de H16 de *R. eutropha*. La cepa que contiene el plásmido pBBR1MCS-3::xyIABFGH también fue capaz de utilizar D-xilosa, L-arabinosa y D-galactosa como únicas fuentes de carbono.

Crecimiento de la cepa recombinante de H16 de *R. eutropha* que alberga el plásmido pBBR1MCS-3::xyIABFGH con glucosa, gluconato o xilosa como única fuente de carbono. Las células de la cepa recombinante de H16 de *R. eutropha* que albergan el plásmido pBBR1MCS-3::xyIABFGH se cultivaron en MSM líquido que contenía diferentes concentraciones (0,5 %, 1 %, 2 % y 4 % [p/v]) de D-xilosa, D-glucosa al 1 % (p/v) y Na-gluconato al 1 % (p/v) como única fuente de carbono. También se cultivó el tipo silvestre de la cepa H16 de *R. eutropha* con Na-gluconato al 1 % (p/v) como única fuente de carbono para comparar la tasa de crecimiento. La cepa recombinante mostró un crecimiento similar en Na-gluconato al 1 % (p/v) ($\mu = 0,47 \text{ h}^{-1}$) y también en glucosa al 1 % (p/v) ($\mu = 0,55 \text{ h}^{-1}$) como única fuente de carbono en comparación con la cepa de tipo silvestre en Na-gluconato 1 % (p/v) ($\mu = 0,47 \text{ h}^{-1}$) (Fig. 9). Después de casi 22 h, se alcanzó la fase de crecimiento estacionario. El crecimiento de la cepa recombinante en D-xilosa como única fuente de carbono fue más lento. Después de casi 40 h, se alcanzó la fase de crecimiento estacionario para los cultivos que contenían 0,5 % (p/v) ($\mu = 0,20 \text{ h}^{-1}$) a 2 % (p/v) ($\mu = 0,18 \text{ h}^{-1}$) de D-xilosa como única fuente de carbono. El cultivo que contenía D-xilosa al 4 % (p/v) como única fuente de carbono alcanzó la fase de crecimiento estacionario después de casi 50 h ($\mu = 0,17 \text{ h}^{-1}$) (Fig. 9).

Control de la expresión de xyIA y xyIB en la cepa recombinante de H16 de *R. eutropha* que contiene el plásmido pBBR1MCS-3::xyIABFGH. La presencia de XylA y XylB activas funcionales en la cepa recombinante se confirmó mediante análisis enzimáticos. Ensayos enzimáticos acoplados que emplean las fracciones de proteínas solubles obtenidas de células de la cepa recombinante de H16 que albergan pBBR1MCS-3::xyIABFGH cultivadas en MSM que contiene xilosa al 1 % (p/v) como única fuente de carbono demostraron la presencia de XylA activo (0,035 U mg⁻¹) y XylB (0,364 U mg⁻¹), respectivamente, mientras que estas actividades estaban ausentes en las fracciones de proteínas solubles del control negativo de la cepa H16 que alberga el plásmido pBBR1MCS-3, lo que indica la ausencia de XylA y XylB en la cepa de control negativo.

Análisis del contenido de PHB de células recombinantes de *R. eutropha* por GC. El material celular liofilizado se sometió a metanolisis en presencia de metanol y ácido sulfúrico (MeOH al 85 %, v/v y H₂SO₄ al 15 %, v/v) durante 4 horas a 100 °C, y los ésteres metílicos resultantes de los constituyentes de PHA fueron caracterizados mediante cromatografía de gases utilizando un Agilent 6850 GC (Agilent Technologies, Waldbronn, Alemania) equipado con una columna capilar BP21 (50 m por 0,22 mm; espesor de la película, 250 nm; SGE, Darmstadt, Alemania) y un detector de ionización de llama (Agilent Technologies). Se analizó una porción de 2 µl de la fase orgánica después de la inyección dividida (relación de división, 1:5), y se utilizó un flujo constante de hidrógeno de 0,6 ml min⁻¹ como gas portador. Las temperaturas del inyector fueron de 250 °C y 220 °C, respectivamente. Se aplicó el siguiente programa de temperatura: 120 °C durante 5 min, aumento de 3 °C min⁻¹ hasta 180 °C y aumento de 10 °C min⁻¹ hasta 220 °C y 220 °C durante 31 min. Se identificaron las sustancias mediante la comparación de sus tiempos de retención con los del éster metílico de ácido graso estándar.

Construcción de diferentes vectores adecuados para la expresión de transcetolasa (Tkt) y transaldolasa (TalB) en cepas de los géneros *Ralstonia* y *Cupriavidus*. Las regiones codificantes de los genes que codifican una proteína homóloga a la transcetolasa (Tkt) y transaldolasa (TalB) se amplificaron mediante PCR a partir de ADN genómico de la respectiva cepa hospedadora. Los productos de la PCR pueden comprender, además de las regiones codificantes, sitios de restricción adecuados para la clonación, elementos reguladores adecuados, por ejemplo, una secuencia promotora y sitios de unión a ribosomas para cepas de los géneros *Ralstonia* y *Cupriavidus* en cadena arriba del codón de inicio respectivo para permitir la expresión en este hospedador. Estas características pueden proporcionarse ya sea mediante oligonucleótidos sintéticos o mediante PCR de fusión. Estos fragmentos se pueden clonar a través de vectores de clonación en vectores adecuados para la expresión en cepas de los géneros *Ralstonia* y *Cupriavidus* bajo el control de un promotor adecuado, que permite la expresión de los genes clonados en la respectiva cepa hospedadora. Las regiones codificantes de los genes que codifican una proteína homóloga a la transcetolasa (Tkt) y transaldolasa (TalB) se pueden clonar por separado o en combinación con diferentes genes, por ejemplo, *xylA*, *xylB*, *xylG*, y *xylF*. La construcción precisa de estos plásmidos se puede confirmar mediante secuenciación. Los plásmidos construidos pueden transferirse mediante conjugación de S17-1 de *E. coli* o mediante electroporación a cepas de los géneros *Ralstonia* y *Cupriavidus*. Los transconjugantes o transformantes, respectivamente, se pueden seleccionar en placas de agar MSM que contienen gluconato de sodio al 1 % (p/v) y una selección adecuada para la estabilidad del plásmido. Las cepas recombinantes de los géneros *Ralstonia* y *Cupriavidus* se pueden caracterizar por su crecimiento en diferentes fuentes de carbono.

Ejemplo 3

Cepas bacterianas, plásmidos, oligonucleótidos y condiciones de cultivo. Todas las bacterias, plásmidos y cebadores utilizados en esta aplicación se enumeran en la Tabla 5. Se cultivaron células de *E. coli* en caldo de lisogenia (LB) (Sambrook et al., 1989) o en medio de sales minerales (MSM) complementado con glucosa al 1 % (p / v) de acuerdo con Schlegel et al. (1961). Se cultivaron células de *R. eutropha* en caldo nutritivo (NB, de sus siglas en inglés) (Sambrook et al., 1989) o en MSM suplementado con gluconato de sodio al 1% (p / v) y oleato de sodio (0,4 %, p / v) para la síntesis de FAEE. Se incubaron los cultivos líquidos en matraces Erlenmeyer desmontados en un agitador rotatorio horizontal a una agitación de 110 rpm a 37 °C (*E. coli*) o 30 °C (cepas de *R. eutropha*). Se aplicaron antibióticos de acuerdo con Sambrook et al. (1989) y como se indica en el texto.

Tabla 5. Cepas bacterianas, plásmidos y oligonucleótidos

Cepa, plásmido o cebador	Características relevantes	Fuente o referencia
Cepas		
TOP10 de <i>E. coli</i>	<i>F-mcrA</i> , $\Delta(mrr-hsdRMS-mcrBC)$ <i>f80lacZ</i> $\Delta M15$, $\Delta lacX74$, <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , $\Delta(ara-leu)7697$, <i>galil</i> , <i>galK</i> , <i>rpsL</i> , <i>endA1</i> , <i>nupG</i>	Invitrogen
S17-1 de <i>E. coli</i>	<i>recA1</i> , <i>thi1</i> , <i>hsdR17(rk-, mk+)</i> , <i>proA</i> , genes <i>tra</i> del plásmido RP4 cromosómicamente integrados (cepa de movilización)	Simon et al. 1983
<i>R. eutropha</i> $\Delta phaC1$	mutante negativo para PHB de H16 de <i>R. eutropha</i>	Peplinski et al., 2010
Plásmidos		
pBBR1MCS-3	Tc ^R , <i>lacZ</i> α , <i>mob</i> , <i>rep</i>	Kovach et al., 1995
pMicrodiesel	Amp ^R , <i>atfA</i> , <i>adhB</i> , <i>pdC</i>	Kalscheuer et al., 2006
pBBR1MCS-3:: <i>atfA/pdc/adhB</i>	con <i>atfA/pdc/adhB</i> como fragmento <i>XbaI/ApaI</i>	en este ejemplo

Oligonucleótidos

M13 directo 5' GTAAAACGACGGCCAGT 3' (SEQ ID NO:94)

M13 inverso 5' AACAGCTATGACCATG 3' (SEQ ID NO: 95)

Aislamiento, análisis y modificación de ADN. Se preparó ADN plasmídico de cepas de *R. eutropha* a partir de lisados crudos por el método de extracción alcalina (Birnboim y Doly, 1979). Se preparó el ADN plasmídico de *E. coli* con un kit comercial de preparación de plásmidos (Kit de plásmidos Roti Prep, Roth, Karlsruhe) aplicando las condiciones recomendadas por el fabricante. Se restringió el ADN con endonucleasas de restricción (Fermentas GmbH, St. Leon-Rot) bajo las condiciones recomendadas por el fabricante. Todos las demás manipulaciones y procedimientos genéticos se realizaron según lo descrito por Sambrook et al. (1989).

Transferencia de ADN mediante conjugación. La transferencia de los plásmidos pBBR1MCS-3 y pBBR1MCS-3::*atfA/pdc/adhB* se realizó mediante conjugación aplicando el protocolo descrito por Friedrich et al. (1981). Se utilizó S17-1 de *E. coli* como cepa donante y cepas Δ *phaC1* de *R. eutropha* como receptor.

Preparación de las fracciones de proteínas solubles de las cepas de *R. eutropha* y *E. coli*. Las células de cepas recombinantes de *R. eutropha* o *E. coli* se cultivaron como se describió anteriormente en presencia de gluconato de sodio al 1 % (p/v) o glucosa al 1 % (p/v), respectivamente. Se recogieron las células mediante centrifugación durante 15 minutos a 4 °C y 3500 x g, se lavaron en un tampón adecuado como se describe anteriormente y se resuspendieron en dos volúmenes del mismo tampón. Se rompieron las células mediante exposición a ultrasonidos empleando una sonda MS72 y un homogeneizador ultrasónico Sonoplus GM200 (Bandelin, Berlín) con una amplitud de 10 µm (2 min/ml; control de salida del 50 %). Durante la exposición a ultrasonidos las muestras se enfriaron con hielo. Las concentraciones de proteína en extractos celulares crudos se determinaron según lo descrito por Bradford (1976).

Análisis de secuencia de ADN. El ADN se secuenció en un sistema de electroforesis del analizador de ADN 3730 de 48 capilares (Applied Biosystems, Foster City, CA). Para la secuenciación de insertos en el plásmido pBBR1MCS-3 (Kovach et al. 1995) se utilizaron los cebadores universales "M13 directo" y "M13 inverso".

Construcción de los plásmidos usados. Como el vector de expresión en H16 de *R. eutropha*, se eligió el vector pBBR1MCS-3 que permite una expresión estable de los genes que contienen en *R. eutropha*. Los genes necesarios para la síntesis de FAEE se cortaron del plásmido pMicrodiesel (Kalscheuer et al., 2006) mediante restricción con *Xba*I y *Ap*aI lo que da como resultado un fragmento de 4,8 kpb que contiene los genes *atfA*, *pdc* y *adhB*.

Este fragmento se ligó al plásmido pBBR1MCS-3 restringido *Xba*I y *Ap*aI utilizando ADN ligasa T4 (Fermentas GmbH, St. Leon-Rot), lo que dio como resultado el plásmido pBBR1MCS-3::*atfA/pdc/adhB* de 9,9 kpb. El plásmido construido se transfirió mediante transformación (Hanahan 1983) a Top10 de *E. coli* para su verificación y finalmente a la cepa S17-1 de *E. coli* que sirve como donante para los experimentos de conjugación.

Ensayos de actividad WS/DGAT, PDC y ADHB. Se aplicaron fracciones de proteínas solubles de cepas recombinantes de *R. eutropha* para las mediciones de actividad WS/DGAT, PDC y ADHB a 30 °C utilizando un espectrofotómetro Nicolet Evolution 100 UV/VIS (Thermo Electron Corporation, Cambridge, UK). La actividad de PDC se ensayó con un sistema acoplado a NADH utilizando ADH como una enzima auxiliar que mide la producción de NAD⁺ mediante la disminución de la absorción a 340 nm. La mezcla de reacción tamponada (fosfato de potasio (KP) 50 mM, pH 6,5) contenía MgSO₄ 5 mM, pirofosfato de tiamina 0,1 mM (TPP), NaDH 0,175 mM, piruvato 17,5 mM, ADH 0,1 U/ml y solución de proteína 100 µg/ml, del extracto soluble.

La actividad de ADH se midió directamente mediante la disminución de NADH debido a la reducción de acetaldehído. Se utilizó la misma mezcla de reacción que para el ensayo de actividad de PDC, aunque el piruvato se reemplazó por acetaldehído 0,1 mM y la ADH se reemplazó por el tampón KP.

La actividad de WS/DGAT se midió radiométricamente mediante la síntesis de ésteres de cera marcados a partir de hexadecanol o etanol y 1-¹⁴C-palmitoil coenzima A. La mezcla de reacción tamponada (tampón de fosfato de sodio 125 mM, pH 7,4) contenía hexadecanol o etanol 3,75 mM, albumina de suero bovino (BSA) 12,5 µg/ml, 1-¹⁴C-palmitoil CoA 4,72 µM y extracto crudo de proteína 200 µg/ml. Antes de utilizar, se preparó una emulsión doble concentrada de hexadecanol, BSA y tampón fosfato de sodio mediante ultrasonidos (1 min/ml). Se incubó la mezcla de reacción durante 30 minutos a 35 °C y la reacción se detuvo mediante la adición de un volumen de cloroformo-metanol (1:1; v/v). Después de girar en vórtice vigoroso durante 1 minuto, se separaron las fases mediante centrifugación durante 2 minutos a 16000 x g. La fase orgánica se eliminó y se aplicó a una placa de cromatografía de capa fina (TLC, de sus siglas en inglés) (gel de sílice 60, láminas de aluminio de 20x20 cm, Merck KGaA, Darmstadt) que luego se desarrolló con hexano:éter dietílico:ácido acético (90:7,5:1; (v/v/v) como el disolvente. Las manchas de éster de cera marcadas radiactivamente se visualizaron mediante un PhosphorImager (Bio-Rad Laboratories GmbH, Munich) después de exponer la placa de TLC desarrollada a una pantalla de fósforo. Como sustancias estándar para la TLC

se utilizaron oleiloleato o etiloleato que se visualizaron mediante tinción con vapor de yodo.

Determinación cuantitativa y cualitativa de etanol sintetizado y FAEE.

Las cantidades de etanol en medios de cultivo libres de células se determinaron mediante cromatografía líquida de alto rendimiento (HPLC, de sus siglas en inglés). Se realizó el análisis de HPLC de muestras de 20 µl con una HPLC LaChrom Elite (VWR-Hitachi International GmbH, Darmstadt) equipada con una columna C avanzada de Metacarb 67H (longitud: 300 mm; diámetro interno de 6,5 mm; Varian, Palo Alto, EE. UU.) y un detector de índice de refracción (RI, de sus siglas en inglés) Tipo 2490 (VWR-Hitachi International GmbH, Darmstadt).

Las FAEE en células liofilizadas se analizaron mediante TLC, cromatografía de gases (GC, de sus siglas en inglés) y cromatografía de gases acoplada y espectrometría de masas (GC/MS, de sus siglas en inglés). Para el análisis de TLC cualitativo, se mezclaron 100 mg de materia seca celular con 1 ml de cloroformo:metanol (2:1; v/v) y se rompieron con un molino de bolas tipo MM301 (Retsch, Haan). Después de la centrifugación durante 10 minutos a 16000 x g, se separó el sobrenadante orgánico mediante TLC como se describe para los ensayos de actividad WS/DGAT. El etiloleato sirvió como sustancia estándar. Las manchas se visualizaron mediante tinción con vapor de yodo y se rasparon las manchas de interés de la placa de TLC, se eluyeron con cloroformo:metanol (2:1, v/v) y se analizaron mediante GC o GC/MS. Los análisis de GC de 10-100 mg de extractos celulares se realizaron con un cromatógrafo de gases capilar Agilent 6850 (Agilent Technologies GmbH, Waldbronn) equipado con una columna capilar BP21 (longitud: 50 m; diámetro interno: 0,22 mm; espesor de la película 250 nm; SGE GmbH, Darmstadt) y un detector de ionización de llama (FID). Con una relación de división de 1:15, se inyectaron 2 µl de muestra mediante un muestreador automático de la serie 6850 de Agilent (Agilent Technologies GmbH, Waldbronn). Se utilizó como gas portador H₂ (99,999 %) a un caudal de 0,6 ml/min. Las temperaturas de entrada y del detector fueron de 250 °C y 275 °C, respectivamente, la temperatura de la columna se mantuvo a 120 °C durante 5 minutos, seguido de un aumento de 3 °C/min hasta 180 °C, un segundo aumento de 10 °C/min hasta 220 °C y finalmente se mantuvo a 220 °C durante 31 minutos.

Los análisis de GC/MS se llevaron a cabo en un sistema GC serie 6890 (Hewlett Packard, Waldbronn) equipado con una columna capilar BPX35 (longitud: 50 m; diámetro interno: 0,25 mm; espesor de la película 250 nm; SGE GmbH, Darmstadt). GC se acopló con un detector selectivo de la masa de ionización electrónica de la Serie 5973 (EI-MSD) (Hewlett Packard, Waldbronn). El helio (99,999%) sirvió como gas portador con un caudal constante de 0,6 ml/min. Se inyectaron 3 µl de muestra con un autoinyector Serie 7683 (Hewlett Packard, Waldbronn) con una relación de división de 1:20. La temperatura de entrada fue de 250 °C, la de la salida fue de 240 °C. Se aplicó el siguiente programa de temperatura: 5 min a 120 °C, incremento de 3 °C/min hasta 180 °C, seguido de un aumento de 5 °C/min hasta 210 °C y, finalmente, la temperatura se mantuvo durante 29 min a 210 °C. Los datos se analizaron con el programa de búsqueda espectral de masas NIST (Stein et al., 1998)

Transferencia de plásmidos a *ΔphaC1* de *R. eutropha* y determinación de actividades enzimáticas. Ambos plásmidos, pBBR1MCS-3 como vector de control y pBBR1MCS-3::*atfA/pdc/adhB*, se transfirieron mediante la conjugación de S17-1 de *E. coli* a *ΔphaC1* de *R. eutropha*, lo que dio como resultado las cepas pBBR1MCS-3 *ΔphaC1* de *R. eutropha* y pBBR1MCS-3::*atfA/pdc/adhB ΔphaC1* de *R. eutropha*. Los transconjugantes podrían seleccionarse en placas de agar MSM que contengan gluconato de sodio al 1 % (p/v) y tetraciclina (12,5 µg/ml) después de 12 h de incubación a 30 °C.

Confirmación de la presencia de WS/DGAT, PDC y ADHB activos funcionales.

La presencia de WS/DGAT, PDC y ADHB activos funcionales en las cepas recombinantes se confirmó mediante análisis enzimáticos.

Ensayos enzimáticos fotométricos directos o acoplados que emplean las fracciones de proteínas solubles obtenidas de células de las cepas recombinantes de pBBR1MCS-3 *ΔphaC1* de *R. eutropha* y pBBR1MCS-3::*atfA/pdc/adhB* de *ΔphaC1* de *R. eutropha* cultivadas durante 12 horas a 30 °C en MSM que contiene gluconato de sodio al 1 % (p/v) como única fuente de carbono, demostró la presencia de PDC activo (18,73 U/mg) y ADH B (1,51 U/mg) en pBBR1MCS-3::*atfA/pdc/adhB* de *ΔphaC1* de *R. eutropha*, en contraste con la cepa de control pBBR1MCS-3 de *ΔphaC1* de *R. eutropha* (PDC: 3,52 U/mg y ADH B: 0,25 U/mg). Los ensayos enzimáticos radiométricos para WS/DGAT revelaron la síntesis de ésteres de cera y ésteres etílicos radiomarcados a partir de hexadecanol o etanol, respectivamente, y 1-¹⁴C-palmitoil CoA en las fracciones de proteína soluble de pBBR1MCS-3::*atfA/pdc/adhB* de *ΔphaC1* de *R. eutropha* pero no de pBBR1MCS-3 de *ΔphaC1* de *R. eutropha*, lo que indica la ausencia de WS/DGAT en la cepa de control negativo. Para la actividad de WS/DGAT no se determinaron actividades específicas.

Análisis de la biosíntesis de etanol y FAEE mediante HPLC, GC, GC/MS y TLC.

Se analizó la biosíntesis de etanol y FAEE en las cepas pBBR1MCS-3 de *ΔphaC1* de *R. eutropha* y pBBR1MCS-3::*atfA/pdc/adhB ΔphaC1* de *R. eutropha*. Las células se cultivaron en precultivos MSM que contenían gluconato al 2 % (p/v) y tetraciclina o medio NB con tetraciclina durante 12 horas a 30 °C y se utilizaron para inocular cultivos principales del mismo medio en matraces Erlenmeyer de 1 l con 2,5 % del volumen de cultivo de 200 ml. Un cultivo

MSM adicional (400 ml) contenía adicionalmente 0,4 % (p/v) de oleato de Na. Cada 4-8 h las muestras se retiraron y se centrifugaron. Se utilizó el sobrenadante para la determinación de etanol por HPLC. Se lavaron los sedimentos celulares dos veces con solución salina estéril, se liofilizaron y se utilizaron para la determinación del éster etílico por GC, GC/MS y TLC.

Sin oleato de Na como cosustrato, se encontraron cantidades máximas de etanol de 1,51 g/l después de 22 h en pBBR1MCS-3::atfA/pdc/adhB de *AphaC1* de *R. eutropha* crecido en NB, mientras que la cepa de control alcanzó 1,16 g/l después de 14 h. Con medio MSM, las cantidades máximas de etanol alcanzaron 0,85 g/l para pBBR1MCS-3::atfA/pdc/adhB de *AphaC1* de *R. eutropha* después de 32 h y 0,44 g/l para la cepa de control después de 8 h. Si se añadió oleato de Na al medio MSM, la concentración máxima de etanol alcanzó 4,38 g/l con pBBR1MCS-3::atfA/pdc/adhB de *AphaC1* de *R. eutropha* después de 36 h y 1,34 g/l con la cepa de control después de 30 h. Por lo tanto, se podría demostrar que pBBR1MCS-3::atfA/pdc/adhB de pBBR1MCS es capaz de sintetizar concentraciones de etanol tres veces más altas como la cepa de control si se cultiva en MSM con gluconato y oleato de Na.

Posiblemente los casos de FAEE se analizaron mediante TLC y GC a partir de la masa celular liofilizada. Sin la adición de oleato de Na, ni las células cultivadas en NB ni las cultivadas en MSM mostraron una síntesis de etiloleato como se reveló mediante el análisis de GC y TLC. Si se añadió oleato de Na a la masa celular seca media de pBBR1MCS-3::atfA/pdc/adhB de pBBR1MCS mostró concentraciones de etiloleato de 0,45 % (p/p) después de 36 h y de 0,23 % (v/v) después de 48 h. La masa celular seca de las cepas de control no proporcionó un pico correspondiente al etiloleato en los análisis de GC. Comparable a los análisis de GC, los análisis de TLC de extractos crudos de las masas celulares secas respectivas revelaron la aparición de sustancias que corresponden a etiloleato solo en pBBR1MCS-3::atfA/pdc/adhB que contiene la cepa *AphaC1* de *R. eutropha*.

Mientras que los análisis de GC/MS de la mancha de etiloleato supuesta raspada de las placas de TLC no dieron resultados, los análisis de GC/MS del etiloleato que contiene la muestra de 36 h y la muestra de 48 h revelaron la presencia de etilpalmitoleato en la masa celular seca. Otros FAEE no pudieron ser detectados. Por tanto, si se complementa con oleato de Na, la cepa pBBR1MCS-3::atfA/pdc/adhB de *AphaC1* de *R. eutropha* es capaz de sintetizar FAEE.

Referencias

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers y D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Altschul, S. F., T.L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller y D. J. Lipman 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402
- Anderson, R. L. y V. L. Sapico. 1975. D-fructose (D-mannose) kinase. *Methods Enzymol.* 42:39-43.
- Aneja, K. K., R. D. Ashby y D. K. Y. Solaiman. 2009. Altered composition of *Ralstonia eutropha* poly(hydroxyalkanoate) through expression of PHA synthase from *Allochrocatium vinosum* ATCC 35206. *Biotechnol. Lett.* 31:1601-1612.
- Birnboim, H. C. y J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513-1523.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248-254.
- Buchholz, B., Nordsiek, G., Meister, M. Bowien, B. 1994. Transfer of genes from *Pseudomonas saccharophila* to construct xylose-utilizing strains of *Alcaligenes eutrophus*. *Current Microbiology* 29:157-162.
- Coulombel, C., M. J. Foglietti y F. Percheron. 1982. Identification and kinetic studies of an inducible mannokinase from a *Streptomyces* strain. *Biochim. Biophys. Acta* 706:117-122.
- Eliasson, A., E. Boles, B. Johansson, M. österberg, J. M. Thevelein, I. Spencer-Martins, H. Juhnke y B. Hahn-Hagerdal. 2000. Xylulose fermentation by mutant and wild-type strains of *Zygosaccharomyces* and *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 53:376-382.
- Franklin, F. C. H., M. Bagdasarian, M. M. Bagdasarian y K. N. Timmis. 1981. Molecular and functional analysis of the TOL plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring meta cleavage pathway. *Proc. Natl. Acad. Sci. USA* 78:7458-7462.
- Friedrich, B., C. Hogrefe y H. G. Schlegel. 1981. Naturally occurring genetic transfer of hydrogen-oxidizing ability between strains of *Alcaligenes eutrophus*. *J. Bacteriol.* 147:198-205.
- Gao, Q., M. Zhang, J. D. McMillan y D. S. Kompala. 2002. Characterization of heterologous and native enzyme activity profiles in metabolically engineered *Zymomonas mobilis* strains during batch fermentation of glucose and xylose mixtures. *Appl. Biochem. Biotechnol.* 98-100:341-355.
- Gottschalk, G., U. Eberhardt y H. G. Schlegel. 1964. Verwertung von Fructose durch *Hydrogenomonas* H16. *Arch. Mikrobiol.* 48:95-108.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557-580.
- Hirota, Y., H. Suzuki, Y. Nishimura y Yasuda. 1977. On the process of cellular division in *Escherichia coli*: a mutant of *E. coli* lacking a murein-lipoprotein. *Proc. Natl. Acad. Sci. U. S. A.* 74:1417-1420.
- Jeffke, T., N.-H. Gropp, C. Kaiser, C. Grzesik, B. Kusian y B. Bowien. 1999. Mutational analysis of the *cbb* operon (CO₂ Assimilation) promoter of *Ralstonia eutropha*. *J. Bacteriol.* 181:4374-4380.
- Kalscheuer, R., T. Stolling y A. Steinbüchel. 2006. Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology* 152:2529-2536

- Kang, S. y A. Markovitz. 1967. Induction of capsular polysaccharide synthesis by rho-fluorophenylalanine in *Escherichia coli* wild type and strains with altered phenylalanyl soluble ribonucleic acid synthetase. *J. Bacteriol.* 93:584-591.
- König, C., I. Sammler, E. Wilde y H. G. Schlegel. 1969. Constitutive glucose-6-phosphate dehydrogenase in mutants utilizing glucose, which are derived from cryptic wildtype strains. *Arch. Mikrobiol.* 67:51-57.
- Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop 2º y K. M. Peterson. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166:175-176.
- Laemmli, U. K. 1970 Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Meijnen, J.-P., J. H. de Winde y H. J. Ruijsenaars. 2008. Engineering *Pseudomonas putida* S12 for efficient utilization of D-xylose and L-arabinose. *Appl. Env. Microbiol.* 74: 5031-5037.
- Park, H.-C., K.-J. Lim, J.-S. Park, Y.-H. Lee y T.-L. Huh. 1995. High frequency transformation of *Alcaligenes eutrophus* producing poly-β-hydroxybutyric acid by electroporation. *Biotechnol. Tech.* 9:31-34.
- Parker, C., W. O. Barnell, J. L. Snoep, L. O. Ingram y T. Conway. 1995. Characterization of the *Zymomonas mobilis* glucose facilitator gene product (*glf*) in recombinant *Escherichia coli*: examination of transport mechanism, kinetics and the role of glucokinase in glucose transport. *Mol. Microbiol.* 15:795-802.
- Peplinski K., Ehrenreich A., Döring C., Bömeke M., Reinecke F., Hutmacher C. y Steinbüchel A., 2010. Genome-wide transcriptome analyses of the 'Knallgas' bacterium *Raistonia eutropha* H16 with regard to polyhydroxyalkanoate metabolism. *Microbiology*, 156:2136-2152
- Pohlmann, A., W. F. Fricke, F. Reinecke, B. Kusian, H. Liesegang, R. Cramm, T. Eitinger, C. Ewerling, M. Potter, E. Schwartz, A. Strittmatter, I. Voss, G. Gottschalk, A. Steinbüchel, B. Friedrich y B. Bowien. 2006. Genome sequence of the bioplastic-producing "Knallgas" bacterium *Raistonia eutropha* H16. *Nat. Biotechnol.* 24:1257-1262.
- Pries, A., A. Steinbüchel y H. G. Schlegel. 1990. Lactose and galactose utilizing strains of poly(hydroxyalkanoic acid) accumulating *Alcaligenes eutrophus* and *Pseudomonas saccharophila* obtained by recombinant DNA technology. *Appl. Microbiol. Biotechnol.* 33:410-417.
- Porthun, A., M. Bernhard, B. Friedrich. 2002. Expression of a functional NAD-reducing [NiFe] hydrogenase from the Gram-positive *Rhodococcus opacus* in the Gram-negative *Raistonia eutropha*. *Arch. Microbiol.* 177:159-166.
- Reinecke, F. y A. Steinbüchel. 2009. *Raistonia eutropha* strain H16 as model organism for PHA metabolism and for biotechnological production of technically interesting biopolymers. *J. Mol. Microbiol. Biotechnol.* 16:91-108.
- Sambrook, J., E. F. Fritsch y T. Maniatis, 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Press, Cold Spring Harbor, Nueva York.
- Schlegel, H. G., H. Kaltwasser y G. Gottschalk. 1961. Ein Submersverfahren zur Kultur wasserstoffoxydierender Bakterien: Wachstumsphysiologische Untersuchungen. *Arch. Mikrobiol.* 38:209-222.
- Schlegel, H. G. y G. Gottschalk. 1965. Verwertung von Glucose durch eine Mutante von *Hydrogenomonas* H16. *Biochem. Z.* 341:249-259.
- Schaferjohann, J., R. Bednarski y B. Bowien. 1996. Regulation of CO₂ assimilation in *Raistonia eutropha*: Premature transcription termination within the *cbb* operon. *J. Bacteriol.* 178: 6714-6719.
- Schwartz, E., U. Gerischer y B. Friedrich. 1998. Transcriptional regulation of *Alcaligenes eutrophus* hydrogenase genes. *J. Bacteriol.* 180:3197-3204.
- Sebastian, J. y C. Asensio. 1972. Purification and properties of the mannokinase from *Escherichia coli*. *Arch. Biochem. Biophys.* 151:227-233.
- Shamanna, D. K y K. E. Sanderson. 1979. Uptake and catabolism of D-xylose in *Salmonella typhimurium* LT2. *J. Bacteriol.* 139:64-70.
- Siedow, A., R. Cramm, R. A. Siddiqui y B. Friedrich. 1999. A megaplasmid- borne anaerobic ribonucleotide reductase in *Alcaligenes eutrophus* H16. *J. Bacteriol.* 181:4919-4928.
- Simon R., Priefer U. y A. Punier. 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria, *Biotechnology* 1:784-794
- Solaiman, D. K. Y. y B. M. Swingle. 2010. Isolation of novel *Pseudomonas syringae* promoters and functional characterization in polyhydroxyalkanoate- producing pseudomonads. *New Biotechnol.* doi:10.1016/j.nbt.2009.12.003
- Solaiman, D. K., B. M. Swingle y R. D. Ashby. 2010. A new shuttle vector for gene expression in biopolymer-producing *Raistonia eutropha*. *J. Microbiol. Methods* doi:10.1016/j.mimet.2010.04.010.
- Stein, S., Levitsky, A., Fateev, O. & Mallard, G. (1998) The NIST Mass Spectral Search Program. Windows-Software Version 1.6d.
- Verlinden, R.A.J., Hill, D.J., Kenward, M.A. Williams, C.D. y Radecka, I. 2007. Bacterial synthesis of biodegradable polyhydroxyalkanoates. *J. Appl. Microbiol.* 102:1437-1449.
- Weber, K., Osborn, M. (1969) The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Biol. Chem.* 244:4406-4412.
- West, S. E., H. P. Schweizer, C. Dali, A. K. Sample, and L. J. Runyen-Janecky. 1994. Construction of improved *Escherichia-Pseudomonas* shuttle vectors derived from pUC18/19 and sequence of the region required for their replication in *Pseudomonas aeruginosa*. *Gene* 148:81-86.
- Windhövel, U. y B. Bowien. 1990. On the operon structure of the *cfx* gene clusters in *Alcaligenes eutrophus*. *Arch. Microbiol.* 154:85-91.

LISTADO DE SECUENCIAS [0183]

<110> Neste Oil OYJ

<120> Microorganismos con amplia gama de utilización de sustratos

<130> BP204594EP

<150> EP10169808.2

<151> 16/07/2010

<150> US61/364898

<151> 16/07/2010

<160> 95

<170> PatentIn versión 3.5

<210> 1

<211> 391

<212> PRT

<213> *Escherichia coli*

<400> 1

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Met Glu Asn Pro Ser Ser Gln Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Val
35 40 45

Gln Asn Ala Ala Gly Asp Ile Val Ser Leu Arg Asp Val Ile Glu Ser
50 55 60

Asp Lys Ser Thr Leu Leu Gly Glu Ala Val Ala Lys Arg Phe Gly Glu
65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
85 90 95

Gln Val His Pro Asn Lys His Asn Ser Glu Ile Gly Phe Ala Lys Glu
100 105 110

Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
115 120 125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
130 135 140

Met Asn Ala Phe Arg Glu Phe Ser Glu Ile Val Ser Leu Leu Gln Pro
145 150 155 160

Val Ala Gly Ala His Pro Ala Ile Ala His Phe Leu Gln Gln Pro Asp
 165 170 175
 Ala Glu Arg Leu Ser Glu Leu Phe Ala Ser Leu Leu Asn Met Gln Gly
 180 185 190
 Glu Glu Lys Ser Arg Ala Leu Ala Ile Leu Lys Ser Ala Leu Asp Ser
 195 200 205
 Gln Gln Gly Glu Pro Trp Gln Thr Ile Arg Leu Ile Ser Glu Phe Tyr
 210 215 220
 Pro Glu Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
 225 230 235 240
 Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
 245 250 255
 Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val
 260 265 270
 Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
 275 280 285
 Ala Asn Val Lys Phe Glu Ala Lys Pro Ala Asn Gln Leu Leu Thr Gln
 290 295 300
 Pro Val Lys Gln Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Asp Asp
 305 310 315 320
 Phe Ala Phe Ser Leu His Asp Leu Ser Asp Lys Glu Thr Thr Ile Ser
 325 330 335
 Gln Gln Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Asp Ala Thr Leu
 340 345 350
 Trp Lys Gly Ser Gln Gln Leu Gln Leu Lys Pro Gly Glu Ser Ala Phe
 355 360 365
 Ile Ala Ala Asn Glu Ser Pro Val Thr Val Lys Gly His Gly Arg Leu
 370 375 380
 Ala Arg Val Tyr Asn Lys Leu
 385 390

<210> 2
 <211> 473

ES 2 696 235 T3

<212> PRT

<213> *Zymomonas mobilis*

<400> 2

5

```

Met Ser Ser Glu Ser Ser Gln Gly Leu Val Thr Arg Leu Ala Leu Ile
 1          5          10          15

Ala Ala Ile Gly Gly Leu Leu Phe Gly Tyr Asp Ser Ala Val Ile Ala
 20          25          30

Ala Ile Gly Thr Pro Val Asp Ile His Phe Ile Ala Pro Arg His Leu
 35          40          45

Ser Ala Thr Ala Ala Ala Ser Leu Ser Gly Met Val Val Val Ala Val
 50          55          60

Leu Val Gly Cys Val Thr Gly Ser Leu Leu Ser Gly Trp Ile Gly Ile
 65          70          75          80

Arg Phe Gly Arg Arg Gly Gly Leu Leu Met Ser Ser Ile Cys Phe Val
 85          90          95

Ala Ala Gly Phe Gly Ala Ala Leu Thr Glu Lys Leu Phe Gly Thr Gly
100          105          110

Gly Ser Ala Leu Gln Ile Phe Cys Phe Phe Arg Phe Leu Ala Gly Leu
115          120          125

Gly Ile Gly Val Val Ser Thr Leu Thr Pro Thr Tyr Ile Ala Glu Ile
130          135          140

Ala Pro Pro Asp Lys Arg Gly Gln Met Val Ser Gly Gln Gln Met Ala
145          150          155          160

Ile Val Thr Gly Ala Leu Thr Gly Tyr Ile Phe Thr Trp Leu Leu Ala
165          170          175

His Phe Gly Ser Ile Asp Trp Val Asn Ala Ser Gly Trp Cys Trp Ser
180          185          190

Pro Ala Ser Glu Gly Leu Ile Gly Ile Ala Phe Leu Leu Leu Leu
195          200          205

Thr Ala Pro Asp Thr Pro His Trp Leu Val Met Lys Gly Arg His Ser
210          215          220

Glu Ala Ser Lys Ile Leu Ala Arg Leu Glu Pro Gln Ala Asp Pro Asn

```

ES 2 696 235 T3

225		230		235		240
Leu Thr Ile Gln Lys	Ile Lys Ala Gly Phe Asp Lys Ala Met Asp Lys					
	245		250		255	
Ser Ser Ala Gly Leu Phe Ala Phe Gly Ile Thr Val Val Phe Ala Gly						
	260		265		270	
Val Ser Val Ala Ala Phe Gln Gln Leu Val Gly Ile Asn Ala Val Leu						
	275		280		285	
Tyr Tyr Ala Pro Gln Met Phe Gln Asn Leu Gly Phe Gly Ala Asp Thr						
	290		295		300	
Ala Leu Leu Gln Thr Ile Ser Ile Gly Val Val Asn Phe Ile Phe Thr						
	305		310		315	
Met Ile Ala Ser Arg Val Val Asp Arg Phe Gly Arg Lys Pro Leu Leu						
	325		330		335	
Ile Trp Gly Ala Leu Gly Met Ala Ala Met Met Ala Val Leu Gly Cys						
	340		345		350	
Cys Phe Trp Phe Lys Val Gly Gly Val Leu Pro Leu Ala Ser Val Leu						
	355		360		365	
Leu Tyr Ile Ala Val Phe Gly Met Ser Trp Gly Pro Val Cys Trp Val						
	370		375		380	
Val Leu Ser Glu Met Phe Pro Ser Ser Ile Lys Gly Ala Ala Met Pro						
	385		390		395	
Ile Ala Val Thr Gly Gln Trp Leu Ala Asn Ile Leu Val Asn Phe Leu						
	405		410		415	
Phe Lys Val Ala Asp Gly Ser Pro Ala Leu Asn Gln Thr Phe Asn His						
	420		425		430	
Gly Phe Ser Tyr Leu Val Phe Ala Ala Leu Ser Ile Leu Gly Gly Leu						
	435		440		445	
Ile Val Ala Arg Phe Val Pro Glu Thr Lys Gly Arg Ser Leu Asp Glu						
	450		455		460	
Ile Glu Glu Met Trp Arg Ser Gln Lys						
	465		470			

ES 2 696 235 T3

<210> 3
 <211> 302
 <212> PRT
 <213> *Escherichia coli*

5

<400> 3

```

Met Arg Ile Gly Ile Asp Leu Gly Gly Thr Lys Thr Glu Val Ile Ala
 1           5           10           15

Leu Gly Asp Ala Gly Glu Gln Leu Tyr Arg His Arg Leu Pro Thr Pro
          20           25           30

Arg Asp Asp Tyr Arg Gln Thr Ile Glu Thr Ile Ala Thr Leu Val Asp
          35           40           45

Met Ala Glu Gln Ala Thr Gly Gln Arg Gly Thr Val Gly Met Gly Ile
 50           55           60

Pro Gly Ser Ile Ser Pro Tyr Thr Gly Val Val Lys Asn Ala Asn Ser
65           70           75           80

Thr Trp Leu Asn Gly Gln Pro Phe Asp Lys Asp Leu Ser Ala Arg Leu
          85           90           95

Gln Arg Glu Val Arg Leu Ala Asn Asp Ala Asn Cys Leu Ala Val Ser
          100          105          110

Glu Ala Val Asp Gly Ala Ala Ala Gly Ala Gln Thr Val Phe Ala Val
115          120          125

Ile Ile Gly Thr Gly Cys Gly Ala Gly Val Ala Phe Asn Gly Arg Ala
130          135          140

His Ile Gly Gly Asn Gly Thr Ala Gly Glu Trp Gly His Asn Pro Leu
145          150          155          160

Pro Trp Met Asp Glu Asp Glu Leu Arg Tyr Arg Glu Glu Val Pro Cys
          165          170          175

Tyr Cys Gly Lys Gln Gly Cys Ile Glu Thr Phe Ile Ser Gly Thr Gly
          180          185          190

Phe Ala Met Asp Tyr Arg Arg Leu Ser Gly His Ala Leu Lys Gly Ser
          195          200          205

Glu Ile Ile Arg Leu Val Glu Glu Ser Asp Pro Val Ala Glu Leu Ala
210          215          220

```

Leu Arg Arg Tyr Glu Leu Arg Leu Ala Lys Ser Leu Ala His Val Val
225 230 235 240

Asn Ile Leu Asp Pro Asp Val Ile Val Leu Gly Gly Gly Met Ser Asn
245 250 255

Val Asp Arg Leu Tyr Gln Thr Val Gly Gln Leu Ile Lys Gln Phe Val
260 265 270

Phe Gly Gly Glu Cys Glu Thr Pro Val Arg Lys Ala Lys His Gly Asp
275 280 285

Ser Ser Gly Val Arg Gly Ala Ala Trp Leu Trp Pro Gln Glu
290 295 300

<210> 4
<211> 440
<212> PRT
<213> *Escherichia coli*

<400> 4

Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg Tyr Glu Gly Ser
1 5 10 15

Lys Ser Ser Asn Pro Leu Ala Phe Arg His Tyr Asn Pro Asp Glu Leu
20 25 30

Val Leu Gly Lys Arg Met Glu Glu His Leu Arg Phe Ala Ala Cys Tyr
35 40 45

Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe Gly Val Gly Ala
50 55 60

Phe Asn Arg Pro Trp Gln Gln Pro Gly Glu Ala Leu Ala Leu Ala Lys
65 70 75 80

Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His Lys Leu His Val Pro
85 90 95

Phe Tyr Cys Phe His Asp Val Asp Val Ser Pro Glu Gly Ala Ser Leu
100 105 110

Lys Glu Tyr Ile Asn Asn Phe Ala Gln Met Val Asp Val Leu Ala Gly
115 120 125

Lys Gln Glu Glu Ser Gly Val Lys Leu Leu Trp Gly Thr Ala Asn Cys
130 135 140

5

10

ES 2 696 235 T3

Phe	Thr	Asn	Pro	Arg	Tyr	Gly	Ala	Gly	Ala	Ala	Thr	Asn	Pro	Asp	Pro	145	150	155	160
Glu	Val	Phe	Ser	Trp	Ala	Ala	Thr	Gln	Val	Val	Thr	Ala	Met	Glu	Ala	165	170	175	
Thr	His	Lys	Leu	Gly	Gly	Glu	Asn	Tyr	Val	Leu	Trp	Gly	Gly	Arg	Glu	180	185	190	
Gly	Tyr	Glu	Thr	Leu	Leu	Asn	Thr	Asp	Leu	Arg	Gln	Glu	Arg	Glu	Gln	195	200	205	
Leu	Gly	Arg	Phe	Met	Gln	Met	Val	Val	Glu	His	Lys	His	Lys	Ile	Gly	210	215	220	
Phe	Gln	Gly	Thr	Leu	Leu	Ile	Glu	Pro	Lys	Pro	Gln	Glu	Pro	Thr	Lys	225	230	235	240
His	Gln	Tyr	Asp	Tyr	Asp	Ala	Ala	Thr	Val	Tyr	Gly	Phe	Leu	Lys	Gln	245	250	255	
Phe	Gly	Leu	Glu	Lys	Glu	Ile	Lys	Leu	Asn	Ile	Glu	Ala	Asn	His	Ala	260	265	270	
Thr	Leu	Ala	Gly	His	Ser	Phe	His	His	Glu	Ile	Ala	Thr	Ala	Ile	Ala	275	280	285	
Leu	Gly	Leu	Phe	Gly	Ser	Val	Asp	Ala	Asn	Arg	Gly	Asp	Ala	Gln	Leu	290	295	300	
Gly	Trp	Asp	Thr	Asp	Gln	Phe	Pro	Asn	Ser	Val	Glu	Glu	Asn	Ala	Leu	305	310	315	320
Val	Met	Tyr	Glu	Ile	Leu	Lys	Ala	Gly	Gly	Phe	Thr	Thr	Gly	Gly	Leu	325	330	335	
Asn	Phe	Asp	Ala	Lys	Val	Arg	Arg	Gln	Ser	Thr	Asp	Lys	Tyr	Asp	Leu	340	345	350	
Phe	Tyr	Gly	His	Ile	Gly	Ala	Met	Asp	Thr	Met	Ala	Leu	Ala	Leu	Lys	355	360	365	
Ile	Ala	Ala	Arg	Met	Ile	Glu	Asp	Gly	Glu	Leu	Asp	Lys	Arg	Ile	Ala	370	375	380	
Gln	Arg	Tyr	Ser	Gly	Trp	Asn	Ser	Glu	Leu	Gly	Gln	Gln	Ile	Leu	Lys	385	390	395	400

Gly Gln Met Ser Leu Ala Asp Leu Ala Lys Tyr Ala Gln Glu His His
 405 410 415

Leu Ser Pro Val His Gln Ser Gly Arg Gln Glu Gln Leu Glu Asn Leu
 420 425 430

Val Asn His Tyr Leu Phe Asp Lys
 435 440

<210> 5
 <211> 484
 <212> PRT
 <213> *Escherichia coli*

<400> 5

Met Tyr Ile Gly Ile Asp Leu Gly Thr Ser Gly Val Lys Val Ile Leu
 1 5 10 15

Leu Asn Glu Gln Gly Glu Val Val Ala Ala Gln Thr Glu Lys Leu Thr
 20 25 30

Val Ser Arg Pro His Pro Leu Trp Ser Glu Gln Asp Pro Glu Gln Trp
 35 40 45

Trp Gln Ala Thr Asp Arg Ala Met Lys Ala Leu Gly Asp Gln His Ser
 50 55 60

Leu Gln Asp Val Lys Ala Leu Gly Ile Ala Gly Gln Met His Gly Ala
 65 70 75 80

Thr Leu Leu Asp Ala Gln Gln Arg Val Leu Arg Pro Ala Ile Leu Trp
 85 90 95

Asn Asp Gly Arg Cys Ala Gln Glu Cys Thr Leu Leu Glu Ala Arg Val
 100 105 110

Pro Gln Ser Arg Val Ile Thr Gly Asn Leu Met Met Pro Gly Phe Thr
 115 120 125

Ala Pro Lys Leu Leu Trp Val Gln Arg His Glu Pro Glu Ile Phe Arg
 130 135 140

Gln Ile Asp Lys Val Leu Leu Pro Lys Asp Tyr Leu Arg Leu Arg Met
 145 150 155 160

Thr Gly Glu Phe Ala Ser Asp Met Ser Asp Ala Ala Gly Thr Met Trp
 165 170 175

5

10

ES 2 696 235 T3

Leu Asp Val Ala Lys Arg Asp Trp Ser Asp Val Met Leu Gln Ala Cys
 180 185 190
 Asp Leu Ser Arg Asp Gln Met Pro Ala Leu Tyr Glu Gly Ser Glu Ile
 195 200 205
 Thr Gly Ala Leu Leu Pro Glu Val Ala Lys Ala Trp Gly Met Ala Thr
 210 215 220
 Val Pro Val Val Ala Gly Gly Gly Asp Asn Ala Ala Gly Ala Val Gly
 225 230 235 240
 Val Gly Met Val Asp Ala Asn Gln Ala Met Leu Ser Leu Gly Thr Ser
 245 250 255
 Gly Val Tyr Phe Ala Val Ser Glu Gly Phe Leu Ser Lys Pro Glu Ser
 260 265 270
 Ala Val His Ser Phe Cys His Ala Leu Pro Gln Arg Trp His Leu Met
 275 280 285
 Ser Val Met Leu Ser Ala Ala Ser Cys Leu Asp Trp Ala Ala Lys Leu
 290 295 300
 Thr Gly Leu Ser Asn Val Pro Ala Leu Ile Ala Ala Ala Gln Gln Ala
 305 310 315 320
 Asp Glu Ser Ala Glu Pro Val Trp Phe Leu Pro Tyr Leu Ser Gly Glu
 325 330 335
 Arg Thr Pro His Asn Asn Pro Gln Ala Lys Gly Val Phe Phe Gly Leu
 340 345 350
 Thr His Gln His Gly Pro Asn Glu Leu Ala Arg Ala Val Leu Glu Gly
 355 360 365
 Val Gly Tyr Ala Leu Ala Asp Gly Met Asp Val Val His Ala Cys Gly
 370 375 380
 Ile Lys Pro Gln Ser Val Thr Leu Ile Gly Gly Gly Ala Arg Ser Glu
 385 390 395 400
 Tyr Trp Arg Gln Met Leu Ala Asp Ile Ser Gly Gln Gln Leu Asp Tyr
 405 410 415
 Arg Thr Gly Gly Asp Val Gly Pro Ala Leu Gly Ala Ala Arg Leu Ala
 420 425 430

ES 2 696 235 T3

Gln Ile Ala Ala Asn Pro Glu Lys Ser Leu Ile Glu Leu Leu Pro Gln
435 440 445

Leu Pro Leu Glu Gln Ser His Leu Pro Asp Ala Gln Arg Tyr Ala Ala
450 455 460

Tyr Gln Pro Arg Arg Glu Thr Phe Arg Arg Leu Tyr Gln Gln Leu Leu
465 470 475 480

Pro Leu Met Ala

<210> 6

<211> 491

<212> PRT

<213> *Escherichia coli*

<400> 6

Met Asn Thr Gln Tyr Asn Ser Ser Tyr Ile Phe Ser Ile Thr Leu Val
1 5 10 15

Ala Thr Leu Gly Gly Leu Leu Phe Gly Tyr Asp Thr Ala Val Ile Ser
20 25 30

Gly Thr Val Glu Ser Leu Asn Thr Val Phe Val Ala Pro Gln Asn Leu
35 40 45

Ser Glu Ser Ala Ala Asn Ser Leu Leu Gly Phe Cys Val Ala Ser Ala
50 55 60

Leu Ile Gly Cys Ile Ile Gly Gly Ala Leu Gly Gly Tyr Cys Ser Asn
65 70 75 80

Arg Phe Gly Arg Arg Asp Ser Leu Lys Ile Ala Ala Val Leu Phe Phe
85 90 95

Ile Ser Gly Val Gly Ser Ala Trp Pro Glu Leu Gly Phe Thr Ser Ile
100 105 110

Asn Pro Asp Asn Thr Val Pro Val Tyr Leu Ala Gly Tyr Val Pro Glu
115 120 125

Phe Val Ile Tyr Arg Ile Ile Gly Gly Ile Gly Val Gly Leu Ala Ser
130 135 140

Met Leu Ser Pro Met Tyr Ile Ala Glu Leu Ala Pro Ala His Ile Arg
145 150 155 160

5

10

Gly Lys Leu Val Ser Phe Asn Gln Phe Ala Ile Ile Phe Gly Gln Leu
 165 170 175

Leu Val Tyr Cys Val Asn Tyr Phe Ile Ala Arg Ser Gly Asp Ala Ser
 180 185 190

Trp Leu Asn Thr Asp Gly Trp Arg Tyr Met Phe Ala Ser Glu Cys Ile
 195 200 205

Pro Ala Leu Leu Phe Leu Met Leu Leu Tyr Thr Val Pro Glu Ser Pro
 210 215 220

Arg Trp Leu Met Ser Arg Gly Lys Gln Glu Gln Ala Glu Gly Ile Leu
 225 230 235 240

Arg Lys Ile Met Gly Asn Thr Leu Ala Thr Gln Ala Val Gln Glu Ile
 245 250 255

Lys His Ser Leu Asp His Gly Arg Lys Thr Gly Gly Arg Leu Leu Met
 260 265 270

Phe Gly Val Gly Val Ile Val Ile Gly Val Met Leu Ser Ile Phe Gln
 275 280 285

Gln Phe Val Gly Ile Asn Val Val Leu Tyr Tyr Ala Pro Glu Val Phe
 290 295 300

Lys Thr Leu Gly Ala Ser Thr Asp Ile Ala Leu Leu Gln Thr Ile Ile
 305 310 315 320

Val Gly Val Ile Asn Leu Thr Phe Thr Val Leu Ala Ile Met Thr Val
 325 330 335

Asp Lys Phe Gly Arg Lys Pro Leu Gln Ile Ile Gly Ala Leu Gly Met
 340 345 350

Ala Ile Gly Met Phe Ser Leu Gly Thr Ala Phe Tyr Thr Gln Ala Pro
 355 360 365

Gly Ile Val Ala Leu Leu Ser Met Leu Phe Tyr Val Ala Ala Phe Ala
 370 375 380

Met Ser Trp Gly Pro Val Cys Trp Val Leu Leu Ser Glu Ile Phe Pro
 385 390 395 400

Asn Ala Ile Arg Gly Lys Ala Leu Ala Ile Ala Val Ala Ala Gln Trp

405	410	415
Leu Ala Asn Tyr Phe Val Ser Trp Thr Phe Pro Met Met Asp Lys Asn		
420	425	430
Ser Trp Leu Val Ala His Phe His Asn Gly Phe Ser Tyr Trp Ile Tyr		
435	440	445
Gly Cys Met Gly Val Leu Ala Ala Leu Phe Met Trp Lys Phe Val Pro		
450	455	460
Glu Thr Lys Gly Lys Thr Leu Glu Glu Leu Glu Ala Leu Trp Glu Pro		
465	470	475
Glu Thr Lys Lys Thr Gln Gln Thr Ala Thr Leu		
485	490	

<210> 7
 <211> 330
 <212> PRT
 <213> *Escherichia coli*

<400> 7

Met Lys Ile Lys Asn Ile Leu Leu Thr Leu Cys Thr Ser Leu Leu Leu		
1	5	10
Thr Asn Val Ala Ala His Ala Lys Glu Val Lys Ile Gly Met Ala Ile		
20	25	30
Asp Asp Leu Arg Leu Glu Arg Trp Gln Lys Asp Arg Asp Ile Phe Val		
35	40	45
Lys Lys Ala Glu Ser Leu Gly Ala Lys Val Phe Val Gln Ser Ala Asn		
50	55	60
Gly Asn Glu Glu Thr Gln Met Ser Gln Ile Glu Asn Met Ile Asn Arg		
65	70	75
Gly Val Asp Val Leu Val Ile Ile Pro Tyr Asn Gly Gln Val Leu Ser		
85	90	95
Asn Val Val Lys Glu Ala Lys Gln Glu Gly Ile Lys Val Leu Ala Tyr		
100	105	110
Asp Arg Met Ile Asn Asp Ala Asp Ile Asp Phe Tyr Ile Ser Phe Asp		
115	120	125
Asn Glu Lys Val Gly Glu Leu Gln Ala Lys Ala Leu Val Asp Ile Val		

ES 2 696 235 T3

130	135	140
Pro Gln Gly Asn Tyr Phe Leu Met Gly Gly Ser Pro Val Asp Asn Asn 145 150 155 160		
Ala Lys Leu Phe Arg Ala Gly Gln Met Lys Val Leu Lys Pro Tyr Val 165 170 175		
Asp Ser Gly Lys Ile Lys Val Val Gly Asp Gln Trp Val Asp Gly Trp 180 185 190		
Leu Pro Glu Asn Ala Leu Lys Ile Met Glu Asn Ala Leu Thr Ala Asn 195 200 205		
Asn Asn Lys Ile Asp Ala Val Val Ala Ser Asn Asp Ala Thr Ala Gly 210 215 220		
Gly Ala Ile Gln Ala Leu Ser Ala Gln Gly Leu Ser Gly Lys Val Ala 225 230 235 240		
Ile Ser Gly Gln Asp Ala Asp Leu Ala Gly Ile Lys Arg Ile Ala Ala 245 250 255		
Gly Thr Gln Thr Met Thr Val Tyr Lys Pro Ile Thr Leu Leu Ala Asn 260 265 270		
Thr Ala Ala Glu Ile Ala Val Glu Leu Gly Asn Gly Gln Glu Pro Lys 275 280 285		
Ala Asp Thr Thr Leu Asn Asn Gly Leu Lys Asp Val Pro Ser Arg Leu 290 295 300		
Leu Thr Pro Ile Asp Val Asn Lys Asn Asn Ile Lys Asp Thr Val Ile 305 310 315 320		
Lys Asp Gly Phe His Lys Glu Ser Glu Leu 325 330		

<210> 8
 <211> 513
 <212> PRT
 <213> *Escherichia coli*

<400> 8

Met Pro Tyr Leu Leu Glu Met Lys Asn Ile Thr Lys Thr Phe Gly Ser 1 5 10 15
Val Lys Ala Ile Asp Asn Val Cys Leu Arg Leu Asn Ala Gly Glu Ile

ES 2 696 235 T3

20					25					30					
Val	Ser	Leu	Cys	Gly	Glu	Asn	Gly	Ser	Gly	Lys	Ser	Thr	Leu	Met	Lys
		35					40					45			
Val	Leu	Cys	Gly	Ile	Tyr	Pro	His	Gly	Ser	Tyr	Glu	Gly	Glu	Ile	Ile
	50					55					60				
Phe	Ala	Gly	Glu	Glu	Ile	Gln	Ala	Ser	His	Ile	Arg	Asp	Thr	Glu	Arg
65					70					75					80
Lys	Gly	Ile	Ala	Ile	Ile	His	Gln	Glu	Leu	Ala	Leu	Val	Lys	Glu	Leu
				85					90					95	
Thr	Val	Leu	Glu	Asn	Ile	Phe	Leu	Gly	Asn	Glu	Ile	Thr	His	Asn	Gly
			100					105					110		
Ile	Met	Asp	Tyr	Asp	Leu	Met	Thr	Leu	Arg	Cys	Gln	Lys	Leu	Leu	Ala
		115					120					125			
Gln	Val	Ser	Leu	Ser	Ile	Ser	Pro	Asp	Thr	Arg	Val	Gly	Asp	Leu	Gly
	130					135					140				
Leu	Gly	Gln	Gln	Gln	Leu	Val	Glu	Ile	Ala	Lys	Ala	Leu	Asn	Lys	Gln
145					150					155					160
Val	Arg	Leu	Leu	Ile	Leu	Asp	Glu	Pro	Thr	Ala	Ser	Leu	Thr	Glu	Gln
				165					170					175	
Glu	Thr	Ser	Ile	Leu	Leu	Asp	Ile	Ile	Arg	Asp	Leu	Gln	Gln	His	Gly
			180					185					190		
Ile	Ala	Cys	Ile	Tyr	Ile	Ser	His	Lys	Leu	Asn	Glu	Val	Lys	Ala	Ile
		195					200					205			
Ser	Asp	Thr	Ile	Cys	Val	Ile	Arg	Asp	Gly	Gln	His	Ile	Gly	Thr	Arg
	210					215					220				
Asp	Ala	Ala	Gly	Met	Ser	Glu	Asp	Asp	Ile	Ile	Thr	Met	Met	Val	Gly
225					230					235					240
Arg	Glu	Leu	Thr	Ala	Leu	Tyr	Pro	Asn	Glu	Pro	His	Thr	Thr	Gly	Asp
				245					250					255	
Glu	Ile	Leu	Arg	Ile	Glu	His	Leu	Thr	Ala	Trp	His	Pro	Val	Asn	Arg
			260					265					270		

ES 2 696 235 T3

His Ile Lys Arg Val Asn Asp Val Ser Phe Ser Leu Lys Arg Gly Glu
275 280 285

Ile Leu Gly Ile Ala Gly Leu Val Gly Ala Gly Arg Thr Glu Thr Ile
290 295 300

Gln Cys Leu Phe Gly Val Trp Pro Gly Gln Trp Glu Gly Lys Ile Tyr
305 310 315 320

Ile Asp Gly Lys Gln Val Asp Ile Arg Asn Cys Gln Gln Ala Ile Ala
325 330 335

Gln Gly Ile Ala Met Val Pro Glu Asp Arg Lys Arg Asp Gly Ile Val
340 345 350

Pro Val Met Ala Val Gly Lys Asn Ile Thr Leu Ala Ala Leu Asn Lys
355 360 365

Phe Thr Gly Gly Ile Ser Gln Leu Asp Asp Ala Ala Glu Gln Lys Cys
370 375 380

Ile Leu Glu Ser Ile Gln Gln Leu Lys Val Lys Thr Ser Ser Pro Asp
385 390 395 400

Leu Ala Ile Gly Arg Leu Ser Gly Gly Asn Gln Gln Lys Ala Ile Leu
405 410 415

Ala Arg Cys Leu Leu Leu Asn Pro Arg Ile Leu Ile Leu Asp Glu Pro
420 425 430

Thr Arg Gly Ile Asp Ile Gly Ala Lys Tyr Glu Ile Tyr Lys Leu Ile
435 440 445

Asn Gln Leu Val Gln Gln Gly Ile Ala Val Ile Val Ile Ser Ser Glu
450 455 460

Leu Pro Glu Val Leu Gly Leu Ser Asp Arg Val Leu Val Met His Glu
465 470 475 480

Gly Lys Leu Lys Ala Asn Leu Ile Asn His Asn Leu Thr Gln Glu Gln
485 490 495

Val Met Glu Ala Ala Leu Arg Ser Glu His His Val Glu Lys Gln Ser
500 505 510

Val

<210> 9
 <211> 393
 <212> PRT
 <213> *Escherichia coli*

5

<400> 9

```

Met Ser Lys Ser Asn Pro Ser Glu Val Lys Leu Ala Val Pro Thr Ser
1          5          10          15

Gly Gly Phe Ser Gly Leu Lys Ser Leu Asn Leu Gln Val Phe Val Met
20          25          30

Ile Ala Ala Ile Ile Ala Ile Met Leu Phe Phe Thr Trp Thr Thr Asp
35          40          45

Gly Ala Tyr Leu Ser Ala Arg Asn Val Ser Asn Leu Leu Arg Gln Thr
50          55          60

Ala Ile Thr Gly Ile Leu Ala Val Gly Met Val Phe Val Ile Ile Ser
65          70          75          80

Ala Glu Ile Asp Leu Ser Val Gly Ser Met Met Gly Leu Leu Gly Gly
85          90          95

Val Ala Ala Ile Cys Asp Val Trp Leu Gly Trp Pro Leu Pro Leu Thr
100          105          110

Ile Ile Val Thr Leu Val Leu Gly Leu Leu Leu Gly Ala Trp Asn Gly
115          120          125

Trp Trp Val Ala Tyr Arg Lys Val Pro Ser Phe Ile Val Thr Leu Ala
130          135          140

Gly Met Leu Ala Phe Arg Gly Ile Leu Ile Gly Ile Thr Asn Gly Thr
145          150          155          160

Thr Val Ser Pro Thr Ser Ala Ala Met Ser Gln Ile Gly Gln Ser Tyr
165          170          175

Leu Pro Ala Ser Thr Gly Phe Ile Ile Gly Ala Leu Gly Leu Met Ala
180          185          190

Phe Val Gly Trp Gln Trp Arg Gly Arg Met Arg Arg Gln Ala Leu Gly
195          200          205

Leu Gln Ser Pro Ala Ser Thr Ala Val Val Gly Arg Gln Ala Leu Thr
210          215          220

```

Ala Ile Ile Val Leu Gly Ala Ile Trp Leu Leu Asn Asp Tyr Arg Gly
225 230 235 240

Val Pro Thr Pro Val Leu Leu Leu Thr Leu Leu Leu Leu Gly Gly Met
245 250 255

Phe Met Ala Thr Arg Thr Ala Phe Gly Arg Arg Ile Tyr Ala Ile Gly
260 265 270

Gly Asn Leu Glu Ala Ala Arg Leu Ser Gly Ile Asn Val Glu Arg Thr
275 280 285

Lys Leu Ala Val Phe Ala Ile Asn Gly Leu Met Val Ala Ile Ala Gly
290 295 300

Leu Ile Leu Ser Ser Arg Leu Gly Ala Gly Ser Pro Ser Ala Gly Asn
305 310 315 320

Ile Ala Glu Leu Asp Ala Ile Ala Ala Cys Val Ile Gly Gly Thr Ser
325 330 335

Leu Ala Gly Gly Val Gly Ser Val Ala Gly Ala Val Met Gly Ala Phe
340 345 350

Ile Met Ala Ser Leu Asp Asn Gly Met Ser Met Met Asp Val Pro Thr
355 360 365

Phe Trp Gln Tyr Ile Val Lys Gly Ala Ile Leu Leu Leu Ala Val Trp
370 375 380

Met Asp Ser Ala Thr Lys Arg Arg Ser
385 390

<210> 10
<211> 317
<212> PRT
<213> *Escherichia coli*

<400> 10

Met Thr Asp Lys Leu Thr Ser Leu Arg Gln Tyr Thr Thr Val Val Ala
1 5 10 15

Asp Thr Gly Asp Ile Ala Ala Met Lys Leu Tyr Gln Pro Gln Asp Ala
20 25 30

Thr Thr Asn Pro Ser Leu Ile Leu Asn Ala Ala Gln Ile Pro Glu Tyr
35 40 45

5

10

ES 2 696 235 T3

Arg	Lys	Leu	Ile	Asp	Asp	Ala	Val	Ala	Trp	Ala	Lys	Gln	Gln	Ser	Asn	50	55	60
Asp	Arg	Ala	Gln	Gln	Ile	Val	Asp	Ala	Thr	Asp	Lys	Leu	Ala	Val	Asn	65	70	75
Ile	Gly	Leu	Glu	Ile	Leu	Lys	Leu	Val	Pro	Gly	Arg	Ile	Ser	Thr	Glu	85	90	95
Val	Asp	Ala	Arg	Leu	Ser	Tyr	Asp	Thr	Glu	Ala	Ser	Ile	Ala	Lys	Ala	100	105	110
Lys	Arg	Leu	Ile	Lys	Leu	Tyr	Asn	Asp	Ala	Gly	Ile	Ser	Asn	Asp	Arg	115	120	125
Ile	Leu	Ile	Lys	Leu	Ala	Ser	Thr	Trp	Gln	Gly	Ile	Arg	Ala	Ala	Glu	130	135	140
Gln	Leu	Glu	Lys	Glu	Gly	Ile	Asn	Cys	Asn	Leu	Thr	Leu	Leu	Phe	Ser	145	150	155
Phe	Ala	Gln	Ala	Arg	Ala	Cys	Ala	Glu	Ala	Gly	Val	Phe	Leu	Ile	Ser	165	170	175
Pro	Phe	Val	Gly	Arg	Ile	Leu	Asp	Trp	Tyr	Lys	Ala	Asn	Thr	Asp	Lys	180	185	190
Lys	Glu	Tyr	Ala	Pro	Ala	Glu	Asp	Pro	Gly	Val	Val	Ser	Val	Ser	Glu	195	200	205
Ile	Tyr	Gln	Tyr	Tyr	Lys	Glu	His	Gly	Tyr	Glu	Thr	Val	Val	Met	Gly	210	215	220
Ala	Ser	Phe	Arg	Asn	Ile	Gly	Glu	Ile	Leu	Glu	Leu	Ala	Gly	Cys	Asp	225	230	235
Arg	Leu	Thr	Ile	Ala	Pro	Ala	Leu	Leu	Lys	Glu	Leu	Ala	Glu	Ser	Glu	245	250	255
Gly	Ala	Ile	Glu	Arg	Lys	Leu	Ser	Tyr	Thr	Gly	Glu	Val	Lys	Ala	Arg	260	265	270
Pro	Ala	Arg	Ile	Thr	Glu	Ser	Glu	Phe	Leu	Trp	Gln	His	Asn	Gln	Asp	275	280	285
Pro	Met	Ala	Val	Asp	Lys	Leu	Ala	Glu	Gly	Ile	Arg	Lys	Phe	Ala	Ile	290	295	300

Asp Gln Glu Lys Leu Glu Lys Met Ile Gly Asp Leu Leu
 305 310 315

<210> 11
 <211> 663
 <212> PRT
 <213> *Escherichia coli*

<400> 11

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
 1 5 10 15

Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
 20 25 30

Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
 35 40 45

Pro Gln Asn Pro Ser Trp Ala Asp Arg Asp Arg Phe Val Leu Ser Asn
 50 55 60

Gly His Gly Ser Met Leu Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr
 65 70 75 80

Asp Leu Pro Met Glu Glu Leu Lys Asn Phe Arg Gln Leu His Ser Lys
 85 90 95

Thr Pro Gly His Pro Glu Val Gly Tyr Thr Ala Gly Val Glu Thr Thr
 100 105 110

Thr Gly Pro Leu Gly Gln Gly Ile Ala Asn Ala Val Gly Met Ala Ile
 115 120 125

Ala Glu Lys Thr Leu Ala Ala Gln Phe Asn Arg Pro Gly His Asp Ile
 130 135 140

Val Asp His Tyr Thr Tyr Ala Phe Met Gly Asp Gly Cys Met Met Glu
 145 150 155 160

Gly Ile Ser His Glu Val Cys Ser Leu Ala Gly Thr Leu Lys Leu Gly
 165 170 175

Lys Leu Ile Ala Phe Tyr Asp Asp Asn Gly Ile Ser Ile Asp Gly His
 180 185 190

Val Glu Gly Trp Phe Thr Asp Asp Thr Ala Met Arg Phe Glu Ala Tyr
 195 200 205

ES 2 696 235 T3

Gly 210	Trp	His	Val	Ile	Arg	Asp 215	Ile	Asp	Gly	His	Asp 220	Ala	Ala	Ser	Ile
Lys 225	Arg	Ala	Val	Glu	Glu	Ala	Arg	Ala	Val	Thr	Asp 235	Lys	Pro	Ser	Leu 240
Leu	Met	Cys	Lys	Thr 245	Ile	Ile	Gly	Phe	Gly 250	Ser	Pro	Asn	Lys	Ala 255	Gly
Thr	His	Asp	Ser 260	His	Gly	Ala	Pro	Leu	Gly 265	Asp	Ala	Glu	Ile 270	Ala	Leu
Thr	Arg	Glu 275	Gln	Leu	Gly	Trp	Lys 280	Tyr	Ala	Pro	Phe	Glu 285	Ile	Pro	Ser
Glu 290	Ile	Tyr	Ala	Gln	Trp	Asp 295	Ala	Lys	Glu	Ala	Gly 300	Gln	Ala	Lys	Glu
Ser 305	Ala	Trp	Asn	Glu	Lys 310	Phe	Ala	Ala	Tyr	Ala 315	Lys	Ala	Tyr	Pro	Gln 320
Glu	Ala	Ala	Glu	Phe 325	Thr	Arg	Arg	Met	Lys 330	Gly	Glu	Met	Pro	Ser 335	Asp
Phe	Asp	Ala	Lys 340	Ala	Lys	Glu	Phe	Ile 345	Ala	Lys	Leu	Gln	Ala 350	Asn	Pro
Ala	Lys	Ile 355	Ala	Ser	Arg	Lys	Ala 360	Ser	Gln	Asn	Ala	Ile 365	Glu	Ala	Phe
Gly 370	Pro	Leu	Leu	Pro	Glu	Phe 375	Leu	Gly	Gly	Ser	Ala 380	Asp	Leu	Ala	Pro
Ser 385	Asn	Leu	Thr	Leu	Trp 390	Ser	Gly	Ser	Lys	Ala 395	Ile	Asn	Glu	Asp	Ala 400
Ala	Gly	Asn	Tyr	Ile 405	His	Tyr	Gly	Val	Arg 410	Glu	Phe	Gly	Met	Thr 415	Ala
Ile	Ala	Asn	Gly 420	Ile	Ser	Leu	His	Gly 425	Gly	Phe	Leu	Pro	Tyr 430	Thr	Ser
Thr	Phe	Leu 435	Met	Phe	Val	Glu	Tyr 440	Ala	Arg	Asn	Ala 445	Val	Arg	Met	Ala
Ala	Leu	Met	Lys	Gln	Arg	Gln	Val	Met	Val	Tyr	Thr	His	Asp	Ser	Ile

450		455		460
Gly Leu Gly Glu Asp	Gly Pro Thr His Gln	Pro Val Glu Gln Val Ala		
465	470	475		480
Ser Leu Arg Val Thr	Pro Asn Met Ser Thr	Trp Arg Pro Cys Asp Gln		
	485	490		495
Val Glu Ser Ala Val	Ala Trp Lys Tyr Gly	Val Glu Arg Gln Asp Gly		
	500	505		510
Pro Thr Ala Leu Ile	Leu Ser Arg Gln Asn	Leu Ala Gln Gln Glu Arg		
	515	520		525
Thr Glu Glu Gln Leu	Ala Asn Ile Ala Arg	Gly Gly Tyr Val Leu Lys		
	530	535		540
Asp Cys Ala Gly Gln	Pro Glu Leu Ile Phe	Ile Ala Thr Gly Ser Glu		
545	550	555		560
Val Glu Leu Ala Val	Ala Ala Tyr Glu Lys	Leu Thr Ala Glu Gly Val		
	565	570		575
Lys Ala Arg Val Val	Ser Met Pro Ser Thr	Asp Ala Phe Asp Lys Gln		
	580	585		590
Asp Ala Ala Tyr Arg	Glu Ser Val Leu Pro	Lys Ala Val Thr Ala Arg		
	595	600		605
Val Ala Val Glu Ala	Gly Ile Ala Asp Tyr	Trp Tyr Lys Tyr Val Gly		
	610	615		620
Leu Asn Gly Ala Ile	Val Gly Met Thr Thr	Phe Gly Glu Ser Ala Pro		
625	630	635		640
Ala Glu Leu Leu Phe	Glu Glu Phe Gly Phe	Thr Val Asp Asn Val Val		
	645	650		655
Ala Lys Ala Lys Glu	Leu Leu			
	660			

<210> 12
 <211> 321
 <212> PRT
 <213> *Escherichia coli*

5

<400> 12

Met Thr Lys Tyr Ala Leu Val Gly Asp Val Gly Gly Thr Asn Ala Arg

ES 2 696 235 T3

1		5						10					15		
Leu	Ala	Leu	Cys	Asp	Ile	Ala	Ser	Gly	Glu	Ile	Ser	Gln	Ala	Lys	Thr
		20						25					30		
Tyr	Ser	Gly	Leu	Asp	Tyr	Pro	Ser	Leu	Glu	Ala	Val	Ile	Arg	Val	Tyr
		35					40					45			
Leu	Glu	Glu	His	Lys	Val	Glu	Val	Lys	Asp	Gly	Cys	Ile	Ala	Ile	Ala
	50					55					60				
Cys	Pro	Ile	Thr	Gly	Asp	Trp	Val	Ala	Met	Thr	Asn	His	Thr	Trp	Ala
65					70					75				80	
Phe	Ser	Ile	Ala	Glu	Met	Lys	Lys	Asn	Leu	Gly	Phe	Ser	His	Leu	Glu
				85					90					95	
Ile	Ile	Asn	Asp	Phe	Thr	Ala	Val	Ser	Met	Ala	Ile	Pro	Met	Leu	Lys
			100					105					110		
Lys	Glu	His	Leu	Ile	Gln	Phe	Gly	Gly	Ala	Glu	Pro	Val	Glu	Gly	Lys
		115					120					125			
Pro	Ile	Ala	Val	Tyr	Gly	Ala	Gly	Thr	Gly	Leu	Gly	Val	Ala	His	Leu
	130					135					140				
Val	His	Val	Asp	Lys	Arg	Trp	Val	Ser	Leu	Pro	Gly	Glu	Gly	Gly	His
145					150					155					160
Val	Asp	Phe	Ala	Pro	Asn	Ser	Glu	Glu	Glu	Ala	Ile	Ile	Leu	Glu	Ile
				165				170						175	
Leu	Arg	Ala	Glu	Ile	Gly	His	Val	Ser	Ala	Glu	Arg	Val	Leu	Ser	Gly
			180					185					190		
Pro	Gly	Leu	Val	Asn	Leu	Tyr	Arg	Ala	Ile	Val	Lys	Ala	Asp	Asn	Arg
		195					200					205			
Leu	Pro	Glu	Asn	Leu	Lys	Pro	Lys	Asp	Ile	Thr	Glu	Arg	Ala	Leu	Ala
	210					215					220				
Asp	Ser	Cys	Thr	Asp	Cys	Arg	Arg	Ala	Leu	Ser	Leu	Phe	Cys	Val	Ile
225					230					235					240
Met	Gly	Arg	Phe	Gly	Gly	Asn	Leu	Ala	Leu	Asn	Leu	Gly	Thr	Phe	Gly
				245					250					255	

Gly Val Phe Ile Ala Gly Gly Ile Val Pro Arg Phe Leu Glu Phe Phe
 260 265 270

Lys Ala Ser Gly Phe Arg Ala Ala Phe Glu Asp Lys Gly Arg Phe Lys
 275 280 285

Glu Tyr Val His Asp Ile Pro Val Tyr Leu Ile Val His Asp Asn Pro
 290 295 300

Gly Leu Leu Gly Ser Gly Ala His Leu Arg Gln Thr Leu Gly His Ile
 305 310 315 320

Leu

<210> 13

<211> 391

<212> PRT

<213> *Escherichia albertii*

<400> 13

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
 1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Met Glu Asn Pro Ser Asn Gln Pro
 20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Val
 35 40 45

His Asn Ala Ala Gly Glu Val Val Ser Leu Arg Asp Val Ile Glu Ser
 50 55 60

Asn Lys Ser Thr Leu Leu Gly Glu Ala Val Ala Gln Arg Phe Gly Glu
 65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
 85 90 95

Gln Val His Pro Asn Lys Arg Asn Ser Glu Ile Gly Phe Ala Lys Glu
 100 105 110

Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
 115 120 125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
 130 135 140

5

10

ES 2 696 235 T3

Met	Asn	Ala	Phe	Arg	Glu	Phe	Ser	Glu	Ile	Val	Ser	Leu	Leu	Gln	Pro	145	150	155				160
Val	Ala	Asp	Ala	His	Pro	Ala	Ile	Thr	His	Phe	Leu	Gln	Gln	Pro	Asn		165	170				175
Ala	Glu	Arg	Leu	Ser	Glu	Leu	Phe	Ala	Ser	Leu	Leu	Asn	Met	Gln	Gly		180	185				190
Glu	Glu	Lys	Ser	Arg	Ala	Leu	Ala	Ile	Leu	Lys	Ser	Ala	Leu	His	Ser		195	200				205
Gln	Gln	Gly	Glu	Pro	Trp	Gln	Thr	Ile	Arg	Leu	Ile	Ser	Glu	Phe	Tyr		210	215				220
Pro	Asp	Asp	Ser	Gly	Leu	Phe	Ser	Pro	Leu	Leu	Leu	Asn	Val	Val	Lys	225	230	235				240
Leu	Asn	Pro	Gly	Glu	Ala	Met	Phe	Leu	Phe	Ala	Glu	Thr	Pro	His	Ala		245	250				255
Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn	Ser	Asp	Asn	Val		260	265				270
Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Tyr	Ile	Asp	Ile	Pro	Glu	Leu	Val		275	280				285
Ala	Asn	Val	Lys	Phe	Glu	Ala	Lys	Pro	Ala	Asn	Gln	Leu	Leu	Thr	Gln	290	295	300				
Pro	Val	Lys	His	Gly	Ala	Glu	Leu	Asp	Phe	Pro	Ile	Pro	Val	Asp	Asp	305	310	315				320
Phe	Ala	Phe	Ser	Leu	His	Asp	Leu	Asn	Asp	Lys	Glu	Cys	Thr	Ile	Gly		325	330				335
Gln	Gln	Ser	Ala	Ala	Ile	Leu	Phe	Cys	Val	Glu	Gly	Glu	Ala	Thr	Leu		340	345				350
Asn	Lys	Asp	Ser	Gln	Gln	Leu	Leu	Leu	Lys	Pro	Gly	Glu	Ser	Ala	Phe		355	360				365
Ile	Ala	Ala	Asn	Glu	Ser	Pro	Val	Lys	Val	Lys	Gly	His	Gly	Arg	Leu	370	375	380				
Ala	Arg	Val	Tyr	Asn	Lys	Leu										385	390					

<210> 14
 <211> 302
 <212> PRT
 <213> *Escherichia blattae*

5

<400> 14

```

Met Arg Ile Gly Ile Asp Leu Gly Gly Thr Lys Thr Glu Val Ile Ala
1           5           10           15

Leu Gly Asp Ala Gly Glu Gln Leu Tyr Arg His Arg Leu Pro Thr Pro
20           25           30

Arg Asp Asp Tyr Arg Gln Thr Ile Glu Thr Ile Ala Thr Leu Val Asp
35           40           45

Met Ala Glu Gln Thr Thr Gly Gln Cys Gly Thr Val Gly Ile Gly Ile
50           55           60

Pro Gly Ser Ile Ser Pro Tyr Thr Gly Val Val Lys Asn Ala Asn Ser
65           70           75           80

Thr Trp Leu Asn Gly Gln Pro Phe Asp Lys Asp Leu Ser Ala Arg Leu
85           90           95

Gln Arg Glu Val Arg Leu Ala Asn Asp Ala Asn Cys Leu Ala Val Ser
100          105          110

Glu Ala Val Asp Gly Ala Ala Ala Gly Ala Lys Thr Val Phe Ala Val
115          120          125

Ile Ile Gly Thr Gly Cys Gly Ala Gly Val Ala Phe Asn Gly Arg Ala
130          135          140

His Ile Gly Gly Asn Gly Thr Ala Gly Glu Trp Gly His Asn Pro Leu
145          150          155          160

Pro Trp Met Asp Thr Asp Glu Leu Arg Tyr Arg Glu Glu Val Pro Cys
165          170          175

Tyr Cys Gly Lys Gln Gly Cys Ile Glu Thr Phe Ile Ser Gly Thr Gly
180          185          190

Phe Ala Thr Asp Tyr His Arg Leu Ser Gly His Val Leu Lys Gly Ser
195          200          205

Glu Ile Ile Arg Leu Val Glu Glu Asn Asp Pro Val Ala Glu Leu Ala
210          215          220

```

Leu Asp Arg Tyr Glu Arg Arg Leu Ala Lys Ser Leu Ala His Val Val
225 230 235 240

Asn Ile Leu Asp Pro Asp Val Ile Val Leu Gly Gly Gly Met Ser Asn
245 250 255

Val Glu Arg Leu Tyr Gln Thr Val Pro Pro Leu Ile Lys Arg Phe Val
260 265 270

Phe Gly Gly Glu Cys Glu Thr Pro Val Arg Lys Ala Lys His Gly Asp
275 280 285

Ser Ser Gly Val Arg Gly Ala Ala Trp Leu Trp Pro Gln Glu
290 295 300

<210> 15

<211> 411

<212> PRT

<213> *Escherichia fergusonii*

<400> 15

Met Pro Glu Asn Ser Cys Phe Asn Asn Ala Asn Lys Leu His Met Lys
1 5 10 15

Gln Gly Leu Ile Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala
20 25 30

Trp Gly Ser Lys Thr Ala Leu Thr Asp Leu Tyr Gly Ile Ala Asn Pro
35 40 45

Ser Gln Gln Pro Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser
50 55 60

Ser Ser Lys Val Glu Thr Ser Glu Gly Gln Ile Val Ser Leu Arg Asp
65 70 75 80

Val Ile Glu Asn Asp Lys Ser Ala Leu Leu Gly Asp Ala Val Ala Glu
85 90 95

Arg Phe Gly Glu Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln
100 105 110

Pro Leu Ser Ile Gln Val His Pro Asn Lys Arg Asn Ser Glu Ile Gly
115 120 125

Phe Ala Lys Glu Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg
130 135 140

ES 2 696 235 T3

Asn	Tyr	Lys	Asp	Pro	Asn	His	Lys	Pro	Glu	Leu	Val	Phe	Ala	Leu	Thr
145					150					155					160
Pro	Phe	Leu	Ala	Met	Asn	Ala	Phe	Arg	Glu	Phe	Ser	Glu	Ile	Val	Ser
				165					170					175	
Leu	Leu	Gln	Pro	Val	Ala	Gly	Ala	His	Pro	Ala	Ile	Ala	His	Phe	Leu
			180					185					190		
Gln	Gln	Pro	Asp	Ala	Glu	Arg	Leu	Ser	Glu	Leu	Phe	Ala	Ala	Leu	Leu
		195					200					205			
Asn	Met	Gln	Gly	Glu	Glu	Lys	Ser	Arg	Ala	Leu	Ala	Ile	Leu	Lys	Ser
	210					215					220				
Ala	Leu	Asp	Thr	Gln	Gln	Gly	Glu	Pro	Trp	Gln	Thr	Ile	Arg	Val	Ile
225					230					235					240
Ala	Gln	Phe	Tyr	Pro	Asp	Asp	Ser	Gly	Leu	Phe	Ser	Pro	Leu	Leu	Leu
				245					250					255	
Asn	Val	Val	Lys	Leu	Asn	Pro	Gly	Glu	Ala	Met	Phe	Leu	Phe	Ala	Glu
			260					265					270		
Thr	Pro	His	Ala	Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn
		275					280					285			
Ser	Asp	Asn	Val	Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Tyr	Ile	Asp	Ile
	290					295					300				
Pro	Glu	Leu	Val	Ala	Asn	Val	Lys	Phe	Glu	Ala	Lys	Pro	Ala	Gly	Gln
305					310					315					320
Leu	Leu	Thr	Gln	Pro	Ile	Lys	Asn	Gly	His	Glu	Leu	Asp	Phe	Pro	Ile
				325					330					335	
Pro	Val	Glu	Asp	Phe	Ala	Phe	Ala	Leu	His	Asp	Leu	Thr	Lys	Asp	Glu
			340					345					350		
Ala	Ala	Val	Asn	Gln	Gln	Ser	Ala	Ala	Ile	Leu	Phe	Cys	Val	Glu	Gly
		355					360					365			
Glu	Ala	Cys	Leu	Arg	Lys	Gly	Asp	Gln	Gln	Leu	His	Leu	Lys	Pro	Gly
	370					375					380				
Glu	Ser	Ala	Phe	Ile	Ala	Ala	Asn	Glu	Ser	Pro	Val	Leu	Val	Ser	Gly
385					390					395					400

Asn Gly Arg Leu Ala Arg Val Tyr Asn Lys Leu
405 410

<210> 16
<211> 440
<212> PRT
<213> *Escherichia fergusonii*

<400> 16

```

Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg Tyr Glu Gly Pro
1      5      10      15

Lys Thr Thr Asn Pro Leu Ala Phe Arg His Tyr Asn Pro Asp Glu Leu
      20      25      30

Val Leu Gly Lys Arg Met Glu Asp His Leu Arg Phe Ala Ala Cys Tyr
      35      40      45

Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe Gly Val Gly Ala
      50      55      60

Phe Asp Arg Pro Trp Gln Gln Pro Gly Glu Ala Leu Met Leu Ala Lys
65      70      75      80

Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His Lys Leu Asn Val Pro
      85      90      95

Phe Tyr Cys Phe His Asp Val Asp Val Ser Pro Glu Gly Ala Ser Leu
      100      105      110

Lys Glu Tyr Lys Asn Asn Phe Ala Glu Met Val Asp Val Leu Ala Ala
      115      120      125

Lys Gln Glu Gln Ser Gly Val Lys Leu Leu Trp Gly Thr Ala Asn Cys
      130      135      140

Phe Thr Asn Pro Arg Tyr Gly Ala Gly Ala Ala Thr Asn Pro Asp Pro
145      150      155      160

Glu Val Phe Ser Trp Ala Ala Thr Gln Val Val Thr Ala Met Glu Ala
      165      170      175

Thr Lys Arg Leu Gly Gly Glu Asn Tyr Val Leu Trp Gly Gly Arg Glu
      180      185      190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Arg Gln Glu Arg Glu Gln
      195      200      205

```

Leu 210	Gly	Arg	Phe	Met	Gln	Met	Val	Val	Glu	His	Lys	His	Lys	Ile	Gly
						215					220				
Phe 225	Gln	Gly	Thr	Leu	Leu	Ile	Glu	Pro	Lys	Pro	Gln	Glu	Pro	Thr	Lys 240
					230					235					
His	Gln	Tyr	Asp	Tyr	Asp	Ala	Ala	Thr	Val	Tyr	Gly	Phe	Leu	Lys 255	Gln
				245					250						
Phe	Gly	Leu	Glu	Lys	Glu	Ile	Lys	Leu	Asn	Ile	Glu	Ala	Asn	His	Ala
			260					265					270		
Thr	Leu	Ala	Gly	His	Ser	Phe	His	His	Glu	Ile	Ala	Thr	Ala	Ile	Ala
		275					280					285			
Leu	Gly	Leu	Phe	Gly	Ser	Val	Asp	Ala	Asn	Arg	Gly	Asp	Ala	Gln	Leu
	290					295					300				
Gly	Trp	Asp	Thr	Asp	Gln	Phe	Pro	Asn	Ser	Val	Glu	Glu	Asn	Ala	Leu 320
305					310					315					
Val	Met	Tyr	Glu	Ile	Leu	Lys	Ala	Gly	Gly	Phe	Thr	Thr	Gly	Gly	Leu
				325					330					335	
Asn	Phe	Asp	Ala	Lys	Val	Arg	Arg	Gln	Ser	Thr	Asp	Lys	Tyr	Asp	Leu
			340					345					350		
Phe	Tyr	Gly	His	Ile	Gly	Ala	Met	Asp	Thr	Met	Ala	Leu	Ala	Leu	Lys
		355					360					365			
Val	Ala	Ala	Arg	Met	Ile	Glu	Asp	Gly	Glu	Leu	Asp	Lys	Arg	Val	Ala
	370					375					380				
Gln	Arg	Tyr	Ser	Gly	Trp	Asn	Ser	Glu	Leu	Gly	Gln	Gln	Ile	Leu	Lys 400
385					390					395					
Gly	Gln	Met	Ser	Leu	Ser	Asp	Leu	Ala	Lys	Tyr	Ala	Gln	Asp	His	Asn
				405					410					415	
Leu	Ser	Pro	Val	His	Gln	Ser	Gly	His	Gln	Glu	Leu	Leu	Glu	Ser	Leu
			420					425					430		
Val	Asn	His	Tyr	Leu	Phe	Asp	Lys								
		435					440								

<210> 17
<211> 391

<212> PRT

<213> *Shigella flexneri*

<400> 17

5

```

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
 1          5          10          15

Thr Ala Leu Thr Glu Leu Tyr Gly Met Glu Asn Pro Ser Ser Gln Pro
 20          25          30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Val
 35          40          45

Gln Asn Ala Ala Gly Asp Ile Val Ser Leu Arg Asp Val Ile Glu Ser
 50          55          60

Asp Lys Ser Thr Leu Leu Gly Glu Ala Val Ala Lys Arg Phe Gly Glu
 65          70          75          80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
 85          90          95

Gln Val His Pro Asn Lys His Asn Ser Glu Ile Gly Phe Ala Lys Glu
100          105          110

Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
115          120          125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
130          135          140

Met Asn Ala Phe Arg Glu Phe Ser Glu Ile Val Ser Leu Leu Gln Pro
145          150          155          160

Val Ala Gly Ala His Pro Ala Ile Ala His Phe Leu Gln Gln Pro His
165          170          175

Ala Glu Arg Leu Ser Glu Leu Phe Ala Asn Leu Leu Asn Met Gln Gly
180          185          190

Glu Glu Lys Ser Arg Ala Leu Ala Ile Leu Lys Ser Ala Leu Asp Ser
195          200          205

Gln Gln Gly Glu Pro Trp Gln Thr Ile Arg Leu Ile Ser Glu Phe Tyr
210          215          220

Pro Glu Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys

```

ES 2 696 235 T3

225											230											235											240
Leu	Asn	Pro	Gly	Glu	Ala	Met	Phe	Leu	Phe	Ala	Glu	Thr	Pro	His	Ala																		
				245					250				255																				
Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn	Ser	Asp	Asn	Val																		
				260					265				270																				
Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Tyr	Ile	Asp	Ile	Pro	Glu	Leu	Val																		
				275					280				285																				
Ala	Asn	Val	Lys	Phe	Glu	Ala	Lys	Pro	Ala	Asn	Gln	Leu	Leu	Thr	Gln																		
				290					295				300																				
Pro	Val	Lys	Gln	Gly	Ala	Glu	Leu	Asp	Phe	Pro	Ile	Pro	Val	Asp	Asp																		
305								310				315				320																	
Phe	Ala	Phe	Ser	Leu	His	Asp	Leu	Ser	Asp	Lys	Glu	Thr	Thr	Ile	Ser																		
				325					330				335																				
Gln	Gln	Ser	Ala	Ala	Ile	Leu	Phe	Cys	Val	Glu	Gly	Asp	Ala	Thr	Leu																		
				340					345				350																				
Trp	Lys	Gly	Ser	Gln	Gln	Leu	Gln	Leu	Lys	Pro	Gly	Glu	Ser	Ala	Phe																		
				355					360				365																				
Ile	Ala	Ala	Asn	Glu	Ser	Pro	Val	Thr	Val	Lys	Gly	His	Gly	Arg	Leu																		
				370					375				380																				
Ala	Arg	Val	Tyr	Asn	Lys	Leu																											
385								390																									

<210> 18
<211> 391
<212> PRT
<213> *Shigella sonnei*

<400> 18

Met	Gln	Lys	Leu	Ile	Asn	Ser	Val	Gln	Asn	Tyr	Ala	Trp	Gly	Ser	Lys
1				5					10					15	
Thr	Ala	Leu	Thr	Glu	Leu	Tyr	Gly	Met	Glu	Asn	Pro	Ser	Ser	Gln	Pro
			20					25					30		
Met	Ala	Glu	Leu	Trp	Met	Gly	Ala	His	Pro	Lys	Ser	Ser	Ser	Arg	Val
		35					40					45			
Gln	Asn	Ala	Ala	Gly	Asp	Ile	Val	Ser	Leu	Arg	Asp	Val	Ile	Glu	Ser

ES 2 696 235 T3

50						55						60			
Asp	Lys	Ser	Thr	Leu	Leu	Gly	Glu	Ala	Val	Ala	Lys	Arg	Phe	Gly	Glu
65					70					75					80
Leu	Pro	Phe	Leu	Phe	Lys	Val	Leu	Cys	Ala	Ala	Gln	Pro	Leu	Ser	Ile
			85						90					95	
Gln	Val	His	Pro	Asn	Lys	Arg	Asn	Ser	Glu	Ile	Gly	Phe	Ala	Lys	Glu
			100					105					110		
Asn	Ala	Ala	Gly	Ile	Pro	Met	Asp	Ala	Ala	Glu	Arg	Asn	Tyr	Lys	Asp
		115					120					125			
Pro	Asn	His	Lys	Pro	Glu	Leu	Val	Phe	Ala	Leu	Thr	Pro	Phe	Leu	Ala
	130					135					140				
Met	Asn	Ala	Phe	Arg	Glu	Phe	Ser	Glu	Ile	Val	Ser	Leu	Leu	Gln	Pro
145					150					155					160
Val	Ala	Gly	Ala	His	Pro	Ala	Ile	Ala	His	Phe	Leu	Gln	Gln	Pro	Asn
				165					170					175	
Ala	Glu	Arg	Leu	Ser	Glu	Leu	Phe	Ala	Ser	Leu	Leu	Asn	Met	Gln	Gly
			180					185					190		
Glu	Glu	Lys	Ser	Arg	Ala	Leu	Ala	Ile	Leu	Lys	Ser	Ala	Leu	Asp	Ser
		195					200					205			
Gln	Gln	Gly	Glu	Pro	Trp	Gln	Thr	Ile	Arg	Leu	Ile	Ser	Glu	Phe	Tyr
	210					215					220				
Pro	Glu	Asp	Ser	Gly	Leu	Phe	Ser	Pro	Leu	Leu	Leu	Asn	Val	Val	Lys
225					230					235					240
Leu	Asn	Pro	Gly	Glu	Ala	Met	Phe	Leu	Phe	Ala	Glu	Thr	Pro	His	Ala
			245						250					255	
Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn	Ser	Asp	Asn	Val
			260					265					270		
Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Tyr	Ile	Asp	Ile	Pro	Glu	Leu	Val
		275					280					285			
Ala	Asn	Val	Lys	Phe	Glu	Ala	Lys	Pro	Ala	Asn	Gln	Leu	Leu	Thr	Gln
	290					295					300				

Pro Val Lys Gln Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Asp Asp
305 310 315 320

Phe Ala Phe Ser Leu His Asp Leu Ser Asp Lys Glu Thr Thr Ile Ser
325 330 335

Gln Gln Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Asp Ala Thr Leu
340 345 350

Trp Lys Gly Ser Gln Gln Leu Gln Leu Lys Pro Gly Glu Ser Ala Phe
355 360 365

Ile Ala Ala Asn Glu Ser Pro Val Thr Val Lys Gly His Gly Arg Leu
370 375 380

Ala Arg Val Tyr Asn Lys Leu
385 390

<210> 19
<211> 440
<212> PRT
<213> *Shigella sonnei*

<400> 19

Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg Tyr Glu Gly Ser
1 5 10 15

Lys Ser Ser Asn Pro Leu Ala Phe Arg His Tyr Asn Pro Asp Glu Leu
20 25 30

Val Leu Gly Lys Arg Met Glu Glu His Leu Arg Phe Ala Ala Cys Tyr
35 40 45

Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe Gly Val Gly Ala
50 55 60

Phe Asn Arg Pro Trp Gln Gln Pro Gly Glu Ala Leu Ala Leu Ala Lys
65 70 75 80

Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His Lys Leu His Val Pro
85 90 95

Phe Tyr Cys Phe His Asp Val Asp Val Ser Pro Glu Gly Ala Ser Leu
100 105 110

Lys Glu Tyr Ile Asn Asn Phe Ala Gln Met Val Asp Val Leu Ala Ala
115 120 125

5

10

ES 2 696 235 T3

Lys	Gln	Glu	Glu	Ser	Gly	Val	Lys	Leu	Leu	Trp	Gly	Thr	Ala	Asn	Cys	130	135	140
Phe	Thr	Asn	Pro	Arg	Tyr	Gly	Ala	Gly	Ala	Ala	Thr	Asn	Pro	Asp	Pro	145	150	155
Glu	Val	Phe	Ser	Trp	Ala	Ala	Thr	Gln	Val	Val	Thr	Ala	Met	Glu	Ala	165	170	175
Thr	His	Lys	Leu	Gly	Gly	Glu	Asn	Tyr	Val	Leu	Trp	Gly	Gly	Arg	Glu	180	185	190
Gly	Tyr	Glu	Thr	Leu	Leu	Asn	Thr	Asp	Leu	Arg	Gln	Glu	Arg	Glu	Gln	195	200	205
Leu	Gly	Arg	Phe	Met	Gln	Met	Val	Val	Glu	His	Lys	His	Lys	Ile	Gly	210	215	220
Phe	Gln	Gly	Thr	Leu	Leu	Ile	Glu	Pro	Lys	Pro	Gln	Glu	Pro	Thr	Lys	225	230	235
His	Gln	Tyr	Asp	Tyr	Asp	Ala	Ala	Thr	Val	Tyr	Gly	Phe	Leu	Lys	Gln	245	250	255
Phe	Gly	Leu	Glu	Lys	Glu	Ile	Lys	Leu	Asn	Ile	Glu	Ala	Asn	His	Ala	260	265	270
Thr	Leu	Ala	Gly	His	Ser	Phe	His	His	Glu	Ile	Ala	Thr	Ala	Ile	Ala	275	280	285
Leu	Gly	Leu	Phe	Gly	Ser	Val	Asp	Ala	Asn	Arg	Gly	Asp	Ala	Gln	Leu	290	295	300
Gly	Trp	Asp	Thr	Asp	Gln	Phe	Pro	Asn	Ser	Val	Glu	Glu	Asn	Ala	Leu	305	310	315
Val	Met	Tyr	Glu	Ile	Leu	Lys	Ala	Gly	Gly	Phe	Thr	Thr	Gly	Gly	Leu	325	330	335
Asn	Phe	Asp	Ala	Lys	Val	Arg	Arg	Gln	Ser	Thr	Asp	Lys	Tyr	Asp	Leu	340	345	350
Phe	Tyr	Gly	His	Ile	Gly	Ala	Met	Asp	Thr	Met	Ala	Leu	Ala	Leu	Lys	355	360	365
Ile	Ala	Ala	Arg	Met	Ile	Glu	Asp	Gly	Glu	Leu	Asp	Lys	Arg	Ile	Ala	370	375	380

Gln Arg Tyr Ser Gly Trp Asn Ser Glu Leu Gly Gln Gln Ile Leu Lys
385 390 395 400

Gly Gln Met Ser Leu Ala Asp Leu Ala Lys Tyr Ala Gln Glu His Asn
405 410 415

Leu Ser Pro Val His Gln Ser Gly Arg Gln Glu Gln Leu Glu Asn Leu
420 425 430

Val Asn His Tyr Leu Phe Asp Lys
435 440

<210> 20

<211> 391

<212> PRT

<213> *Shigella dysenteriae*

<400> 20

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Met Glu Asn Pro Ser Ser Gln Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Val
35 40 45

Gln Asn Ala Ala Gly Asp Ile Val Ser Leu Arg Asp Val Ile Glu Ser
50 55 60

Asp Lys Ser Thr Leu Leu Gly Glu Ala Val Ala Lys Arg Phe Gly Glu
65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
85 90 95

Gln Val His Pro Asn Lys His Asn Ser Glu Ile Gly Phe Ala Lys Glu
100 105 110

Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
115 120 125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
130 135 140

Met Asn Ala Phe Arg Glu Phe Ser Glu Ile Val Ser Leu Leu Gln Pro
145 150 155 160

ES 2 696 235 T3

Val Ala Gly Ala His Pro Ala Ile Ala His Phe Leu Gln Gln Pro His
165 170 175

Ala Glu His Leu Ser Glu Leu Phe Ala Ser Leu Leu Asn Met Gln Gly
180 185 190

Glu Glu Lys Ser Arg Ala Leu Ala Ile Leu Lys Ser Thr Leu Asp Ser
195 200 205

Gln Gln Gly Glu Pro Trp Gln Thr Ile Arg Leu Ile Ser Glu Phe Tyr
210 215 220

Pro Glu Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
225 230 235 240

Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
245 250 255

Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val
260 265 270

Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
275 280 285

Ala Asn Val Lys Phe Glu Ala Lys Pro Ala Asn Gln Leu Leu Thr Gln
290 295 300

Pro Val Lys Gln Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Asp Asp
305 310 315 320

Phe Ala Phe Ser Leu His Asp Leu Ser Asp Lys Glu Thr Thr Ile Ser
325 330 335

Gln Gln Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Asp Ala Thr Leu
340 345 350

Trp Lys Gly Ser Gln Gln Leu Gln Leu Lys Pro Gly Glu Ser Ala Phe
355 360 365

Ile Ala Ala Asn Glu Ser Pro Val Thr Val Lys Gly His Gly Arg Leu
370 375 380

Ala Arg Val Tyr Asn Lys Leu
385 390

<210> 21
<211> 300

ES 2 696 235 T3

<212> PRT

<213> *Shigella dysenteriae*

<400> 21

5

```

Met Arg Ile Gly Ile Asp Leu Gly Gly Thr Lys Thr Glu Val Ile Ala
 1          5          10          15

Leu Gly Asp Ala Gly Glu Gln Leu Tyr Arg His Arg Leu Pro Thr Pro
          20          25          30

Arg Asp Asp Tyr Arg Gln Thr Ile Glu Thr Ile Ala Thr Leu Val Asp
          35          40          45

Met Ala Glu Gln Ala Thr Gly Gln Arg Gly Thr Val Gly Met Gly Ile
 50          55          60

Pro Gly Ala Ile Ser Pro Tyr Thr Gly Val Val Lys Asn Ala Asn Ser
65          70          75          80

Thr Trp Leu Asn Gly Gln Pro Phe Asp Lys Asp Leu Ser Ala Arg Leu
          85          90          95

Gln Arg Glu Val Arg Leu Ala Asn Asp Ala Asn Cys Leu Ala Val Ser
          100          105          110

Glu Ala Val Asp Gly Ala Ala Ala Gly Ala Gln Thr Val Phe Ala Val
115          120          125

Ile Ile Gly Thr Gly Cys Gly Ala Gly Val Ala Phe Asn Gly Arg Ala
130          135          140

His Ile Gly Gly Asn Gly Thr Ala Gly Glu Trp Gly His Asn Pro Leu
145          150          155          160

Pro Trp Met Asp Glu Asp Glu Leu Arg Tyr Arg Glu Glu Val Pro Cys
          165          170          175

Tyr Cys Gly Lys Gln Gly Cys Ile Glu Thr Phe Ile Ser Gly Thr Gly
          180          185          190

Phe Ala Thr Asp Tyr Arg Arg Leu Ser Gly His Ala Leu Lys Gly Ser
          195          200          205

Glu Ile Ile Ser Leu Val Glu Glu Ser Asp Pro Val Ala Glu Leu Ala
210          215          220

Leu Arg Arg Tyr Glu Leu Arg Leu Ala Lys Ser Leu Ala His Val Val
225          230          235          240

```

Asn Ile Leu Asp Pro Asp Val Ile Val Leu Gly Gly Gly Met Ser Asn
245 250 255

Val Asp Arg Leu Tyr Gln Thr Val Pro Gln Leu Ile Lys Gln Phe Val
260 265 270

Phe Gly Gly Glu Cys Glu Thr Pro Val Arg Lys Ala Lys His Gly Asp
275 280 285

Ser Ser Gly Val Arg Gly Ala Ala Trp Leu Trp Pro
290 295 300

<210> 22

<211> 440

<212> PRT

<213> *Shigella dysenteriae*

<400> 22

Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg Tyr Glu Gly Ser
1 5 10 15

Lys Ser Ser Asn Pro Leu Ala Phe Arg His Tyr Asn Pro Asp Glu Leu
20 25 30

Val Leu Gly Lys Arg Met Glu Glu His Leu Arg Phe Ala Ala Cys Tyr
35 40 45

Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe Gly Val Gly Ala
50 55 60

Phe Asn Arg Pro Trp Gln Gln Pro Gly Glu Ala Leu Ala Leu Ala Lys
65 70 75 80

Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His Lys Leu His Val Pro
85 90 95

Phe Tyr Cys Phe His Asp Val Asp Val Ser Pro Glu Gly Ala Ser Leu
100 105 110

Lys Glu Tyr Ile Asn Asn Phe Ala Gln Met Val Asp Val Leu Ala Gly
115 120 125

Lys Gln Glu Glu Ser Gly Val Lys Leu Leu Trp Gly Thr Ala Asn Cys
130 135 140

Phe Thr Asn Pro Arg Tyr Gly Ala Gly Ala Ala Thr Asn Pro Asp Pro
145 150 155 160

5

10

ES 2 696 235 T3

Glu	Val	Phe	Ser	Trp	Ala	Ala	Thr	Gln	Val	Val	Thr	Ala	Met	Glu	Ala	165	170	175
Thr	His	Lys	Leu	Gly	Gly	Glu	Asn	Tyr	Val	Leu	Trp	Gly	Gly	Arg	Glu	180	185	190
Gly	Tyr	Glu	Thr	Leu	Leu	Asn	Thr	Asp	Leu	Arg	Gln	Glu	Arg	Glu	Gln	195	200	205
Leu	Gly	Arg	Phe	Met	Gln	Met	Val	Val	Glu	His	Lys	His	Lys	Ile	Gly	210	215	220
Phe	Gln	Gly	Thr	Leu	Leu	Ile	Glu	Pro	Lys	Pro	Gln	Glu	Pro	Thr	Lys	225	230	235
His	Gln	Tyr	Asp	Tyr	Asp	Ala	Ala	Thr	Val	Tyr	Gly	Phe	Leu	Lys	Gln	245	250	255
Phe	Gly	Leu	Glu	Lys	Glu	Ile	Lys	Leu	Asn	Ile	Glu	Ala	Asn	His	Ala	260	265	270
Thr	Leu	Ala	Gly	His	Ser	Phe	His	His	Glu	Ile	Ala	Thr	Ala	Ile	Ala	275	280	285
Leu	Gly	Leu	Phe	Gly	Ser	Val	Asp	Ala	Asn	Arg	Gly	Asp	Ala	Gln	Leu	290	295	300
Gly	Trp	Asp	Thr	Asp	Gln	Phe	Pro	Asn	Ser	Val	Glu	Glu	Asn	Ala	Leu	305	310	315
Val	Met	Tyr	Glu	Ile	Leu	Lys	Ala	Gly	Gly	Phe	Thr	Thr	Gly	Gly	Leu	325	330	335
Asn	Phe	Asp	Ala	Lys	Val	Arg	Arg	Gln	Ser	Thr	Asp	Lys	Tyr	Asp	Leu	340	345	350
Phe	Tyr	Gly	His	Ile	Gly	Ala	Met	Asp	Thr	Met	Ala	Leu	Ala	Leu	Lys	355	360	365
Ile	Ala	Ala	Cys	Met	Ile	Glu	Asp	Gly	Glu	Leu	Asp	Lys	Arg	Ile	Ala	370	375	380
Gln	Arg	Tyr	Ser	Gly	Trp	Asn	Ser	Glu	Leu	Gly	Gln	His	Ile	Leu	Lys	385	390	395
Gly	Gln	Met	Ser	Leu	Ala	Asp	Leu	Ala	Lys	Tyr	Ala	Gln	Glu	His	Asn			

ES 2 696 235 T3

405

410

415

Leu Ser Pro Val His Gln Ser Gly Arg Gln Glu Gln Leu Glu Asn Leu
420 425 430

Val Asn His Tyr Leu Phe Asp Lys
435 440

<210> 23

<211> 350

<212> PRT

<213> *Citrobacter koseri*

<400> 23

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Ile Ala Asn Pro Ser Gln Gln Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Ile
35 40 45

Glu Gly Ala Asp Gly Lys Ala Val Ser Leu Arg Asp Val Ile Glu His
50 55 60

Asp Lys Ala Thr Leu Leu Gly Glu Ser Val Ala Lys Arg Phe Gly Glu
65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
85 90 95

Gln Val His Pro Asn Lys His Asn Ser Glu Met Gly Phe Ala Lys Glu
100 105 110

Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
115 120 125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
130 135 140

Met Asn Ala Phe Arg Glu Phe Ser Glu Val Val Ser Leu Leu Gln Pro
145 150 155 160

Val Ser Gly Ala His Thr Ala Ile Ala His Phe Leu Gln Glu Pro Asn
165 170 175

Ala Glu Arg Leu Ser Gln Leu Phe Ala Ser Leu Leu Asn Met Gln Gly

5

10

ES 2 696 235 T3

```

180              185              190

Asp Glu Lys Ser Arg Ala Leu Ala Val Leu Lys Ala Ala Leu Asn Ser
195              200              205

Gln Gln Gly Glu Pro Trp Gln Thr Ile Arg Val Ile Ser Glu Phe Tyr
210              215              220

Pro Asp Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
225              230              235              240

Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
245              250              255

Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val
260              265              270

Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
275              280              285

Ala Asn Val Lys Phe Val Ala Lys Pro Ala Asp Gln Leu Leu Thr Thr
290              295              300

Pro Val Lys His Gly Thr Glu Leu Asp Phe Pro Ile Pro Val Glu Asp
305              310              315              320

Phe Ala Phe Ser Leu His Asp Leu Ser Ala Gln Glu Thr Ala Ile Ser
325              330              335

Gln Gln Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Glu Ala
340              345              350

```

<210> 24
 <211> 444
 <212> PRT
 <213> *Citrobacter koseri*

<400> 24

```

Met Glu Phe Ile Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg
1              5              10              15

Tyr Glu Gly Pro Lys Ser Thr Asn Pro Leu Ala Phe Arg His Tyr Asn
20              25              30

Pro Asp Glu Ile Val Leu Gly Lys Arg Met Glu Asp His Leu Arg Phe
35              40              45

Ala Ala Cys Tyr Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe

```

ES 2 696 235 T3

50		55		60
Gly Val Gly Ala Phe Asn Arg Pro Trp Gln Gln Pro Gly Asp Ala Leu				
65		70		75
Glu Leu Ala Lys Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His Lys				
	85		90	95
Leu Asn Val Pro Tyr Tyr Cys Phe His Asp Val Asp Val Ser Pro Glu				
	100		105	110
Gly Ala Ser Leu Lys Glu Tyr Lys Asn Asn Phe Ala Gln Met Val Asp				
	115		120	125
Val Leu Glu Ser Lys Gln Glu Gln Ser Gly Val Lys Leu Leu Trp Gly				
	130		135	140
Thr Ala Asn Cys Phe Thr Asn Pro Arg Tyr Gly Ala Gly Ala Ala Thr				
145		150		155
Asn Pro Asp Pro Glu Val Phe Ser Trp Ala Ala Thr Gln Val Val Asn				
	165		170	175
Ala Met Asn Ala Thr His Lys Leu Gly Gly Glu Asn Tyr Val Leu Trp				
	180		185	190
Gly Gly Arg Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Arg Gln				
	195		200	205
Glu Arg Glu Gln Ile Gly Arg Phe Met Gln Met Val Val Glu His Lys				
	210		215	220
His Lys Thr Gly Phe Gln Gly Thr Leu Leu Ile Glu Pro Lys Pro Gln				
225		230		235
Glu Pro Thr Lys His Gln Tyr Asp Tyr Asp Val Ala Thr Val Tyr Gly				
	245		250	255
Phe Leu Lys Gln Phe Gly Leu Glu Lys Glu Ile Lys Val Asn Ile Glu				
	260		265	270
Ala Asn His Ala Thr Leu Ala Gly His Ser Phe His His Glu Ile Ala				
	275		280	285
Ser Ala Ile Ala Leu Gly Ile Phe Gly Ser Val Asp Ala Asn Arg Gly				
	290		295	300

Asp Ala Gln Leu Gly Trp Asp Thr Asp Gln Phe Pro Val Ser Val Glu
305 310 315 320

Glu Asn Ala Leu Val Met Tyr Glu Ile Leu Lys Ala Gly Gly Phe Thr
325 330 335

Thr Gly Gly Leu Asn Phe Asp Ala Lys Val Arg Arg Gln Ser Thr Asp
340 345 350

Lys Tyr Asp Leu Phe Tyr Gly His Ile Gly Ala Met Asp Thr Met Ala
355 360 365

Leu Ser Leu Lys Val Ala Ala Arg Met Ile Glu Asp Gly Glu Leu Asp
370 375 380

Lys Arg Val Ala Lys Arg Tyr Ser Gly Trp Asn Ser Glu Leu Gly Gln
385 390 395 400

Gln Ile Leu Lys Gly Gln Leu Ser Leu Ser Asp Val Ala His Tyr Ala
405 410 415

Glu Gln Arg Asn Leu Ala Pro Val His Gln Ser Gly His Gln Glu Leu
420 425 430

Leu Glu Asn Leu Val Asn His Tyr Leu Phe Asp Lys
435 440

<210> 25

<211> 391

5 <212> PRT

<213> *Citrobacter rodentium*

<400> 25

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Ile Ala Asn Pro Gln Gln Leu Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Gln Val
35 40 45

Val Asp Ala Ser Gly Gln Arg Val Ser Leu Arg Asp Val Ile Glu Ala
50 55 60

Asp Lys Ser Arg Leu Leu Gly Asp Ala Val Ala Arg Arg Phe Gly Glu
65 70 75 80

ES 2 696 235 T3

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
 85 90 95
 Gln Val His Pro Asn Lys Arg Asn Ser Glu Leu Gly Phe Ala Lys Glu
 100 105 110
 Asn Val Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
 115 120 125
 Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
 130 135 140
 Met Asn Ala Phe Arg Glu Phe Ser Glu Ile Val Ser Leu Leu Gln Pro
 145 150 155 160
 Val Ala Gly Ala His Thr Ala Ile Ala His Phe Leu Glu Ala Pro Asp
 165 170 175
 Ala Glu Arg Leu Ser Ala Leu Phe Ala Ala Leu Leu Asn Met Gln Gly
 180 185 190
 Glu Glu Lys Ser Arg Ala Leu Ala Ile Leu Lys Ala Ala Leu Glu Thr
 195 200 205
 Gln Ser Gly Glu Pro Trp Gln Thr Ile Arg Leu Ile Gly Glu Phe Tyr
 210 215 220
 Pro Asp Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
 225 230 235 240
 Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
 245 250 255
 Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val
 260 265 270
 Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
 275 280 285
 Ala Asn Val Lys Phe Glu Ala Lys Pro Ala Asp Arg Leu Leu Thr Ala
 290 295 300
 Pro Leu Thr Ser Gly Ala Glu Gln Asp Phe Pro Ile Pro Val Asp Asp
 305 310 315 320
 Phe Ala Phe Ser Leu His Asp Leu Ser Ala Gln Gln Ala Glu Ile Ala
 325 330 335

Gln Gln Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Glu Ala Thr Leu
340 345 350

Arg Lys Gly Glu Gln Ser Leu Thr Leu Lys Pro Gly Glu Ser Ala Phe
355 360 365

Ile Ser Ala Ala Glu Ser Pro Val Ser Val Ser Gly Ala Gly Arg Leu
370 375 380

Ala Arg Val Tyr Asn Lys Leu
385 390

<210> 26

<211> 302

<212> PRT

<213> *Citrobacter rodentium*

<400> 26

Met Arg Ile Gly Ile Asp Leu Gly Gly Thr Lys Thr Glu Val Ile Ala
1 5 10 15

Leu Gly Glu Ser Gly Glu Gln Leu Phe Arg His Arg Leu Pro Thr Pro
20 25 30

Arg Asn Asp Tyr Arg Gln Thr Ile Glu Thr Ile Ala Thr Leu Val Glu
35 40 45

Met Ala Glu Lys Ala Thr Gly Gln Thr Gly Thr Val Gly Met Gly Ile
50 55 60

Pro Gly Ser Leu Ser Pro Tyr Thr Gly Val Val Lys Asn Ala Asn Ser
65 70 75 80

Thr Trp Leu Asn Gly Gln Pro Phe Asp Arg Asp Val Ser Ala Arg Leu
85 90 95

Gln Arg Glu Val Arg Leu Ala Asn Asp Ala Asn Cys Leu Ala Val Ser
100 105 110

Glu Ala Val Asp Gly Ala Ala Ala Gly Ala Gln Thr Val Phe Ala Val
115 120 125

Ile Ile Gly Thr Gly Cys Gly Ala Gly Val Ala Leu Asn Gly Arg Ala
130 135 140

His Ala Gly Gly Asn Gly Thr Ala Gly Glu Trp Gly His Asn Pro Leu
145 150 155 160

5

Pro Trp Met Asp Asp Asp Glu Leu Arg Tyr Arg Glu Glu Val Pro Cys
165 170 175

Tyr Cys Gly Lys Gln Gly Cys Ile Glu Thr Phe Ile Ser Gly Thr Gly
180 185 190

Phe Ala Thr Asp Tyr Gln Arg Leu Ser Gly Asn Ser Leu Thr Gly Asn
195 200 205

Glu Ile Met Arg Arg Val Glu Glu Gln Asp Pro Leu Ala Glu Leu Ala
210 215 220

Leu Gly Arg Tyr Glu Leu Arg Leu Ala Lys Ser Leu Ala His Val Val
225 230 235 240

Asn Ile Leu Asp Pro Asp Val Ile Val Leu Gly Gly Gly Met Ser Asn
245 250 255

Val Asp Arg Leu Tyr Asn Thr Leu Pro Ala Leu Ile Gly Gln Phe Val
260 265 270

Phe Gly Gly Glu Cys Glu Thr Pro Val Arg Lys Ala Leu His Gly Asp
275 280 285

Ser Ser Gly Val Arg Gly Ala Ala Trp Leu Trp Pro Gln Arg
290 295 300

<210> 27

<211> 440

<212> PRT

<213> *Citrobacter rodentium*

<400> 27

Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg Phe Glu Gly Pro
1 5 10 15

Lys Ser Thr Asn Pro Leu Ala Phe Arg His Tyr Asn Pro Asp Glu Ile
20 25 30

Val Leu Gly Lys Arg Met Glu Asp His Leu Arg Phe Ala Ala Cys Tyr
35 40 45

Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe Gly Val Gly Ala
50 55 60

Phe Asp Arg Pro Trp Gln Gln Pro Gly Glu Ala Leu Ala Leu Ala Lys
65 70 75 80

5

10

ES 2 696 235 T3

Arg	Lys	Ala	Asp	Val	Ala	Phe	Glu	Phe	Phe	His	Lys	Leu	Asn	Val	Pro	85	90	95
Phe	Tyr	Cys	Phe	His	Asp	Val	Asp	Val	Ser	Pro	Glu	Gly	Ala	Ser	Leu	100	105	110
Lys	Glu	Tyr	Lys	Asn	Asn	Phe	Ala	Gln	Met	Val	Asp	Val	Leu	Ala	Ala	115	120	125
Lys	Gln	Glu	Glu	Ser	Gly	Val	Lys	Leu	Leu	Trp	Gly	Thr	Ala	Asn	Cys	130	135	140
Phe	Thr	Asn	Pro	Arg	Tyr	Gly	Ala	Gly	Ala	Ala	Thr	Asn	Pro	Asp	Pro	145	150	155
Glu	Val	Phe	Ser	Trp	Ala	Ala	Thr	Gln	Val	Val	Thr	Ala	Met	Asp	Ala	165	170	175
Thr	His	Lys	Leu	Gly	Gly	Glu	Asn	Tyr	Val	Leu	Trp	Gly	Gly	Arg	Glu	180	185	190
Gly	Tyr	Glu	Thr	Leu	Leu	Asn	Thr	Asp	Leu	Arg	Gln	Glu	Arg	Glu	Gln	195	200	205
Ile	Gly	Arg	Phe	Met	Gln	Met	Val	Val	Glu	His	Lys	His	Lys	Thr	Gly	210	215	220
Phe	Gln	Gly	Thr	Leu	Leu	Ile	Glu	Pro	Lys	Pro	Gln	Glu	Pro	Thr	Lys	225	230	235
His	Gln	Tyr	Asp	Tyr	Asp	Ala	Ala	Thr	Val	Tyr	Gly	Phe	Leu	Lys	Gln	245	250	255
Phe	Gly	Leu	Glu	Lys	Glu	Ile	Lys	Leu	Asn	Ile	Glu	Ala	Asn	His	Ala	260	265	270
Thr	Leu	Ala	Gly	His	Ser	Phe	His	His	Glu	Ile	Ala	Thr	Ala	Ile	Ala	275	280	285
Leu	Gly	Leu	Phe	Gly	Ser	Val	Asp	Ala	Asn	Arg	Gly	Asp	Ala	Gln	Leu	290	295	300
Gly	Trp	Asp	Thr	Asp	Gln	Phe	Pro	Asn	Ser	Val	Glu	Glu	Asn	Ala	Leu	305	310	315
Val	Met	Tyr	Glu	Ile	Leu	Lys	Ala	Gly	Gly	Phe	Thr	Thr	Gly	Gly	Leu	325	330	335

Asn Phe Asp Ala Lys Val Arg Arg Gln Ser Thr Asp Lys Tyr Asp Leu
340 345 350

Phe Tyr Gly His Ile Gly Ala Met Asp Thr Met Ala Leu Ser Leu Lys
355 360 365

Ile Ala Ala Arg Met Ile Glu Asp Gly Glu Leu Asp Lys Arg Val Ala
370 375 380

Lys Arg Tyr Ala Gly Trp Asn Ser Glu Leu Gly Gln Gln Ile Leu Lys
385 390 395 400

Gly Gln Met Ser Leu Ser Glu Leu Ala Gln Tyr Ala Glu Gln His Asn
405 410 415

Leu Ala Pro Val His Gln Ser Gly His Gln Glu Leu Leu Glu Asn Leu
420 425 430

Val Asn His Tyr Leu Phe Asn Lys
435 440

<210> 28

<211> 416

<212> PRT

<213> *Citrobacter youngae*

<400> 28

Met Cys Arg Leu Leu Pro Leu Gly Phe Leu Gly Leu Ile Ala Ala Glu
1 5 10 15

Asn Ser Thr Leu Lys Gln Gly Leu Arg Met Gln Lys Leu Ile Asn Ser
20 25 30

Val Gln Asn Tyr Ala Trp Gly Ser Lys Thr Ala Leu Thr Glu Leu Tyr
35 40 45

Gly Met Ala Asn Pro Ala Gln Gln Pro Met Ala Glu Leu Trp Met Gly
50 55 60

Ala His Pro Lys Ser Ser Ser Gln Val Leu Gly Ala Asp Gly Gln Thr
65 70 75 80

Ile Ala Leu Arg Asp Val Ile Glu Arg Asp Lys Ser Ala Leu Leu Gly
85 90 95

Glu Ala Val Ala Lys Arg Phe Gly Glu Leu Pro Phe Leu Phe Lys Val
100 105 110

5

10

ES 2 696 235 T3

Leu Cys Ala Ala Gln Pro Leu Ser Ile Gln Val His Pro Asn Lys His
 115 120 125
 Asn Ser Glu Ile Gly Phe Ala Lys Glu Asn Ala Ala Gly Ile Pro Met
 130 135 140
 Asp Ala Ala Glu Arg Asn Tyr Lys Asp Pro Asn His Lys Pro Glu Leu
 145 150 155 160
 Val Phe Ala Leu Thr Pro Phe Leu Ala Met Asn Ala Phe Arg Glu Phe
 165 170 175
 Ser Asp Ile Val Ala Leu Leu Gln Pro Val Ser Ala Ala His Pro Ala
 180 185 190
 Ile Ala His Phe Leu Gln Glu Pro Cys Ala Glu Arg Leu Ser Tyr Leu
 195 200 205
 Phe Ala Ser Leu Leu Asn Met Gln Asp Glu Glu Lys Ser His Ala Leu
 210 215 220
 Ala Ile Leu Lys Ser Ala Leu Asn Ser Gln Gln Gly Glu Pro Trp Gln
 225 230 235 240
 Thr Ile Arg Leu Val Ala Asp Phe Tyr Pro Asn Asp Ser Gly Leu Phe
 245 250 255
 Ser Pro Leu Leu Leu Asn Val Val Lys Leu Asn Pro Gly Glu Ala Met
 260 265 270
 Phe Leu Phe Ala Glu Thr Pro His Ala Tyr Leu Gln Gly Val Ala Leu
 275 280 285
 Glu Val Met Ala Asn Ser Asp Asn Val Leu Arg Ala Gly Leu Thr Pro
 290 295 300
 Lys Tyr Ile Asp Ile Pro Glu Leu Val Ala Asn Val Lys Phe Val Ala
 305 310 315 320
 Lys Pro Ala Asn Gln Leu Leu Thr Thr Pro Ile Lys Thr Ala Gly Glu
 325 330 335
 Leu Asp Phe Pro Ile Pro Val Asp Asp Phe Ala Phe Ser Leu His Glu
 340 345 350
 Leu Thr Ala Asp Gly Ala Asp Ile Ser Gln Gln Ser Ala Ala Ile Leu

355

360

365

Phe Cys Ile Glu Gly Gln Ala Val Leu Ser Lys Gly Glu Gln Arg Leu
 370 375 380

Val Leu Gln Pro Gly Glu Ser Ala Phe Ile Ser Ala Asn Glu Ser Pro
 385 390 395 400

Val Asn Val Ser Gly Thr Gly Arg Leu Ala Arg Val Tyr Asn Lys Leu
 405 410 415

<210> 29

<211> 444

<212> PRT

<213> *Citrobacter youngae*

<400> 29

Met Glu Leu Ile Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg
 1 5 10 15

Phe Glu Gly Thr Lys Ser Thr Asn Pro Leu Ala Phe Arg His Tyr Asn
 20 25 30

Pro Asp Glu Ile Val Leu Gly Lys Arg Met Glu Asp His Leu Arg Phe
 35 40 45

Ala Ala Cys Tyr Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe
 50 55 60

Gly Met Gly Ala Phe Asp Arg Pro Trp Gln Gln Pro Gly Glu Ala Leu
 65 70 75 80

Ala Leu Ala Lys Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His Lys
 85 90 95

Leu Asn Val Pro Tyr Tyr Cys Phe His Asp Val Asp Val Ser Pro Glu
 100 105 110

Gly Ala Ser Leu Lys Glu Tyr Lys Asn Asn Phe Ala Gln Met Val Asp
 115 120 125

Val Leu Ala Ala Lys Gln Glu Gln Ser Gly Val Lys Leu Leu Trp Gly
 130 135 140

Thr Ala Asn Cys Phe Thr Asn Pro Arg Tyr Gly Ala Gly Ala Ala Thr
 145 150 155 160

Asn Pro Asp Pro Glu Val Phe Ser Trp Ala Ala Thr Gln Val Val Thr

5

10

ES 2 696 235 T3

165								170					175				
Ala	Met	Asp	Ala	Thr	His	Lys	Leu	Gly	Gly	Glu	Asn	Tyr	Val	Leu	Trp		
			180					185					190				
Gly	Gly	Arg	Glu	Gly	Tyr	Glu	Thr	Leu	Leu	Asn	Thr	Asp	Leu	Arg	Gln		
		195					200					205					
Glu	Arg	Glu	Gln	Ile	Gly	Arg	Phe	Met	Gln	Leu	Val	Val	Glu	His	Lys		
	210					215					220						
His	Lys	Ile	Gly	Phe	Gln	Gly	Thr	Leu	Leu	Ile	Glu	Pro	Lys	Pro	Gln		
225					230					235					240		
Glu	Pro	Thr	Lys	His	Gln	Tyr	Asp	Tyr	Asp	Ala	Ala	Thr	Val	Tyr	Gly		
				245					250					255			
Phe	Leu	Lys	Gln	Phe	Gly	Leu	Glu	Lys	Glu	Ile	Lys	Leu	Asn	Ile	Glu		
			260					265					270				
Ala	Asn	His	Ala	Thr	Leu	Ala	Gly	His	Ser	Phe	His	His	Glu	Ile	Ala		
		275					280					285					
Thr	Ala	Ile	Ala	Leu	Gly	Leu	Phe	Gly	Ser	Val	Asp	Ala	Asn	Arg	Gly		
	290					295					300						
Asp	Ala	Gln	Leu	Gly	Trp	Asp	Thr	Asp	Gln	Phe	Pro	Asn	Ser	Val	Glu		
305					310					315					320		
Glu	Asn	Ala	Leu	Val	Met	Tyr	Glu	Ile	Leu	Lys	Ala	Gly	Gly	Phe	Thr		
				325					330					335			
Thr	Gly	Gly	Leu	Asn	Phe	Asp	Ala	Lys	Val	Arg	Arg	Gln	Ser	Thr	Asp		
			340					345					350				
Lys	Tyr	Asp	Leu	Phe	Tyr	Gly	His	Ile	Gly	Ala	Met	Asp	Thr	Met	Ala		
		355					360					365					
Leu	Ser	Leu	Lys	Ile	Ala	Ala	Arg	Met	Ile	Glu	Asp	Gly	Gly	Leu	Asp		
	370					375					380						
Gln	Arg	Val	Ala	Lys	Arg	Tyr	Ala	Gly	Trp	Asn	Gly	Glu	Leu	Gly	Gln		
385					390					395					400		
Gln	Ile	Leu	Lys	Gly	Gln	Met	Thr	Leu	Thr	Glu	Ile	Ala	Gln	Tyr	Ala		
				405					410					415			

ES 2 696 235 T3

Glu Gln His Asn Leu Ala Pro Val His Gln Ser Gly His Gln Glu Gln
420 425 430

Leu Glu Asn Leu Val Asn His Tyr Leu Phe Asp Lys
435 440

<210> 30
<211> 391
<212> PRT
<213> *Salmonella enterica*

<400> 30

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Ile Ala Asn Pro Gln Gln Gln Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Ile
35 40 45

Thr Thr Ala Asn Gly Glu Thr Val Ser Leu Arg Asp Ala Ile Glu Lys
50 55 60

Asn Lys Thr Ala Met Leu Gly Glu Ala Val Ala Asn Arg Phe Gly Glu
65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
85 90 95

Gln Val His Pro Asn Lys Arg Asn Ser Glu Ile Gly Phe Ala Lys Glu
100 105 110

Asn Ala Ala Gly Ile Pro Met Asp Ala Ala Glu Arg Asn Tyr Lys Asp
115 120 125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
130 135 140

Met Asn Ala Phe Arg Glu Phe Ser Asp Ile Val Ser Leu Leu Gln Pro
145 150 155 160

Val Ala Gly Ala His Ser Ala Ile Ala His Phe Leu Gln Ala Pro Asn
165 170 175

Ala Glu Arg Leu Ser Gln Leu Phe Ala Ser Leu Leu Asn Met Gln Gly
180 185 190

5

10

ES 2 696 235 T3

Glu Glu Lys Ser Arg Ala Leu Ala Val Leu Lys Ala Ala Leu Asn Ser
 195 200 205
 Gln Gln Gly Glu Pro Trp Gln Thr Ile Arg Val Ile Ser Glu Tyr Tyr
 210 215 220
 Pro Asp Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
 225 230 235 240
 Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
 245 250 255
 Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val
 260 265 270
 Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
 275 280 285
 Ala Asn Val Lys Phe Glu Pro Lys Pro Ala Ser Glu Leu Leu Thr Ala
 290 295 300
 Pro Val Lys Ser Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Asp Asp
 305 310 315 320
 Phe Ala Phe Ser Leu His Asp Leu Ala Leu Gln Glu Thr Ser Ile Gly
 325 330 335
 Gln His Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Glu Ala Val Leu
 340 345 350
 Arg Lys Asp Glu Gln Arg Leu Val Leu Lys Pro Gly Glu Ser Ala Phe
 355 360 365
 Ile Gly Ala Asp Glu Ser Pro Val Asn Ala Ser Gly Thr Gly Arg Leu
 370 375 380
 Ala Arg Val Tyr Asn Lys Leu
 385 390

<210> 31
 <211> 440
 <212> PRT
 <213> *Salmonella enterica*

<400> 31

Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg Tyr Glu Gly Pro
 1 5 10 15

ES 2 696 235 T3

Gln	Ser	Thr	Asn	Pro	Leu	Ala	Phe	Arg	His	Tyr	Asn	Pro	Asp	Glu	Leu		
			20					25					30				
Val	Leu	Gly	Lys	Arg	Met	Glu	Asp	His	Leu	Arg	Phe	Ala	Ala	Cys	Tyr		
		35					40					45					
Trp	His	Thr	Phe	Cys	Trp	Asn	Gly	Ala	Asp	Met	Phe	Gly	Val	Gly	Ala		
	50					55					60						
Phe	Asn	Arg	Pro	Trp	Gln	Gln	Pro	Gly	Glu	Ala	Leu	Glu	Leu	Ala	Lys		
65					70					75					80		
Arg	Lys	Ala	Asp	Val	Ala	Phe	Glu	Phe	Phe	His	Lys	Leu	Asn	Val	Pro		
				85					90					95			
Phe	Tyr	Cys	Phe	His	Asp	Val	Asp	Val	Ser	Pro	Glu	Gly	Ala	Ser	Leu		
			100					105					110				
Lys	Glu	Tyr	Lys	Asn	Asn	Phe	Ala	Gln	Met	Val	Asp	Val	Leu	Ala	Ala		
		115					120					125					
Lys	Gln	Glu	Gln	Ser	Gly	Val	Lys	Leu	Leu	Trp	Gly	Thr	Ala	Asn	Cys		
	130					135					140						
Phe	Thr	Asn	Pro	Arg	Tyr	Gly	Ala	Gly	Ala	Ala	Thr	Asn	Pro	Asp	Pro		
145					150					155					160		
Glu	Val	Phe	Ser	Trp	Ala	Ala	Thr	Gln	Val	Val	Thr	Ala	Met	Asn	Ala		
				165					170					175			
Thr	His	Lys	Leu	Gly	Gly	Glu	Asn	Tyr	Val	Leu	Trp	Gly	Gly	Arg	Glu		
			180					185					190				
Gly	Tyr	Glu	Thr	Leu	Leu	Asn	Thr	Asp	Leu	Arg	Gln	Glu	Arg	Glu	Gln		
		195					200					205					
Ile	Gly	Arg	Phe	Met	Gln	Met	Val	Val	Glu	His	Lys	His	Lys	Met	Gly		
	210					215					220						
Phe	Gln	Gly	Thr	Leu	Leu	Ile	Glu	Pro	Lys	Pro	Gln	Glu	Pro	Thr	Lys		
225					230					235					240		
His	Gln	Tyr	Asp	Tyr	Asp	Val	Ala	Thr	Val	Tyr	Gly	Phe	Leu	Lys	Gln		
				245					250					255			
Phe	Gly	Leu	Glu	Lys	Glu	Ile	Lys	Val	Asn	Ile	Glu	Ala	Asn	His	Ala		
			260					265					270				

Thr Leu Ala Gly His Ser Phe His His Glu Ile Ala Thr Ala Ile Ala
275 280 285

Leu Gly Ile Phe Gly Ser Val Asp Ala Asn Arg Gly Asp Ala Gln Leu
290 295 300

Gly Trp Asp Thr Asp Gln Phe Pro Ile Ser Val Glu Glu Asn Ala Leu
305 310 315 320

Val Met Tyr Glu Ile Leu Lys Ala Gly Gly Phe Thr Thr Gly Gly Leu
325 330 335

Asn Phe Asp Ala Lys Val Arg Arg Gln Ser Thr Asp Lys Tyr Asp Leu
340 345 350

Phe Tyr Gly His Ile Gly Ala Met Asp Thr Met Ala Leu Ser Leu Lys
355 360 365

Ile Ala Ala Arg Met Val Glu Asp Gly Glu Leu Asp Lys Arg Val Ala
370 375 380

Lys Arg Tyr Ala Gly Trp Asn Gly Glu Leu Gly Gln Gln Ile Leu Lys
385 390 395 400

Gly Gln Leu Ser Leu Gly Glu Leu Ala Gln Tyr Ala Glu Gln His His
405 410 415

Leu Ala Pro Val His Gln Ser Gly His Gln Glu Leu Leu Glu Asn Leu
420 425 430

Val Asn Arg Tyr Leu Phe Asp Lys
435 440

<210> 32

<211> 391

<212> PRT

<213> *Salmonella typhimurium*

<400> 32

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Ile Ala Asn Pro Gln Gln Gln Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Arg Ile
35 40 45

5

10

ES 2 696 235 T3

Thr	Thr	Ala	Asn	Gly	Glu	Thr	Val	Ser	Leu	Arg	Asp	Ala	Ile	Glu	Lys	50	55	60	
Asn	Lys	Thr	Ala	Met	Leu	Gly	Glu	Ala	Val	Ala	Asn	Arg	Phe	Gly	Glu	65	70	75	80
Leu	Pro	Phe	Leu	Phe	Lys	Val	Leu	Cys	Ala	Ala	Lys	Pro	Leu	Ser	Ile	85	90	95	
Gln	Val	His	Pro	Asn	Lys	Arg	Asn	Ser	Glu	Ile	Gly	Phe	Ala	Lys	Glu	100	105	110	
Asn	Ala	Ala	Gly	Ile	Pro	Met	Asp	Ala	Ala	Glu	Arg	Asn	Tyr	Lys	Asp	115	120	125	
Pro	Asn	His	Lys	Pro	Glu	Leu	Val	Phe	Ala	Leu	Thr	Pro	Phe	Leu	Ala	130	135	140	
Met	Asn	Ala	Phe	Arg	Glu	Phe	Ser	Asp	Ile	Val	Ser	Leu	Leu	Gln	Pro	145	150	155	160
Val	Ala	Gly	Ala	His	Ser	Ala	Ile	Ala	His	Phe	Leu	Gln	Val	Pro	Asn	165	170	175	
Ala	Glu	Arg	Leu	Ser	Gln	Leu	Phe	Ala	Ser	Leu	Leu	Asn	Met	Gln	Gly	180	185	190	
Glu	Glu	Lys	Ser	Arg	Ala	Leu	Ala	Val	Leu	Lys	Ala	Ala	Leu	Asn	Ser	195	200	205	
Gln	Gln	Gly	Glu	Pro	Trp	Gln	Thr	Ile	Arg	Val	Ile	Ser	Glu	Tyr	Tyr	210	215	220	
Pro	Asp	Asp	Ser	Gly	Leu	Phe	Ser	Pro	Leu	Leu	Leu	Asn	Val	Val	Lys	225	230	235	240
Leu	Asn	Pro	Gly	Glu	Ala	Met	Phe	Leu	Phe	Ala	Glu	Thr	Pro	His	Ala	245	250	255	
Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn	Ser	Asp	Asn	Val	260	265	270	
Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Tyr	Ile	Asp	Ile	Pro	Glu	Leu	Val	275	280	285	
Ala	Asn	Val	Lys	Phe	Glu	Pro	Lys	Pro	Ala	Gly	Glu	Leu	Leu	Thr	Ala	290	295	300	

Pro Val Lys Ser Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Asp Asp
305 310 315 320

Phe Ala Phe Ser Leu His Asp Leu Ala Leu Gln Glu Thr Ser Ile Gly
325 330 335

Gln His Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Glu Ala Val Leu
340 345 350

Arg Lys Asp Glu Gln Arg Leu Val Leu Lys Pro Gly Glu Ser Ala Phe
355 360 365

Ile Gly Ala Asp Glu Ser Pro Val Asn Ala Ser Gly Thr Gly Arg Leu
370 375 380

Ala Arg Val Tyr Asn Lys Leu
385 390

<210> 33

<211> 391

<212> PRT

<213> *Enterobacter cloacae*

<400> 33

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Ile Ala Asn Pro Asp Asn Leu Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Lys Val
35 40 45

Glu Asp Ala Ser Gly Gln Val Arg Ala Leu Arg Asp Val Ile Asp Ala
50 55 60

Asp Lys Ala Thr Leu Leu Gly Asn Ser Val Ala Ser Arg Phe Gly Glu
65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Asp Gln Pro Leu Ser Ile
85 90 95

Gln Val His Pro Asn Lys Lys Ala Ser Glu Ile Gly Phe Ala Lys Glu
100 105 110

Asn Ala Ala Gly Ile Pro Leu Asp Ala Ala Glu Arg Asn Tyr Lys Asp
115 120 125

5

10

Pro	Asn	His	Lys	Pro	Glu	Leu	Val	Phe	Ala	Leu	Thr	Pro	Phe	Leu	Ala
130						135					140				
Met	Asn	Ala	Phe	Arg	Glu	Phe	Ser	Glu	Ile	Ile	Ser	Leu	Leu	Gln	Pro
145					150					155					160
Val	Ala	Gly	Ala	His	Asn	Ala	Ile	Ala	His	Phe	Leu	Ala	Asn	Pro	Asn
				165					170					175	
Ala	Glu	Ser	Leu	Ser	Gln	Leu	Phe	Ala	Ser	Leu	Leu	Asn	Met	Gln	Gly
			180					185					190		
Asp	Glu	Lys	Ser	His	Ala	Arg	Ala	Val	Leu	Lys	Ala	Ala	Leu	Asp	Gly
		195					200					205			
Gln	Gln	Gly	Glu	Pro	Trp	Asp	Thr	Ile	Arg	Leu	Ile	Ser	Ala	Phe	Tyr
	210					215					220				
Pro	Asp	Asp	Ser	Gly	Leu	Phe	Ser	Pro	Leu	Leu	Leu	Asn	Val	Val	Lys
225					230					235					240
Leu	Asn	Pro	Gly	Glu	Ala	Met	Phe	Leu	Phe	Ala	Glu	Thr	Pro	His	Ala
				245					250					255	
Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn	Ser	Asp	Asn	Val
			260					265					270		
Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Tyr	Ile	Asp	Ile	Pro	Glu	Leu	Val
		275					280					285			
Ala	Asn	Val	Lys	Phe	Val	Ala	Lys	Pro	Ala	Ala	Glu	Leu	Leu	Thr	Gln
	290					295					300				
Pro	Val	Lys	Asn	Gly	Ala	Glu	Leu	Asp	Phe	Pro	Ile	Pro	Val	Asp	Asp
305					310					315					320
Phe	Ala	Phe	Ser	Leu	His	Asp	Leu	Ser	Gln	Thr	Glu	Thr	Thr	Ile	Ala
				325					330					335	
Gln	Glu	Ser	Ala	Ala	Ile	Leu	Phe	Cys	Val	Glu	Gly	Glu	Ala	Thr	Leu
			340					345					350		
His	Lys	Gly	Glu	Gln	Arg	Leu	Val	Leu	Lys	Pro	Gly	Glu	Ser	Ala	Phe
		355					360					365			
Val	Ala	Ala	Asn	Glu	Ser	Pro	Val	Ser	Val	Ser	Gly	Thr	Gly	Arg	Leu

370

375

380

Ala Arg Val Phe Asn Lys Leu
385 390

5

<210> 34
<211> 461
<212> PRT
<213> *Enterobacter cloacae*

<400> 34

Met Ser Ile His Tyr Tyr Ser Cys Ile Asn Asp Ser Pro His Thr Leu
1 5 10 15

Ile Met Glu Leu Asn Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val
20 25 30

Arg Tyr Glu Gly Pro Lys Thr Thr Asn Pro Leu Ala Phe Arg His Tyr
35 40 45

Asn Pro Asp Glu Leu Val Leu Gly Lys Arg Met Glu Asp His Leu Arg
50 55 60

Phe Ala Ala Cys Tyr Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met
65 70 75 80

Phe Gly Val Gly Ser Phe Asp Arg Pro Trp Gln Gln Pro Gly Glu Ala
85 90 95

Ile Glu Leu Ala Lys Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His
100 105 110

Lys Leu Asn Val Pro Tyr Tyr Cys Phe His Asp Val Asp Val Ser Pro
115 120 125

Glu Gly Ala Ser Leu Lys Glu Tyr Leu Asn Asn Phe Ala Gln Met Val
130 135 140

Asp Val Leu Ala Ala Lys Gln Gln Gln Ser Gly Val Lys Leu Leu Trp
145 150 155 160

Gly Thr Ala Asn Cys Phe Thr Asn Pro Arg Tyr Gly Ala Gly Ala Ala
165 170 175

Thr Asn Pro Asp Pro Glu Val Phe Ser Trp Ala Ala Thr Gln Val Val
180 185 190

Thr Ala Met Asn Ala Thr His Gln Leu Gly Gly Glu Asn Tyr Val Leu

10

ES 2 696 235 T3

195					200					205					
Trp	Gly	Gly	Arg	Glu	Gly	Tyr	Glu	Thr	Leu	Leu	Asn	Thr	Asp	Leu	Arg
	210					215					220				
Gln	Glu	Arg	Glu	Gln	Ile	Gly	Arg	Phe	Met	Gln	Met	Val	Val	Glu	His
225					230					235					240
Lys	His	Lys	Ile	Gly	Phe	Arg	Gly	Thr	Leu	Leu	Ile	Glu	Pro	Lys	Pro
				245					250					255	
Gln	Glu	Pro	Thr	Lys	His	Gln	Tyr	Asp	Tyr	Asp	Val	Ala	Thr	Val	Tyr
			260					265					270		
Gly	Phe	Leu	Lys	Gln	Phe	Gly	Leu	Glu	Lys	Glu	Ile	Lys	Val	Asn	Ile
		275					280					285			
Glu	Ala	Asn	His	Ala	Thr	Leu	Ala	Gly	His	Ser	Phe	His	His	Glu	Ile
	290					295					300				
Ala	Ser	Ala	Ile	Ala	Leu	Gly	Ile	Phe	Gly	Ser	Val	Asp	Ala	Asn	Arg
305					310					315					320
Gly	Asp	Ala	Gln	Leu	Gly	Trp	Asp	Thr	Asp	Gln	Phe	Pro	Asn	Ser	Val
				325					330					335	
Glu	Glu	Asn	Ala	Leu	Val	Met	Tyr	Glu	Ile	Ile	Lys	Ala	Gly	Gly	Phe
			340					345					350		
Thr	Thr	Gly	Gly	Leu	Asn	Phe	Asp	Ala	Lys	Val	Arg	Arg	Gln	Ser	Thr
		355					360					365			
Asp	Lys	Tyr	Asp	Leu	Phe	Tyr	Gly	His	Ile	Gly	Ala	Met	Asp	Thr	Met
	370					375					380				
Ala	Leu	Ala	Leu	Lys	Val	Ala	Ala	Arg	Met	Ile	Glu	Asp	Gly	Glu	Leu
385					390					395					400
Asp	Lys	Arg	Val	Ala	Lys	Arg	Tyr	Ser	Gly	Trp	Asn	Ser	Glu	Leu	Gly
				405					410					415	
Gln	Gln	Ile	Leu	Lys	Gly	Gln	Leu	Ser	Leu	Ala	Glu	Ile	Ala	Lys	Tyr
			420					425					430		
Ala	Glu	Gln	His	Gln	Leu	Ala	Pro	Gln	His	Gln	Ser	Gly	His	Gln	Glu
		435					440					445			

Leu Leu Glu Asn Leu Val Asn His Tyr Leu Phe Asp Lys
 450 455 460

<210> 35<211> 391

<212> PRT

5 <213> *Enterobacter cancerogenus*

<400> 35

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
 1 5 10 15

Thr Ala Leu Thr Asp Leu Tyr Gly Ile Ala Asn Pro Asp Gln Leu Pro
 20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Lys Ile
 35 40 45

Glu Asp Ala Ser Gly Gln Val Arg Ser Leu Arg Asp Ala Ile Asp Ala
 50 55 60

Asp Lys Thr Ala Leu Leu Gly Asp Lys Val Ala Ser Arg Phe Gly Glu
 65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Asp Gln Pro Leu Ser Ile
 85 90 95

Gln Val His Pro Asn Lys Lys Ala Ser Glu Ile Gly Phe Ala Lys Glu
 100 105 110

Asn Ala Ala Gly Ile Pro Leu Asp Ala Ala Glu Arg Asn Tyr Lys Asp
 115 120 125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
 130 135 140

Met Asn Ala Phe Arg Glu Phe Ser Glu Ile Ile Ser Leu Leu Gln Pro
 145 150 155 160

Val Ala Gly Ala Asn Asn Ala Ile Ala His Phe Leu Glu Asn Pro Asn
 165 170 175

Ala Glu Gly Leu Ser Gln Leu Phe Ala Ser Leu Leu Asn Met Gln Gly
 180 185 190

Glu Glu Lys Ser His Ala Leu Ala Ile Leu Lys Ala Ala Leu Asn Asp
 195 200 205

Gln Gln Gly Glu Pro Trp Asp Thr Ile Arg Val Ile Ser Glu Phe Tyr
210 215 220

Pro Asp Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
225 230 235 240

Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
245 250 255

Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val
260 265 270

Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
275 280 285

Ala Asn Val Lys Phe Val Ala Lys Pro Ala Ala Glu Leu Leu Thr Gln
290 295 300

Pro Val Lys Asn Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Asp Asp
305 310 315 320

Phe Ala Phe Ser Leu His Asp Leu Ser Ala Asp Glu Thr Ala Ile Ala
325 330 335

Gln Glu Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Glu Ala Thr Leu
340 345 350

His Lys Gly Glu Gln Arg Leu Val Leu Lys Pro Gly Glu Ser Ala Phe
355 360 365

Val Ala Ala Asn Glu Ser Pro Val Ser Ile Ser Gly Lys Gly Arg Leu
370 375 380

Ala Arg Val Phe Asn Lys Leu
385 390

<210> 36

<211> 392

<212> PRT

<213> *Cronobacter turicensis*

<400> 36

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Val Ala Asn Pro Glu Gly Leu Pro
20 25 30

ES 2 696 235 T3

Met	Ala	Glu	Leu	Trp	Met	Gly	Ala	His	Pro	Lys	Ser	Ser	Ser	Lys	Ile
		35					40					45			
Gln	Asp	Asp	Gln	Gly	Gln	Thr	Arg	Ala	Leu	Arg	Glu	Val	Ile	Asp	Ala
	50					55					60				
Asp	Lys	Thr	Ala	Leu	Leu	Gly	Ala	Pro	Val	Ala	Glu	Arg	Phe	Gly	Glu
65					70					75					80
Leu	Pro	Phe	Leu	Phe	Lys	Val	Leu	Cys	Ala	Asp	Gln	Pro	Leu	Ser	Ile
			85						90					95	
Gln	Val	His	Pro	Asn	Lys	Gln	Ala	Ser	Glu	Glu	Gly	Phe	Ala	Arg	Glu
			100					105					110		
Asn	Ala	Ala	Gly	Ile	Pro	Leu	Asp	Ala	Ala	Glu	Arg	Asn	Tyr	Lys	Asp
	115						120					125			
Pro	Asn	His	Lys	Pro	Glu	Leu	Val	Phe	Ala	Leu	Thr	Pro	Phe	Leu	Ala
	130					135					140				
Met	Asn	Ala	Phe	Arg	Glu	Phe	Ser	Glu	Ile	Val	Ser	Leu	Leu	Gln	Pro
145					150					155					160
Val	Ala	Gly	Ala	His	Thr	Ala	Ile	Ala	His	Phe	Leu	Ala	Glu	Pro	Asn
				165					170					175	
Ala	Asp	Arg	Leu	Arg	Asp	Leu	Phe	Ala	Gly	Leu	Leu	Asn	Met	Gln	Gly
			180					185					190		
Glu	Glu	Lys	Ser	Arg	Ala	Leu	Ala	Val	Leu	Lys	Ala	Thr	Val	Ala	Ser
		195					200					205			
Gln	Gln	Gly	Glu	Pro	Trp	Asp	Thr	Ile	Arg	Phe	Ile	Ala	Gln	Phe	Tyr
	210					215					220				
Pro	Asp	Asp	Ser	Gly	Leu	Phe	Ser	Pro	Leu	Leu	Leu	Asn	Val	Val	Lys
225					230					235					240
Leu	Asn	Pro	Gly	Glu	Ala	Met	Phe	Leu	Phe	Ala	Glu	Thr	Pro	His	Ala
				245					250					255	
Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn	Ser	Asp	Asn	Val
			260					265					270		
Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Tyr	Ile	Asp	Ile	Pro	Glu	Leu	Val
		275					280					285			

Ala Asn Val Lys Phe Val Ala Lys Pro Ala Ala Gln Leu Leu Thr Gln
290 295 300

Pro Glu Lys Asp Gly Ala Ala Leu Glu Phe Pro Ile Pro Val Glu Asp
305 310 315 320

Phe Ala Phe Ser Leu His Asp Leu Gln Ser Gln Pro Gln Arg Leu Ala
325 330 335

Gln Glu Ser Ala Ala Ile Leu Phe Cys Val Glu Gly Glu Ala Val Leu
340 345 350

Ser Lys Gly Asp Glu Arg Leu Thr Leu Lys Pro Gly Glu Ser Ala Phe
355 360 365

Val Ala Ala Asn Glu Ser Pro Val Asn Val Ser Gly Val Gly Arg Leu
370 375 380

Ala Arg Val Tyr Asn Lys Leu Asn
385 390

<210> 37

<211> 392

<212> PRT

<213> *Cronobacter sakazakii*

<400> 37

Met Gln Lys Leu Ile Asn Ser Val Gln Asn Tyr Ala Trp Gly Ser Lys
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Val Ala Asn Pro Glu Gly Leu Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Lys Ile
35 40 45

Gln Asp Ala Gln Gly Gln Thr Arg Ala Leu Arg Glu Val Ile Asp Ala
50 55 60

Asp Lys Thr Ala Leu Leu Gly Ala Pro Val Ala Glu Arg Phe Gly Glu
65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Asp Gln Pro Leu Ser Ile
85 90 95

Gln Val His Pro Asn Lys Gln Ala Ser Glu Glu Gly Phe Ala Arg Glu
100 105 110

5

10

ES 2 696 235 T3

Asn	Ala	Ala	Gly	Ile	Pro	Leu	Ser	Ala	Ala	Glu	Arg	Asn	Tyr	Lys	Asp
		115					120					125			
Pro	Asn	His	Lys	Pro	Glu	Leu	Val	Phe	Ala	Leu	Thr	Pro	Phe	Leu	Ala
	130					135					140				
Met	Asn	Ala	Phe	Arg	Glu	Phe	Ser	Glu	Ile	Val	Ser	Leu	Leu	Gln	Pro
145					150					155					160
Val	Ala	Gly	Ala	His	Thr	Ala	Ile	Ala	His	Phe	Leu	Ala	Glu	Pro	Asn
				165					170					175	
Ala	Asp	Arg	Leu	Arg	Glu	Leu	Phe	Ala	Gly	Leu	Leu	Asn	Met	Gln	Gly
			180					185					190		
Glu	Glu	Lys	Ser	Arg	Ala	Leu	Ala	Val	Leu	Lys	Ala	Thr	Val	Ala	Ser
		195					200					205			
Gln	Gln	Gly	Glu	Pro	Trp	Glu	Thr	Ile	Arg	Phe	Ile	Ala	Gln	Phe	Tyr
	210					215					220				
Pro	Asp	Asp	Ser	Gly	Leu	Phe	Ser	Pro	Leu	Leu	Leu	Asn	Val	Val	Lys
225					230					235					240
Leu	Asn	Pro	Gly	Glu	Ala	Met	Phe	Leu	Phe	Ala	Glu	Thr	Pro	His	Ala
			245						250					255	
Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn	Ser	Asp	Asn	Val
			260					265					270		
Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Tyr	Ile	Asp	Ile	Pro	Glu	Leu	Val
		275					280					285			
Ala	Asn	Val	Lys	Phe	Val	Ala	Lys	Pro	Ala	Gly	Glu	Leu	Leu	Thr	Gln
	290					295					300				
Pro	Val	Lys	Gln	Glu	Ser	Ala	Leu	Glu	Phe	Pro	Ile	Pro	Val	Glu	Asp
305					310					315					320
Phe	Ala	Phe	Ser	Leu	His	Asp	Leu	Asn	Ser	Gln	Pro	Gln	Ser	Leu	Ala
				325					330					335	
Gln	Glu	Ser	Ala	Ala	Ile	Leu	Phe	Cys	Val	Glu	Gly	Glu	Ala	Val	Leu
			340					345					350		
Ser	Lys	Gly	Asp	Glu	Arg	Leu	Thr	Leu	Lys	Pro	Gly	Glu	Ser	Ala	Phe
		355					360					365			

Val Ala Ala Ser Glu Ser Pro Val Arg Ala Ser Gly Val Gly Arg Leu
370 375 380

Ala Arg Val Tyr Asn Lys Leu Asn
385 390

<210> 38

<211> 392

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 38

Met Gln Lys Leu Ile Asn Ala Val Gln Asn Tyr Ala Trp Gly Ser His
1 5 10 15

Thr Ala Leu Thr Glu Leu Tyr Gly Ile Ala Asn Pro Asp Asn Leu Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Gln Ile
35 40 45

Leu Ala Ala Asp Gly Gln Pro Arg Ser Leu Arg Glu Val Ile Asp Ala
50 55 60

Asp Lys Thr Ala Leu Leu Gly Asp Lys Val Ala Ala Arg Phe Gly Glu
65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Ala Gln Pro Leu Ser Ile
85 90 95

Gln Val His Pro Asn Lys Gln Ala Ser Glu Glu Gly Phe Ala Arg Glu
100 105 110

Asn Ala Ala Gly Ile Pro Leu Ser Ala Ala Glu Arg Asn Tyr Lys Asp
115 120 125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
130 135 140

Met Asn Ala Phe Arg Glu Phe Ser Glu Ile Val Thr Leu Leu Gln Pro
145 150 155 160

Val Ala Ser Ala His Pro Ala Ile Gly Ala Phe Leu Gln Gln Pro Asp
165 170 175

Ala Thr His Leu Ser Gln Leu Phe Ala Ser Leu Leu Asn Met Gln Gly
180 185 190

ES 2 696 235 T3

Glu Glu Lys Ala Lys Ala Leu Gln Val Leu Arg Asp Val Leu Ala Arg
195 200 205

Glu Gln Gly Glu Pro Trp Gln Thr Ile Arg Leu Ile Ala Glu Phe Tyr
210 215 220

Pro Asp Asp Ser Gly Leu Phe Ser Pro Leu Leu Leu Asn Val Val Lys
225 230 235 240

Leu Asn Pro Gly Glu Ala Met Phe Leu Phe Ala Glu Thr Pro His Ala
245 250 255

Tyr Leu Gln Gly Val Ala Leu Glu Val Met Ala Asn Ser Asp Asn Val
260 265 270

Leu Arg Ala Gly Leu Thr Pro Lys Tyr Ile Asp Ile Pro Glu Leu Val
275 280 285

Ala Asn Val Lys Phe Glu Ala Lys Pro Ala Gly Glu Leu Leu Thr Gln
290 295 300

Pro Gln Gln His Gly Ala Glu Leu Asp Phe Pro Ile Pro Val Glu Asp
305 310 315 320

Phe Ala Phe Ser Leu His Asp Leu Ser Ala Glu Ala Ser Asp Leu Ala
325 330 335

Gln Ala Ser Ala Ala Ile Ile Phe Cys Val Asp Gly Glu Ala Val Leu
340 345 350

His Lys Gly Asp Gln Ser Leu Thr Leu Lys Pro Gly Glu Ser Ala Phe
355 360 365

Val Ala Ala Asn Glu Ser Pro Val Gln Val Ser Gly Arg Gly Arg Val
370 375 380

Ala Arg Val Phe Asn Lys Leu Gln
385 390

<210> 39

<211> 304

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 39

Met Arg Ile Gly Ile Asp Leu Gly Gly Thr Lys Thr Glu Val Ile Ala
1 5 10 15

ES 2 696 235 T3

Leu Ser Asp Gln Gly Glu Gln Leu Phe Arg His Arg Leu Pro Thr Pro
 20 25 30
 Arg Glu Asp Tyr Arg Gln Thr Ile Glu Thr Ile Ala Thr Leu Val Ala
 35 40 45
 Met Ala Glu Gln Ala Thr Gly Gln Gln Gly Thr Val Gly Met Gly Ile
 50 55 60
 Pro Gly Ala Ile Ser Pro Tyr Thr Gly Val Val Lys Asn Ala Asn Ser
 65 70 75 80
 Thr Trp Leu Asn Gly Gln Pro Phe Asp Lys Asp Leu Ser Leu Arg Leu
 85 90 95
 Glu Arg Glu Val Arg Leu Ala Asn Asp Ala Asn Cys Leu Ala Val Ser
 100 105 110
 Glu Ala Val Asp Gly Ala Ala Ala Gly Ala Gln Thr Val Phe Ala Val
 115 120 125
 Ile Ile Gly Thr Gly Cys Gly Ala Gly Val Ala Leu Asn Gly Arg Ala
 130 135 140
 His Ile Gly Gly Asn Gly Asn Ala Gly Glu Trp Gly His Asn Pro Leu
 145 150 155 160
 Pro Trp Met Asn Asp Asp Glu Leu Arg Tyr Arg Ala Glu Val Pro Cys
 165 170 175
 Tyr Cys Gly Lys Gln Gly Cys Ile Glu Thr Phe Ile Ser Gly Thr Gly
 180 185 190
 Phe Ala Thr Asp Tyr Gln Arg Leu Ser Gly Lys Gly Leu Thr Gly Ser
 195 200 205
 Glu Ile Met Arg Leu Val Gly Glu Gly Asp Glu Lys Ala Glu Leu Ala
 210 215 220
 Leu Ser Arg Tyr Glu Gln Arg Leu Ala Lys Ser Leu Ala His Val Val
 225 230 235 240
 Asn Ile Leu Asp Pro Asp Val Ile Val Leu Gly Gly Gly Met Ser Asn
 245 250 255
 Val Glu Arg Leu Tyr Gln Thr Val Pro Asp Leu Val Lys Gln Trp Val

ES 2 696 235 T3

260

265

270

Phe Gly Gly Glu Cys Glu Thr Pro Ile Arg Lys Ala Leu His Gly Asp
275 280 285

Ser Ser Gly Val Arg Gly Ala Ala Trp Leu Trp Pro Leu Gln Gly Thr
290 295 300

<210> 40

<211> 392

<212> PRT

<213> *Serratia odorifera*

<400> 40

Met Gln Lys Met Thr Asn Ala Val Gln Asn Tyr Ala Trp Gly Ser His
1 5 10 15

Glu Ala Leu Thr Lys Leu Tyr Gly Ile Ala Asn Pro Gln Gly Lys Pro
20 25 30

Met Ala Glu Leu Trp Met Gly Ala His Pro Lys Ser Ser Ser Gln Val
35 40 45

Pro Asp Ala Asn Gly Asp Leu Arg Ser Leu Arg Asp Val Ile Asn Gln
50 55 60

Asp Gln Pro Lys Gln Leu Gly Gln Asp Val Ala Thr Arg Phe Gly Glu
65 70 75 80

Leu Pro Phe Leu Phe Lys Val Leu Cys Ala Asp Gln Pro Leu Ser Ile
85 90 95

Gln Val His Pro Ser Lys Ser Ala Ala Glu Val Gly Tyr Ala Arg Glu
100 105 110

Asn Ala Ala Gly Ile Pro Leu Asp Ala Ala Glu Arg Asn Tyr Lys Asp
115 120 125

Pro Asn His Lys Pro Glu Leu Val Phe Ala Leu Thr Pro Phe Leu Ala
130 135 140

Met Asn Gly Phe Arg Glu Leu Ala Asp Ile Val Ala Leu Leu Gln Pro
145 150 155 160

Ile Ala Gly Ala His His Asp Ile Ala Ala Phe Leu Gln Gln Pro Asp
165 170 175

Thr Glu Arg Leu Ala Lys Leu Phe Ala Ser Leu Leu Gly Met Ser Gly

5

10

ES 2 696 235 T3

				180						185						190
Glu	Gln	Lys	Ser	Leu	Ala	Leu	Gly	Val	Leu	Lys	Ala	Ala	Leu	Asn	Asn	
		195					200					205				
Gln	Gln	Gly	Glu	Pro	Trp	Asp	Thr	Val	Arg	Phe	Ile	Ala	Gly	Phe	Tyr	
	210					215					220					
Pro	Glu	Asp	Ser	Gly	Leu	Phe	Ser	Pro	Leu	Leu	Leu	Asn	Val	Val	Lys	
225					230					235					240	
Leu	Ala	Pro	Gly	Glu	Ala	Met	Phe	Leu	Tyr	Ala	Glu	Thr	Pro	His	Ala	
				245					250					255		
Tyr	Leu	Gln	Gly	Val	Ala	Leu	Glu	Val	Met	Ala	Asn	Ser	Asp	Asn	Val	
		260						265					270			
Leu	Arg	Ala	Gly	Leu	Thr	Pro	Lys	Phe	Ile	Asp	Ile	Pro	Glu	Leu	Leu	
		275					280					285				
Ala	Asn	Leu	Gln	Phe	Arg	Pro	Gln	Pro	Ala	Ser	Gly	Leu	Leu	Thr	Gln	
290						295					300					
Pro	Glu	Lys	Arg	Gly	Asp	Glu	Leu	Phe	Phe	Pro	Ile	Pro	Val	Glu	Asp	
305					310					315					320	
Phe	Ala	Phe	Ser	Leu	His	Asp	Leu	Ser	Ser	Ser	Ala	Gln	Ser	Leu	Ala	
				325					330					335		
Gln	Arg	Ser	Ala	Ala	Ile	Val	Phe	Cys	Val	Glu	Gly	Glu	Ala	Arg	Leu	
			340					345					350			
Glu	Lys	Gln	Gly	Gln	Gln	Ile	Thr	Leu	His	Pro	Gly	Glu	Ser	Cys	Phe	
		355					360					365				
Ile	Gly	Ala	Phe	Glu	Ser	Pro	Val	Ser	Val	Ser	Gly	Ser	Gly	Arg	Ile	
	370					375					380					
Ala	Arg	Val	Tyr	Asn	Gln	Leu	Ser									
385					390											

<210> 41
 <211> 491
 <212> PRT
 <213> *Escherichia fergusonii*

 <400> 41

Met Ser Thr Gln Tyr Asn Ser Ser Tyr Ile Phe Ser Ile Thr Leu Val

ES 2 696 235 T3

1		5		10		15											
Ala	Thr	Leu	Gly	Gly	Leu	Leu	Phe	Gly	Tyr	Asp	Thr	Ala	Val	Ile	Ser		
		20						25					30				
Gly	Thr	Val	Glu	Ser	Leu	Asn	Thr	Val	Phe	Val	Ala	Pro	Gln	Asn	Leu		
		35					40					45					
Ser	Glu	Ser	Ala	Ala	Asn	Ser	Leu	Leu	Gly	Phe	Cys	Val	Ala	Ser	Ala		
	50					55					60						
Leu	Ile	Gly	Cys	Ile	Ile	Gly	Gly	Ala	Leu	Gly	Gly	Tyr	Cys	Ser	Asn		
65					70					75					80		
Arg	Phe	Gly	Arg	Arg	Asp	Ser	Leu	Lys	Ile	Ala	Ala	Val	Leu	Phe	Phe		
				85					90					95			
Ile	Ser	Gly	Val	Gly	Ser	Ala	Trp	Pro	Glu	Leu	Gly	Phe	Thr	Ser	Ile		
			100					105					110				
Asn	Pro	Asp	Asn	Thr	Val	Pro	Val	Tyr	Leu	Ala	Gly	Tyr	Val	Pro	Glu		
		115					120					125					
Phe	Val	Ile	Tyr	Arg	Ile	Ile	Gly	Gly	Ile	Gly	Val	Gly	Leu	Ala	Ser		
	130					135					140						
Met	Leu	Ser	Pro	Met	Tyr	Ile	Ala	Glu	Leu	Ala	Pro	Ala	His	Ile	Arg		
145					150					155					160		
Gly	Lys	Leu	Val	Ser	Phe	Asn	Gln	Phe	Ala	Ile	Ile	Phe	Gly	Gln	Leu		
				165					170					175			
Leu	Val	Tyr	Cys	Val	Asn	Tyr	Phe	Ile	Ala	Arg	Ser	Gly	Asp	Ala	Thr		
			180					185					190				
Trp	Leu	Asn	Thr	Asp	Gly	Trp	Arg	Tyr	Met	Phe	Ala	Ser	Glu	Cys	Ile		
		195					200					205					
Pro	Ala	Leu	Leu	Phe	Leu	Met	Leu	Leu	Tyr	Thr	Val	Pro	Glu	Ser	Pro		
	210					215					220						
Arg	Trp	Leu	Met	Ser	Arg	Gly	Lys	Gln	Glu	Gln	Ala	Glu	Gly	Ile	Leu		
225					230					235					240		
Arg	Lys	Ile	Met	Gly	Asn	Thr	Leu	Ala	Thr	Gln	Ala	Val	Gln	Glu	Ile		
				245					250					255			

Lys His Ser Leu Asp His Gly Arg Lys Thr Gly Gly Arg Leu Leu Met
 260 265 270
 Phe Gly Val Gly Val Ile Val Ile Gly Val Met Leu Ser Ile Phe Gln
 275 280 285
 Gln Phe Val Gly Ile Asn Val Val Leu Tyr Tyr Ala Pro Glu Val Phe
 290 295 300
 Lys Thr Leu Gly Ala Ser Thr Asp Ile Ala Leu Leu Gln Thr Ile Ile
 305 310 315 320
 Val Gly Val Ile Asn Leu Thr Phe Thr Val Leu Ala Ile Met Thr Val
 325 330 335
 Asp Lys Phe Gly Arg Lys Pro Leu Gln Ile Ile Gly Ala Leu Gly Met
 340 345 350
 Ala Ile Gly Met Phe Ser Leu Gly Thr Ala Phe Tyr Thr Gln Ala Pro
 355 360 365
 Gly Ile Val Ala Leu Leu Ser Met Leu Phe Tyr Val Ala Ala Phe Ala
 370 375 380
 Met Ser Trp Gly Pro Val Cys Trp Val Leu Leu Ser Glu Ile Phe Pro
 385 390 395 400
 Asn Ala Ile Arg Gly Lys Ala Leu Ala Ile Ala Val Ala Ala Gln Trp
 405 410 415
 Leu Ala Asn Tyr Phe Val Ser Trp Thr Phe Pro Met Met Asp Lys Asn
 420 425 430
 Ser Trp Leu Val Ala His Phe His Asn Gly Phe Ser Tyr Trp Ile Tyr
 435 440 445
 Gly Cys Met Gly Val Leu Ala Ala Leu Phe Met Trp Lys Phe Val Pro
 450 455 460
 Glu Thr Lys Gly Lys Thr Leu Glu Glu Leu Glu Ala Leu Trp Glu Pro
 465 470 475 480
 Glu Thr Lys Lys Thr Gln Gln Thr Ala Thr Leu
 485 490

<210> 42
 <211> 490
 <212> PRT

<213> *Shigella flexneri*

<400> 42

```

Met Asn Thr Gln Tyr Asn Ser Ser Tyr Ile Phe Ser Ile Thr Leu Val
 1          5          10          15

Ala Thr Leu Gly Gly Leu Leu Phe Gly Tyr Asp Thr Ala Val Ile Ser
      20          25          30

Gly Thr Val Glu Ser Leu Asn Thr Val Phe Val Ala Pro Gln Asn Leu
      35          40          45

Ser Glu Ser Ala Ala Asn Ser Leu Leu Gly Phe Cys Val Ala Ser Ala
 50          55          60

Leu Ile Gly Cys Ile Ile Gly Gly Ala Leu Gly Gly Tyr Cys Ser Asn
65          70          75          80

Arg Phe Gly Arg Arg Asp Ser Leu Lys Ile Ser Ala Val Leu Phe Phe
      85          90          95

Ile Ser Gly Val Gly Ser Ala Trp Pro Glu Leu Gly Phe Thr Ser Ile
      100          105          110

Asn Pro Asp Asn Thr Val Pro Ile Tyr Leu Ala Gly Tyr Val Pro Glu
      115          120          125

Phe Val Ile Tyr Arg Ile Ile Gly Gly Ile Gly Val Gly Leu Ala Ser
      130          135          140

Met Leu Ser Pro Met Tyr Ile Ala Glu Leu Ala Pro Ala His Ile Arg
145          150          155          160

Gly Lys Leu Val Ser Phe Asn Gln Phe Ala Ile Ile Phe Gly Gln Leu
      165          170          175

Leu Val Tyr Cys Val Asn Tyr Phe Ile Ala Arg Ser Gly Asp Ala Ser
      180          185          190

Trp Leu Asn Thr Asp Gly Arg Arg Tyr Met Phe Ala Ser Glu Cys Ile
      195          200          205

Pro Ala Leu Leu Phe Leu Met Leu Leu Tyr Thr Val Pro Glu Ser Pro
      210          215          220

Arg Trp Leu Met Ser Arg Gly Lys Gln Glu Gln Thr Glu Ser Ile Leu
225          230          235          240

```

ES 2 696 235 T3

Arg Lys Ile Met Gly Asn Thr Leu Ala Thr Gln Ala Val Gln Glu Ile
 245 250 255
 Lys His Ser Leu Asp His Gly Arg Lys Thr Gly Gly Arg Leu Leu Met
 260 265 270
 Phe Gly Val Gly Val Ile Val Ile Gly Val Met Leu Ser Ile Phe Gln
 275 280 285
 Gln Phe Val Gly Ile Asn Val Val Leu Tyr Tyr Ala Pro Glu Val Phe
 290 295 300
 Lys Thr Leu Gly Ala Ser Thr Asp Ile Ala Leu Leu Gln Thr Ile Ile
 305 310 315 320
 Val Gly Val Ile Asn Leu Thr Phe Thr Val Leu Ala Ile Met Thr Val
 325 330 335
 Asp Lys Phe Gly Arg Lys Pro Leu Gln Ile Ile Gly Ala Leu Gly Met
 340 345 350
 Ala Ile Gly Met Phe Ser Leu Gly Thr Ala Phe Tyr Thr Gln Ala Ser
 355 360 365
 Gly Ile Val Ala Leu Leu Ser Met Leu Phe Tyr Val Ala Ala Phe Ala
 370 375 380
 Met Ser Trp Gly Pro Val Cys Trp Val Leu Leu Ser Glu Ile Phe Pro
 385 390 395 400
 Asn Ala Ile Arg Gly Lys Ala Leu Ala Ile Ala Val Ala Ala Gln Trp
 405 410 415
 Leu Ala Asn Tyr Phe Val Ser Trp Thr Phe Pro Met Met Asp Lys Asn
 420 425 430
 Ser Trp Leu Val Ala His Phe His Asn Gly Phe Ser Tyr Trp Ile Tyr
 435 440 445
 Gly Cys Met Gly Val Leu Ala Ala Leu Phe Met Trp Lys Phe Val Pro
 450 455 460
 Glu Thr Lys Gly Lys Thr Leu Glu Glu Leu Glu Ala Leu Trp Glu Pro
 465 470 475 480
 Glu Thr Lys Lys Thr Gln Gln Thr Ala Thr
 485 490

ES 2 696 235 T3

<211> 491
 <212> PRT
 <213> *Salmonella enterica*

5 <400> 43

```

Met Asn Thr His Tyr Asn Ser Arg Tyr Ile Phe Ser Ile Thr Leu Val
1          5          10          15

Ala Thr Leu Gly Gly Leu Leu Phe Gly Tyr Asp Thr Ala Val Ile Ser
20          25          30

Gly Thr Val Glu Ser Leu Asn Thr Val Phe Val Ala Pro Gln His Leu
35          40          45

Ser Glu Ser Ala Ala Asn Ser Leu Leu Gly Phe Cys Val Ala Ser Ala
50          55          60

Leu Ile Gly Cys Ile Ile Gly Gly Ala Leu Gly Gly Tyr Cys Ser Asn
65          70          75          80

Arg Phe Gly Arg Arg Asp Ser Leu Lys Ile Ala Ala Leu Leu Phe Phe
85          90          95

Ile Ser Gly Ile Gly Ser Ala Trp Pro Glu Leu Gly Phe Thr Thr Ile
100         105         110

Asn Pro Asp Asn Ala Val Pro Val Tyr Leu Ala Gly Tyr Val Pro Glu
115         120         125

Phe Val Ile Tyr Arg Ile Ile Gly Gly Ile Gly Val Gly Leu Ala Ser
130         135         140

Met Leu Ser Pro Met Tyr Ile Ala Glu Leu Ala Pro Ala His Ile Arg
145         150         155         160

Gly Lys Leu Val Ser Phe Asn Gln Phe Ala Ile Ile Phe Gly Gln Leu
165         170         175

Leu Val Tyr Cys Val Asn Tyr Phe Ile Ala Arg Ser Gly Asp Ala Asn
180         185         190

Trp Leu Asn Thr Asp Gly Trp Arg Tyr Met Phe Ala Ser Glu Ser Ile
195         200         205

Pro Ala Leu Leu Phe Leu Leu Leu Tyr Thr Val Pro Glu Ser Pro
210         215         220
  
```

ES 2 696 235 T3

Arg	Trp	Leu	Met	Ala	Arg	Gly	Lys	His	Glu	Gln	Ala	Glu	Gly	Ile	Leu
225					230					235					240
Arg	Lys	Ile	Met	Gly	Ser	Ser	Leu	Thr	Thr	Gln	Ala	Met	Gln	Glu	Ile
				245					250					255	
Asn	Gln	Ser	Leu	Glu	His	Gly	Arg	Lys	Thr	Gly	Gly	Arg	Leu	Leu	Met
			260					265					270		
Phe	Gly	Val	Gly	Val	Ile	Val	Ile	Gly	Val	Met	Leu	Ser	Val	Phe	Gln
		275					280					285			
Gln	Phe	Val	Gly	Ile	Asn	Val	Val	Leu	Tyr	Tyr	Ala	Pro	Glu	Val	Phe
	290					295					300				
Lys	Thr	Leu	Gly	Ala	Ser	Thr	Asp	Val	Ala	Leu	Leu	Gln	Thr	Ile	Ile
305					310					315					320
Val	Gly	Val	Ile	Asn	Leu	Ser	Phe	Thr	Val	Leu	Ala	Ile	Met	Thr	Val
				325					330					335	
Asp	Lys	Phe	Gly	Arg	Lys	Pro	Leu	Gln	Ile	Ile	Gly	Ala	Leu	Gly	Met
			340					345					350		
Ala	Leu	Gly	Met	Phe	Ser	Leu	Gly	Thr	Ala	Phe	Tyr	Thr	Gln	Ala	Ser
		355					360					365			
Gly	Leu	Ile	Ala	Leu	Leu	Ser	Met	Leu	Phe	Tyr	Val	Ala	Ala	Phe	Ala
	370					375					380				
Met	Ser	Trp	Gly	Pro	Val	Cys	Trp	Val	Leu	Leu	Ala	Glu	Ile	Phe	Pro
385					390					395					400
Asn	Ala	Ile	Arg	Gly	Lys	Ala	Leu	Ala	Ile	Ala	Val	Ala	Ala	Gln	Trp
				405					410					415	
Leu	Ala	Asn	Tyr	Phe	Val	Ser	Trp	Thr	Phe	Pro	Met	Met	Asp	Lys	Asn
			420					425					430		
Ser	Trp	Leu	Val	Ser	His	Phe	His	Asn	Gly	Phe	Ser	Tyr	Trp	Ile	Tyr
		435					440					445			
Gly	Cys	Met	Gly	Ile	Leu	Ala	Ala	Leu	Phe	Met	Trp	Lys	Phe	Val	Pro
	450					455					460				
Glu	Thr	Lys	Gly	Lys	Thr	Leu	Glu	Glu	Leu	Glu	Lys	Leu	Trp	Thr	Pro

465 470 475 480

Ala Glu Glu Lys Thr Pro Lys Ala Ala Ile Gln
485 490

<210> 44

<211> 491

<212> PRT

<213> *Citrobacter youngae*

<400> 44

Met Asn Thr Gln Tyr Asn Ser Ser Tyr Ile Phe Ser Ile Thr Leu Val
1 5 10 15

Ala Thr Leu Gly Gly Leu Leu Phe Gly Tyr Asp Thr Ala Val Ile Ser
20 25 30

Gly Thr Val Glu Ser Leu Asn Thr Val Phe Val Ala Pro Gln His Leu
35 40 45

Ser Glu Ser Ala Ala Asn Ser Leu Leu Gly Phe Cys Val Ala Ser Ala
50 55 60

Leu Ile Gly Cys Ile Ile Gly Gly Ala Leu Gly Gly Tyr Cys Ser Asn
65 70 75 80

Arg Phe Gly Arg Arg Asp Ser Leu Lys Ile Ala Ala Leu Leu Phe Phe
85 90 95

Ile Ser Gly Ile Gly Ser Ala Trp Pro Glu Leu Gly Phe Thr Thr Ile
100 105 110

Asn Pro Asp Asn Ala Val Pro Val Tyr Leu Ala Gly Tyr Val Pro Glu
115 120 125

Phe Val Ile Tyr Arg Ile Ile Gly Gly Ile Gly Val Gly Leu Ala Ser
130 135 140

Met Leu Ser Pro Met Tyr Ile Ala Glu Leu Ala Pro Ala His Ile Arg
145 150 155 160

Gly Lys Leu Val Ser Phe Asn Gln Phe Ala Ile Ile Phe Gly Gln Leu
165 170 175

Leu Val Tyr Cys Val Asn Tyr Phe Ile Ala Arg Ser Gly Asp Ala Asn
180 185 190

Trp Leu Asn Thr Asp Gly Trp Arg Tyr Met Phe Ala Ser Glu Ser Ile

5

10

ES 2 696 235 T3

195					200					205					
Pro	Ala	Leu	Leu	Phe	Leu	Leu	Leu	Leu	Tyr	Thr	Val	Pro	Glu	Ser	Pro
210					215					220					
Arg	Trp	Leu	Met	Ala	Arg	Gly	Lys	His	Glu	Gln	Ala	Glu	Gly	Ile	Leu
225					230					235					240
Arg	Lys	Ile	Met	Gly	Asn	Ser	Leu	Ala	Thr	Gln	Ala	Met	Gln	Glu	Ile
				245					250					255	
Asn	Gln	Ser	Leu	Glu	His	Gly	Arg	Lys	Thr	Gly	Gly	Arg	Leu	Leu	Met
			260					265					270		
Phe	Gly	Val	Gly	Val	Ile	Val	Ile	Gly	Val	Met	Leu	Ser	Ile	Phe	Gln
			275				280					285			
Gln	Phe	Val	Gly	Ile	Asn	Val	Val	Leu	Tyr	Tyr	Ala	Pro	Glu	Val	Phe
			290				295				300				
Lys	Thr	Leu	Gly	Ala	Ser	Thr	Asp	Val	Ala	Leu	Leu	Gln	Thr	Ile	Ile
305					310					315					320
Val	Gly	Val	Ile	Asn	Leu	Ser	Phe	Thr	Val	Leu	Ala	Ile	Met	Thr	Val
				325					330					335	
Asp	Lys	Phe	Gly	Arg	Lys	Pro	Leu	Gln	Ile	Ile	Gly	Ala	Leu	Gly	Met
			340					345					350		
Ala	Leu	Gly	Met	Phe	Ser	Leu	Gly	Thr	Ala	Phe	Tyr	Thr	Gln	Ala	Ser
			355				360					365			
Gly	Leu	Ile	Ala	Leu	Leu	Ser	Met	Leu	Phe	Tyr	Val	Ala	Ala	Phe	Ala
			370				375				380				
Met	Ser	Trp	Gly	Pro	Val	Cys	Trp	Val	Leu	Leu	Ala	Glu	Ile	Phe	Pro
385					390					395					400
Asn	Ala	Ile	Arg	Gly	Lys	Ala	Leu	Ala	Ile	Ala	Val	Ala	Ala	Gln	Trp
				405					410					415	
Leu	Ala	Asn	Tyr	Phe	Val	Ser	Trp	Thr	Phe	Pro	Met	Met	Asp	Lys	Asn
			420					425					430		
Ser	Trp	Leu	Val	Ser	His	Phe	His	Asn	Gly	Phe	Ser	Tyr	Trp	Ile	Tyr
			435				440					445			

Gly Cys Met Gly Ile Leu Ala Ala Leu Phe Met Trp Lys Phe Val Pro
 450 455 460

Glu Thr Lys Gly Lys Thr Leu Glu Glu Leu Glu Glu Leu Trp Glu Pro
 465 470 475 480

Ala Ala Glu Lys Ala Lys Pro Ala Ser Ala Gln
 485 490

<210> 45

<211> 484

5 <212> PRT

<213> *Acidobacterium capsulatum*

<400> 45

Met Ala Arg Leu Asn Lys Tyr Val Val Phe Leu Ala Leu Ile Ala Thr
 1 5 10 15

Phe Gly Gly Leu Leu Phe Gly Tyr Asp Thr Ala Val Ile Asn Gly Ala
 20 25 30

Val Asp Ser Leu Lys Ala Tyr Phe Ile Asn Pro Arg Phe Ser Asp Leu
 35 40 45

Ala Asn Pro Ala Gln Ala Asp Ala Ala Ser Ser Leu Leu Gly Phe Val
 50 55 60

Val Ser Ser Ala Leu Ile Gly Cys Ile Ile Gly Gly Leu Met Gly Gly
 65 70 75 80

Trp Val Ser Thr Val Ile Gly Arg Lys Arg Gly Leu Val Ile Ala Ala
 85 90 95

Val Leu Phe Leu Ile Ser Ala Leu Gly Ala Ser Ala Pro Glu Phe Pro
 100 105 110

Phe Ala Pro Ile Gly His Gly Gly Pro Ala Tyr Met Trp Asn Phe Val
 115 120 125

Ile Tyr Arg Ile Leu Gly Gly Ile Gly Val Gly Leu Ala Ser Met Leu
 130 135 140

Ser Pro Met Tyr Ile Ala Glu Ile Ala Pro Pro Lys Val Arg Gly Asn
 145 150 155 160

Leu Val Ala Trp Asn Gln Phe Ala Ile Ile Phe Gly Met Leu Val Ile
 165 170 175

Tyr Phe Val Asn Tyr Gly Ile Ser Lys Gly Gly Asn Gly Asp Ala Trp
 180 185 190
 Leu Asn Ser Ile Gly Trp Arg Tyr Met Phe Leu Ser Gly Ala Ile Pro
 195 200 205
 Ala Ser Ile Phe Leu Leu Leu Leu Leu Phe Val Pro Glu Thr Pro Arg
 210 215 220
 Tyr Leu Met Met Lys Gly Gln Glu Ala Lys Ala Arg Thr Val Leu Asp
 225 230 235 240
 Lys Leu Val Thr Lys Glu Glu Ala Asp Arg Glu Leu Arg Glu Ile Arg
 245 250 255
 Ala Ser Leu Ser Gln Asn His Ser Gly Lys Leu Phe Ser Phe Gly Ala
 260 265 270
 Phe Leu Ile Phe Ser Gly Met Leu Leu Ser Ile Phe Gln Gln Phe Val
 275 280 285
 Gly Ile Asn Val Val Leu Tyr Tyr Ala Thr Asp Ile Phe Lys Gly Met
 290 295 300
 Gly Met Ser Thr Asn Ala Ala Leu Met Gln Thr Ile Ile Val Gly Ala
 305 310 315 320
 Val Asn Leu Thr Phe Thr Val Ile Ala Ile Leu Thr Val Asp Arg Phe
 325 330 335
 Gly Arg Arg Pro Leu Gln Val Val Gly Gly Leu Ile Met Ala Ala Ser
 340 345 350
 Met Thr Trp Leu Gly Ile Glu Leu Trp Thr Gly Gly Lys Gly Leu Gly
 355 360 365
 Ala Leu Ile Ala Met Leu Val Tyr Thr Ala Gly Phe Ala Val Ser Trp
 370 375 380
 Gly Pro Val Thr Trp Val Leu Leu Ser Glu Ile Phe Pro Asn Gln Ile
 385 390 395 400
 Arg Gly Lys Ala Met Ala Ile Ala Val Ala Val Gln Trp Val Ala Asn
 405 410 415
 Tyr Leu Val Ser Trp Thr Phe Pro Ile Leu Asn Asn Asn Pro Phe Leu
 420 425 430

Val Lys His Phe His His Gly Phe Ala Tyr Trp Ile Tyr Gly Val Met
435 440 445

Ser Ile Leu Ala Ala Leu Phe Val Trp Arg Lys Val Pro Glu Thr Lys
450 455 460

Gly Arg Thr Leu Glu Gln Met Glu Ser Leu Trp Gly Ser Leu Lys Lys
465 470 475 480

Ala Ala Ile Gly

<210> 46

<211> 484

<212> PRT

<213> *Shigella flexneri*

<400> 46

Met Tyr Ile Gly Ile Asp Leu Gly Thr Ser Gly Val Lys Val Ile Leu
1 5 10 15

Leu Asn Glu Gln Gly Glu Val Val Ala Ser Gln Thr Glu Lys Leu Thr
20 25 30

Val Ser Arg Pro His Pro Leu Trp Ser Glu Gln Asp Pro Glu Gln Trp
35 40 45

Trp Gln Ala Thr Asp Arg Ala Met Lys Ala Leu Gly Asp Gln His Ser
50 55 60

Leu Gln Asp Val Lys Ala Leu Gly Ile Ala Gly Gln Met His Gly Ala
65 70 75 80

Thr Leu Leu Asp Ala Gln Gln Arg Val Leu Arg Pro Ala Ile Leu Trp
85 90 95

Asn Asp Gly Arg Cys Ala Gln Glu Cys Thr Leu Leu Glu Ala Arg Val
100 105 110

Pro Gln Ser Arg Val Ile Thr Gly Asn Leu Met Met Pro Gly Phe Thr
115 120 125

Ala Pro Lys Leu Leu Trp Val Gln Arg His Glu Pro Glu Ile Phe Arg
130 135 140

Gln Ile Asp Lys Val Leu Leu Pro Lys Asp Tyr Leu Arg Leu Arg Met
145 150 155 160

5

10

Thr	Gly	Glu	Phe	Ala	Ser	Asp	Met	Ser	Asp	Ala	Ala	Gly	Thr	Met	Trp	165	170	175	
Leu	Asp	Val	Ala	Lys	Arg	Asp	Trp	Ser	Asp	Val	Met	Leu	Gln	Ala	Cys	180	185	190	
Asp	Leu	Ser	Arg	Asp	Gln	Met	Pro	Ala	Leu	Tyr	Glu	Gly	Ser	Glu	Ile	195	200	205	
Thr	Gly	Ala	Leu	Leu	Pro	Glu	Val	Ala	Lys	Ala	Trp	Gly	Met	Ala	Thr	210	215	220	
Val	Pro	Val	Val	Ala	Gly	Gly	Gly	Asp	Asn	Ala	Ala	Gly	Ala	Val	Gly	225	230	235	240
Val	Gly	Met	Val	Asp	Ala	Asn	Gln	Ala	Met	Leu	Ser	Leu	Gly	Thr	Ser	245	250	255	
Gly	Val	Tyr	Phe	Ala	Val	Ser	Glu	Gly	Phe	Leu	Ser	Lys	Pro	Glu	Ser	260	265	270	
Ala	Val	His	Ser	Phe	Cys	His	Ala	Leu	Pro	Gln	Arg	Trp	His	Leu	Met	275	280	285	
Ser	Val	Met	Leu	Ser	Ala	Ala	Ser	Cys	Leu	Asp	Trp	Ala	Ala	Lys	Leu	290	295	300	
Thr	Gly	Leu	Cys	Asn	Val	Pro	Ala	Leu	Ile	Ala	Ala	Ala	Gln	Gln	Ala	305	310	315	320
Asp	Glu	Ser	Ala	Glu	Pro	Val	Trp	Phe	Leu	Pro	Tyr	Leu	Ser	Gly	Glu	325	330	335	
Arg	Thr	Pro	His	Asn	Asn	Pro	Gln	Ala	Lys	Gly	Val	Phe	Phe	Gly	Leu	340	345	350	
Thr	His	Gln	His	Gly	Pro	Asn	Glu	Leu	Ala	Arg	Ala	Val	Leu	Glu	Gly	355	360	365	
Val	Gly	Tyr	Ala	Leu	Ala	Asp	Gly	Met	Asp	Val	Val	His	Ala	Cys	Gly	370	375	380	
Ile	Lys	Pro	Gln	Ser	Val	Thr	Leu	Ile	Gly	Gly	Gly	Ala	Arg	Ser	Glu	385	390	395	400
Tyr	Trp	Arg	Gln	Met	Leu	Ala	Asp	Ile	Ser	Gly	Gln	Gln	Leu	Asp	Tyr	405	410	415	

Arg Thr Gly Gly Asp Val Gly Pro Ala Leu Gly Ala Ala Arg Leu Ala
420 425 430

Gln Ile Ala Ala Asn Pro Glu Lys Ser Leu Ile Glu Leu Leu Pro Gln
435 440 445

Leu Pro Leu Glu Gln Ser His Leu Pro Asp Ala Gln Arg Tyr Ala Ala
450 455 460

Tyr Gln Pro Arg Arg Glu Thr Phe Arg Arg Leu Tyr Gln Gln Leu Leu
465 470 475 480

Pro Leu Met Ala

<210> 47

<211> 484

<212> PRT

<213> *Citrobacter rodentium*

<400> 47

Met Tyr Ile Gly Ile Asp Leu Gly Thr Ser Gly Val Lys Ala Ile Leu
1 5 10 15

Leu Asn Glu Gln Gly Glu Val Val Ala Ala Gln Thr Glu Lys Leu Thr
20 25 30

Val Ser Arg Pro His Pro Leu Trp Ser Glu Gln Asp Pro Glu Gln Trp
35 40 45

Trp Leu Ala Thr Asp Arg Ala Val Lys Ala Leu Gly Lys Gln His Ser
50 55 60

Leu Ser Glu Val Lys Ala Leu Gly Ile Ala Gly Gln Met His Gly Ala
65 70 75 80

Thr Leu Leu Asp Asp Gln Gln Arg Val Leu Arg Pro Ala Ile Leu Trp
85 90 95

Asn Asp Gly Arg Cys Gly Glu Glu Cys Ala Leu Leu Glu Ala Gln Ala
100 105 110

Pro Gln Ser Arg Ala Ile Thr Gly Asn Leu Met Met Pro Gly Phe Thr
115 120 125

Ala Pro Lys Leu Leu Trp Val Lys Cys His Glu Pro Glu Ile Phe Arg
130 135 140

5

10

ES 2 696 235 T3

Gln	Val	Ala	Lys	Val	Leu	Leu	Pro	Lys	Asp	Tyr	Leu	Arg	Leu	Arg	Met	145	150	155	160
Thr	Gly	Glu	Phe	Ala	Ser	Asp	Met	Ser	Asp	Ala	Ala	Gly	Thr	Met	Trp	165	170	175	
Leu	Asp	Val	Glu	Lys	Arg	Asp	Trp	Ser	Asp	Val	Met	Leu	Glu	Ala	Cys	180	185	190	
Gly	Leu	Thr	Arg	Asp	His	Met	Pro	Ala	Leu	Phe	Glu	Gly	Ser	Asp	Ile	195	200	205	
Ser	Gly	Thr	Leu	Leu	Pro	Glu	Ile	Ala	Thr	Ala	Trp	Gly	Met	Pro	Ala	210	215	220	
Ala	Pro	Val	Val	Ala	Gly	Gly	Gly	Asp	Asn	Ala	Ala	Gly	Ala	Val	Gly	225	230	235	240
Val	Gly	Met	Ile	Asp	Ala	Asn	Gln	Ala	Met	Leu	Ser	Leu	Gly	Thr	Ser	245	250	255	
Gly	Val	Tyr	Phe	Ala	Val	Ser	Asp	Gly	Phe	Leu	Ser	Lys	Pro	Glu	Ser	260	265	270	
Ala	Val	His	Ser	Phe	Cys	His	Ala	Leu	Pro	Gly	Arg	Trp	His	Leu	Met	275	280	285	
Ser	Val	Met	Leu	Ser	Ala	Ala	Ser	Cys	Leu	Asp	Trp	Ala	Ala	Lys	Leu	290	295	300	
Thr	Gly	Leu	Glu	Thr	Val	Pro	Ala	Leu	Leu	Asp	Ala	Ala	Gln	Thr	Ala	305	310	315	320
Asp	Thr	Ser	Ala	Asp	Pro	Val	Trp	Phe	Leu	Pro	Tyr	Leu	Ser	Gly	Glu	325	330	335	
Arg	Thr	Pro	His	Asn	Asn	Pro	Gln	Ala	Lys	Gly	Val	Phe	Phe	Gly	Leu	340	345	350	
Thr	His	Gln	His	Gly	Pro	Ala	Glu	Leu	Ala	Arg	Ala	Val	Leu	Glu	Gly	355	360	365	
Val	Gly	Tyr	Ala	Leu	Ala	Asp	Gly	Met	Asp	Val	Val	His	Ala	Cys	Gly	370	375	380	
Val	Lys	Pro	Gln	Ser	Ile	Thr	Leu	Ile	Gly	Gly	Gly	Ala	Arg	Ser	Ala				

385		390		395		400
Tyr Trp Arg Gln Met Leu Ala Asp Ile Ser Gly Leu Gln Leu Asp Tyr						
		405		410		415
Arg Thr Gly Gly Asp Val Gly Pro Ala Leu Gly Ala Ala Arg Leu Ala						
		420		425		430
Gln Ile Ala Met Asn Pro Glu Lys Pro Leu Ser Asp Leu Leu Pro Gln						
		435		440		445
Leu Pro Leu Glu Gln Thr His Arg Pro Asp Ala Gln Arg His Ala Asp						
		450		455		460
Tyr Gln Gln Arg Arg Glu Thr Phe Arg Arg Leu Tyr Gln Gln Leu Leu						
		465		470		475
						480

Pro Leu Met Ser

<210> 48
 <211> 356
 <212> PRT
 <213> *Shigella flexneri*

<400> 48

Met Leu Lys Ile Gly Tyr Val Tyr Arg Gly Asp Cys Tyr Leu Leu Lys															
1			5					10					15		
Leu Ser Ser Asn Tyr Arg Arg Pro Tyr Thr Met Lys Ile Lys Asn Ile															
			20					25					30		
Leu Leu Thr Leu Cys Thr Ser Leu Leu Leu Thr Asn Val Ala Ala His															
			35					40					45		
Ala Lys Glu Val Lys Ile Gly Met Ala Ile Asp Asp Leu Arg Leu Glu															
			50				55				60				
Arg Trp Gln Lys Asp Arg Asp Ile Phe Val Lys Lys Ala Glu Ser Leu															
			65			70				75					80
Gly Ala Lys Val Phe Val Gln Ser Ala Asn Gly Asn Glu Glu Thr Gln															
						85			90					95	
Met Ser Gln Ile Glu Asn Met Ile Asn Arg Gly Val Asp Val Leu Val															
			100					105				110			
Ile Ile Pro Tyr Asn Gly Gln Val Leu Ser Asn Val Val Lys Glu Ala															

ES 2 696 235 T3

115					120					125					
Lys	Gln	Glu	Gly	Ile	Lys	Val	Leu	Ala	Tyr	Asp	Arg	Met	Ile	Asn	Asp
130					135					140					
Ala	Asp	Ile	Asp	Phe	Tyr	Ile	Ser	Phe	Asp	Asn	Glu	Lys	Val	Gly	Glu
145					150					155					160
Leu	Gln	Ala	Lys	Ala	Leu	Val	Asp	Ile	Val	Pro	Gln	Gly	Asn	Tyr	Phe
				165					170					175	
Leu	Met	Gly	Gly	Ser	Pro	Val	Asp	Asn	Asn	Ala	Lys	Leu	Phe	Arg	Ala
			180					185					190		
Gly	Gln	Met	Lys	Val	Leu	Lys	Pro	Tyr	Val	Asp	Ser	Gly	Lys	Ile	Lys
		195					200					205			
Val	Val	Gly	Asp	Gln	Trp	Val	Asp	Gly	Trp	Leu	Pro	Glu	Asn	Ala	Leu
		210					215					220			
Lys	Ile	Met	Glu	Asn	Ala	Leu	Thr	Ala	Asn	Asn	Asn	Lys	Ile	Asp	Ala
225					230					235					240
Val	Val	Ala	Ser	Asn	Asp	Ala	Thr	Ala	Gly	Gly	Ala	Ile	Gln	Ala	Leu
				245					250					255	
Ser	Ala	Gln	Gly	Leu	Ser	Gly	Lys	Val	Ala	Ile	Ser	Gly	Gln	Asp	Ala
			260					265					270		
Asp	Leu	Ala	Gly	Ile	Lys	Arg	Ile	Ala	Ala	Gly	Thr	Gln	Thr	Met	Thr
		275					280					285			
Val	Tyr	Lys	Pro	Ile	Thr	Leu	Leu	Ala	Asn	Thr	Ala	Ala	Glu	Ile	Ala
		290				295					300				
Val	Glu	Leu	Gly	Asn	Gly	Gln	Glu	Pro	Lys	Ala	Asp	Thr	Ser	Leu	Asn
305					310					315					320
Asn	Gly	Leu	Lys	Asp	Val	Pro	Ser	Arg	Leu	Leu	Thr	Pro	Ile	Asp	Val
				325					330					335	
Asn	Lys	Asn	Asn	Ile	Lys	Asp	Thr	Val	Ile	Lys	Asp	Gly	Phe	His	Lys
			340					345					350		
Glu	Ser	Glu	Leu												
		355													

<210> 49
 <211> 330
 <212> PRT
 <213> *Citrobacter koseri*

5

<400> 49

```

Met Lys Ile Lys Asn Ile Leu Leu Thr Val Cys Thr Ser Leu Leu Leu
 1              5              10              15

Thr Ser Ile Thr Gly His Ala Lys Glu Val Lys Ile Gly Met Ala Ile
      20              25              30

Asp Asp Leu Arg Leu Glu Arg Trp Gln Lys Asp Arg Asp Ile Phe Val
      35              40              45

Ser Lys Ala Glu Ser Leu Gly Ala Lys Val Phe Val Gln Ser Ala Asn
 50              55              60

Gly Asn Glu Glu Thr Gln Met Ser Gln Ile Glu Asn Met Ile Asn Arg
65              70              75              80

Gly Val Asp Val Leu Val Ile Ile Pro Tyr Asn Gly Gln Val Leu Ser
      85              90              95

Asn Val Val Lys Glu Ala Lys Gln Glu Gly Ile Lys Val Leu Ala Tyr
      100              105              110

Asp Arg Met Ile Asn Asp Ala Asp Ile Asp Phe Tyr Ile Ser Phe Asp
      115              120              125

Asn Glu Lys Val Gly Glu Leu Gln Ala Gln Ala Leu Val Asp Lys Val
      130              135              140

Pro Glu Gly Asn Tyr Phe Leu Met Gly Gly Ser Pro Val Asp Asn Asn
      145              150              155              160

Ala Lys Leu Phe Arg Ala Gly Gln Met Lys Val Leu Lys Pro Tyr Ile
      165              170              175

Asp Ser Gly Lys Ile Asn Val Val Gly Asp Gln Trp Val Asp Gly Trp
      180              185              190

Leu Pro Glu Asn Ala Leu Lys Ile Met Glu Asn Ala Leu Thr Ala Asn
      195              200              205

Asn Asn Lys Ile Asp Ala Val Val Ala Ser Asn Asp Ala Thr Ala Gly
      210              215              220

```

Gly Ala Ile Gln Ala Leu Ser Ala Gln Gly Leu Ser Gly Lys Val Ala
225 230 235 240

Ile Ser Gly Gln Asp Ala Asp Leu Ala Gly Ile Lys Arg Ile Val Ala
245 250 255

Gly Thr Gln Thr Met Thr Val Tyr Lys Pro Ile Thr Leu Leu Ala Asn
260 265 270

Thr Ala Ala Glu Ile Ala Val Glu Leu Gly Asn Gly Gln Gln Pro Lys
275 280 285

Ala Asp Ala Thr Leu Asn Asn Gly Leu Lys Asp Val Pro Ser Arg Leu
290 295 300

Leu Thr Pro Ile Asp Val Asn Lys Thr Asn Ile Asp Glu Thr Val Ile
305 310 315 320

Lys Asp Gly Phe His Lys Lys Ser Asp Leu
325 330

<210> 50

<211> 513

5 <212> PRT

<213> *Shigella boydii*

<400> 50

Met Pro Tyr Leu Leu Glu Met Lys Asn Ile Thr Lys Thr Phe Gly Ser
1 5 10 15

Val Lys Ala Ile Asp Asn Val Ser Leu Arg Leu Asn Ala Gly Glu Ile
20 25 30

Val Ser Leu Cys Gly Glu Asn Gly Ser Gly Lys Ser Thr Leu Met Lys
35 40 45

Val Leu Cys Gly Ile Tyr Pro His Gly Ser Tyr Glu Gly Glu Ile Ile
50 55 60

Phe Ala Gly Glu Glu Ile Gln Ala Ser His Ile Arg Asp Thr Glu Arg
65 70 75 80

Lys Gly Ile Ala Ile Ile His Gln Glu Leu Ala Leu Val Lys Glu Leu
85 90 95

Thr Val Leu Glu Asn Ile Phe Leu Gly Asn Glu Ile Thr His Asn Gly
100 105 110

Ile Met Asp Tyr Asp Leu Met Thr Leu Arg Cys Gln Lys Leu Leu Ala
 115 120 125

Gln Val Ser Leu Ser Ile Ser Pro Asp Thr Arg Val Gly Asp Leu Gly
 130 135 140

Leu Gly Gln Gln Gln Leu Val Glu Ile Ala Lys Ala Leu Asn Lys Gln
 145 150 155 160

Val Arg Leu Leu Ile Leu Asp Glu Pro Thr Ala Ser Leu Thr Glu Gln
 165 170 175

Glu Thr Ser Val Leu Leu Asp Ile Ile Arg Asp Leu Gln Gln His Gly
 180 185 190

Ile Ala Cys Ile Tyr Ile Ser His Lys Leu Asn Glu Val Lys Ala Ile
 195 200 205

Ser Asp Thr Ile Cys Val Ile Arg Asp Gly Gln His Ile Gly Thr Arg
 210 215 220

Asp Ala Ala Gly Met Ser Glu Asp Asp Ile Ile Thr Met Met Val Gly
 225 230 235 240

Arg Glu Leu Thr Ala Leu Tyr Pro Asn Glu Pro His Thr Thr Gly Asp
 245 250 255

Glu Ile Leu Arg Ile Glu His Leu Thr Ala Trp His Pro Val Asn Arg
 260 265 270

His Ile Lys Arg Val Asn Asp Val Ser Phe Ser Leu Lys Arg Gly Glu
 275 280 285

Ile Leu Gly Ile Ala Gly Leu Val Gly Ala Gly Arg Thr Glu Thr Ile
 290 295 300

Gln Cys Leu Phe Gly Val Trp Pro Gly Gln Trp Glu Gly Lys Ile Tyr
 305 310 315 320

Ile Asp Gly Lys Gln Val Asp Ile Arg Asn Cys Gln Gln Ala Ile Ala
 325 330 335

Gln Gly Ile Ala Met Val Pro Glu Asp Arg Lys Arg Asp Gly Ile Val
 340 345 350

Pro Val Met Ala Val Gly Lys Asn Ile Thr Leu Ala Ala Leu Asn Lys
 355 360 365

Phe Thr Gly Gly Ile Ser Gln Leu Asp Asp Ala Ala Glu Gln Lys Cys
370 375 380

Ile Leu Glu Ser Ile Gln Gln Leu Lys Val Lys Thr Ser Ser Pro Asp
385 390 395 400

Leu Ala Ile Gly Arg Leu Ser Gly Gly Asn Gln Gln Lys Ala Ile Leu
405 410 415

Ala Arg Cys Leu Leu Leu Asn Pro Arg Ile Leu Ile Leu Asp Glu Pro
420 425 430

Thr Arg Gly Ile Asp Ile Gly Val Lys Tyr Glu Ile Tyr Lys Leu Ile
435 440 445

Asn Gln Leu Val Gln Gln Gly Ile Ala Val Ile Val Ile Ser Ser Glu
450 455 460

Leu Pro Glu Val Leu Gly Leu Ser Asp Arg Val Leu Val Met His Glu
465 470 475 480

Gly Lys Leu Lys Ala Asn Leu Ile Asn His Asn Leu Thr Gln Glu Gln
485 490 495

Val Met Glu Ala Ala Leu Arg Ser Glu His His Val Glu Lys Gln Ser
500 505 510

Val

<210> 51

<211> 513

<212> PRT

<213> *Klebsiella pneumoniae*

<400> 51

Met Ala Trp Leu Leu Glu Met Lys Asn Ile Thr Lys Thr Phe Gly Ala
1 5 10 15

Val Lys Ala Val Asp Asn Val Ser Leu Arg Leu Asn Ala Gly Glu Val
20 25 30

Val Ser Leu Cys Gly Glu Asn Gly Ser Gly Lys Ser Thr Leu Met Lys
35 40 45

Val Leu Cys Gly Ile Tyr Pro His Gly Ser Tyr Glu Gly Glu Ile Ile
50 55 60

ES 2 696 235 T3

Phe	Ala	Gly	Glu	Thr	Leu	Gln	Ala	Asn	His	Ile	Arg	Asp	Thr	Glu	Arg	65	70	75	80
Lys	Gly	Ile	Ala	Ile	Ile	His	Gln	Glu	Leu	Ala	Leu	Val	Lys	His	Leu	85	90	95	
Thr	Val	Leu	Glu	Asn	Ile	Phe	Leu	Gly	Ala	Glu	Ile	Ser	Arg	His	Gly	100	105	110	
Leu	Leu	Asp	Tyr	Glu	Thr	Met	Thr	Leu	Arg	Cys	Gln	Lys	Leu	Leu	Ala	115	120	125	
Gln	Val	Asn	Leu	Pro	Ile	Ser	Pro	Asp	Thr	Arg	Val	Gly	Asp	Leu	Gly	130	135	140	
Leu	Gly	Gln	Gln	Gln	Leu	Val	Glu	Ile	Ala	Lys	Ala	Leu	Asn	Lys	Gln	145	150	155	160
Val	Arg	Leu	Leu	Ile	Leu	Asp	Glu	Pro	Thr	Ala	Ser	Leu	Thr	Glu	Gln	165	170	175	
Glu	Thr	Ala	Thr	Leu	Leu	Thr	Ile	Ile	Arg	Asp	Leu	Gln	Asn	His	Gly	180	185	190	
Ile	Ala	Cys	Ile	Tyr	Ile	Ser	His	Lys	Leu	Asn	Glu	Val	Lys	Ala	Ile	195	200	205	
Ser	Asp	Thr	Ile	Cys	Val	Ile	Arg	Asp	Gly	Gln	His	Ile	Gly	Thr	Arg	210	215	220	
Asp	Ala	Ser	Gly	Met	Ser	Glu	Asp	Asp	Ile	Ile	Thr	Met	Met	Val	Gly	225	230	235	240
Arg	Glu	Leu	Thr	Ala	Leu	Tyr	Pro	Ser	Glu	Pro	His	Ala	His	Gly	Glu	245	250	255	
Glu	Ile	Leu	Arg	Val	Glu	His	Leu	Thr	Ala	Trp	His	Pro	Val	Asn	Arg	260	265	270	
His	Ile	Lys	Arg	Val	Asn	Asp	Val	Ser	Phe	Ser	Leu	Arg	Arg	Gly	Glu	275	280	285	
Ile	Leu	Gly	Ile	Ala	Gly	Leu	Val	Gly	Ala	Gly	Arg	Thr	Glu	Ala	Val	290	295	300	
Gln	Cys	Leu	Phe	Gly	Val	Trp	Pro	Gly	Arg	Trp	Gln	Gly	Glu	Ile	Phe	305	310	315	320

ES 2 696 235 T3

Ile Asp Gly Gln Pro Val Ser Ile Ser Asn Cys Gln Gln Ala Ile Ala
325 330 335

His Gly Ile Ala Met Val Pro Glu Asp Arg Lys Lys Asp Gly Ile Val
340 345 350

Pro Val Met Ala Val Gly Lys Asn Ile Thr Leu Ala Ala Leu Asn Gln
355 360 365

Phe Thr Gly Ala Met Ser Ser Leu Asp Asp Ala Ala Glu Gln His Cys
370 375 380

Ile Gln Gln Ser Ile Gln Arg Leu Lys Ile Lys Thr Ser Ser Pro Glu
385 390 395 400

Leu Ala Ile Gly Arg Leu Ser Gly Gly Asn Gln Gln Lys Ala Ile Leu
405 410 415

Ala Arg Cys Leu Leu Leu Asn Pro Arg Ile Leu Ile Leu Asp Glu Pro
420 425 430

Thr Arg Gly Ile Asp Ile Gly Ala Lys Tyr Glu Ile Tyr Lys Leu Ile
435 440 445

Asn Gln Leu Val Gln Gln Gly Ile Ala Val Ile Val Ile Ser Ser Glu
450 455 460

Leu Pro Glu Val Leu Gly Leu Ser Asp Arg Val Leu Val Met His Glu
465 470 475 480

Gly Arg Leu Lys Ala Asn Leu Val Asn Gln His Leu Thr Gln Glu Gln
485 490 495

Val Met Glu Ala Ala Leu Arg Ser Glu Arg His Val Glu Glu His Val
500 505 510

Val

<210> 52

<211> 393

<212> PRT

<213> *Citrobacter koseri*

<400> 52

Met Ser Lys Asn Asn Pro Ser Glu Val Lys Leu Ser Ala Pro Leu Thr
1 5 10 15

ES 2 696 235 T3

Gly Gly Phe Pro Gly Leu Lys Ser Leu Asn Leu Gln Val Phe Val Met
 20 25 30
 Ile Ala Ala Ile Ile Val Ile Met Leu Phe Phe Thr Trp Thr Thr Glu
 35 40 45
 Gly Ala Tyr Leu Ser Ala Arg Asn Val Ser Asn Leu Leu Arg Gln Thr
 50 55 60
 Ala Ile Thr Gly Ile Leu Ala Val Gly Met Val Phe Val Ile Ile Ser
 65 70 75 80
 Ala Glu Ile Asp Leu Ser Val Gly Ser Met Met Gly Leu Leu Gly Gly
 85 90 95
 Val Ala Ala Ile Cys Asp Val Trp Leu Gly Trp Pro Leu Pro Leu Thr
 100 105 110
 Ile Val Val Thr Leu Ala Leu Gly Leu Leu Leu Gly Ala Trp Asn Gly
 115 120 125
 Trp Trp Val Ala Tyr Arg Lys Val Pro Ser Phe Ile Val Thr Leu Ala
 130 135 140
 Gly Met Leu Ala Phe Arg Gly Ile Leu Ile Gly Ile Thr Asn Gly Thr
 145 150 155 160
 Thr Val Ser Pro Thr Ser Ser Ala Met Ser Gln Ile Gly Gln Ser Tyr
 165 170 175
 Leu Ser Asn Gly Leu Gly Phe Thr Ile Gly Ala Ile Ser Leu Met Ala
 180 185 190
 Phe Ile Gly Trp Gln Trp Arg Gly Arg Met Arg Arg Gln Ala Leu Ala
 195 200 205
 Leu Ser Thr Pro Ala Ser Thr Ala Val Val Gly Arg Gln Ala Leu Thr
 210 215 220
 Ala Val Ile Val Leu Gly Ala Ile Trp Leu Leu Asn Asp Tyr Arg Gly
 225 230 235 240
 Val Pro Thr Pro Val Leu Leu Leu Val Phe Leu Leu Leu Gly Gly Met
 245 250 255
 Phe Met Ala Thr Arg Thr Ala Phe Gly Arg Arg Ile Tyr Ala Ile Gly

[illegible]

<210> 53
<211> 393
<212> PRT
<213> *Shigella dysenteriae*

<400> 53

Met	Ser	Lys	Ser	Asn	Pro	Ser	Glu	Val	Lys	Leu	Ala	Val	Pro	Thr	Ser
1				5					10					15	
Gly	Gly	Phe	Ser	Gly	Leu	Lys	Ser	Leu	Asn	Leu	Gln	Val	Phe	Val	Met
			20					25					30		
Ile	Ala	Ala	Ile	Ile	Ala	Ile	Met	Leu	Phe	Phe	Thr	Trp	Thr	Thr	Asp
		35					40					45			
Gly	Ala	Tyr	Leu	Ser	Ala	Arg	Asn	Val	Ser	Asn	Leu	Leu	His	Gln	Thr
	50					55					60				
Ala	Ile	Thr	Gly	Ile	Leu	Ala	Val	Gly	Met	Val	Phe	Val	Ile	Ile	Ser
65					70					75					80
Ala	Glu	Ile	Asp	Leu	Ser	Val	Gly	Ser	Met	Met	Gly	Leu	Leu	Gly	Gly

ES 2 696 235 T3

				85					90					95			
Val	Ala	Ala	Ile	Cys	Asp	Val	Trp	Leu	Gly	Trp	Pro	Leu	Pro	Leu	Thr		
			100					105					110				
Ile	Ile	Val	Thr	Leu	Val	Leu	Gly	Leu	Leu	Leu	Gly	Ala	Trp	Asn	Gly		
		115					120					125					
Trp	Trp	Val	Ala	Tyr	Arg	Lys	Val	Pro	Ser	Phe	Ile	Val	Thr	Leu	Ala		
	130					135					140						
Gly	Met	Leu	Ala	Phe	Arg	Gly	Ile	Leu	Ile	Gly	Ile	Thr	Asn	Gly	Thr		
145					150					155					160		
Thr	Val	Ser	Pro	Thr	Ser	Ala	Ala	Met	Ser	Gln	Ile	Gly	Gln	Ser	Tyr		
				165					170					175			
Leu	Pro	Ala	Ser	Thr	Gly	Phe	Ile	Ile	Gly	Ala	Leu	Gly	Leu	Met	Ala		
			180					185					190				
Phe	Val	Gly	Trp	Gln	Trp	Arg	Gly	Arg	Met	Arg	Arg	Gln	Ala	Leu	Gly		
		195					200					205					
Leu	Gln	Ser	Pro	Ala	Ser	Thr	Ala	Val	Val	Gly	Arg	Gln	Ala	Leu	Thr		
	210					215					220						
Ala	Ile	Ile	Val	Leu	Gly	Ala	Ile	Trp	Leu	Leu	Asn	Asp	Tyr	Arg	Gly		
225					230					235					240		
Val	Pro	Thr	Pro	Val	Leu	Leu	Leu	Thr	Leu	Leu	Leu	Leu	Gly	Gly	Met		
				245					250					255			
Phe	Met	Ala	Thr	Arg	Thr	Ala	Phe	Gly	Arg	Arg	Ile	Tyr	Ala	Ile	Gly		
			260					265					270				
Gly	Asn	Leu	Glu	Ala	Ala	Arg	Leu	Ser	Gly	Ile	Asn	Val	Glu	Arg	Thr		
	275						280					285					
Lys	Leu	Ala	Val	Phe	Ala	Ile	Asn	Gly	Leu	Met	Val	Ala	Ile	Ala	Gly		
	290					295					300						
Leu	Ile	Leu	Ser	Ser	Arg	Leu	Gly	Ala	Gly	Ser	Pro	Ser	Ala	Gly	Asn		
305					310					315					320		
Ile	Ala	Glu	Leu	Asp	Ala	Ile	Ala	Ala	Cys	Val	Ile	Gly	Gly	Thr	Ser		
				325					330					335			

Leu Ala Gly Gly Val Gly Ser Val Ala Gly Ala Val Met Gly Ala Phe
 340 345 350

Ile Met Ala Ser Leu Asp Asn Gly Met Ser Met Met Asp Val Pro Thr
 355 360 365

Phe Trp Gln Tyr Ile Val Lys Gly Ala Ile Leu Leu Leu Ala Val Trp
 370 375 380

Met Asp Ser Ala Thr Lys Arg Arg Ser
 385 390

<210> 54

<211> 393

5 <212> PRT

<213> *Enterobacter cloacae*

<400> 54

Met Ser Lys Ser Asn Pro Ser Asp Ile Lys Val Ala Val Gln Thr Pro
 1 5 10 15

Gly Ala Phe Ala Gly Leu Lys Ala Leu Asn Leu Gln Val Phe Val Met
 20 25 30

Ile Ala Ala Ile Ile Ala Ile Met Leu Phe Phe Thr Trp Met Thr Asp
 35 40 45

Gly Ser Tyr Leu Ser Ala Arg Asn Val Ser Asn Leu Leu Arg Gln Thr
 50 55 60

Ala Ile Thr Gly Ile Leu Ala Val Gly Met Val Phe Val Ile Ile Ser
 65 70 75 80

Ala Glu Ile Asp Leu Ser Val Gly Ser Met Met Gly Leu Leu Gly Gly
 85 90 95

Val Ala Ala Ile Phe Asp Val Trp Leu Gly Trp Pro Leu Pro Leu Thr
 100 105 110

Val Ala Val Thr Leu Val Leu Gly Leu Val Leu Gly Ala Trp Asn Gly
 115 120 125

Trp Trp Val Ala Tyr Arg Lys Val Pro Ser Phe Ile Val Thr Leu Ala
 130 135 140

Gly Met Leu Ala Phe Arg Gly Val Leu Ile Gly Ile Thr Asn Gly Thr
 145 150 155 160

Thr Val Ser Pro Thr Ser Ala Ala Met Ser Gln Ile Gly Gln Ser Tyr
 165 170 175

Leu Ser Asp Ser Met Gly Phe Thr Ile Gly Val Val Gly Leu Leu Ala
 180 185 190

Phe Val Val Trp Gln Trp Arg Gly Arg Met Arg Arg Gln Ser Leu Gly
 195 200 205

Leu Ala Ser Ser Pro Ser Thr Ser Val Val Gly Arg Gln Ala Leu Thr
 210 215 220

Ala Val Ile Val Leu Gly Ala Ile Trp Leu Leu Asn Asp Tyr Arg Gly
 225 230 235 240

Val Pro Thr Pro Val Leu Leu Leu Ala Leu Leu Leu Leu Gly Gly Met
 245 250 255

Phe Met Ala Thr Arg Thr Ala Phe Gly Arg Arg Ile Tyr Ala Ile Gly
 260 265 270

Gly Asn Leu Glu Ala Ala Arg Leu Ser Gly Ile Asn Val Glu Arg Thr
 275 280 285

Lys Leu Ala Val Phe Ala Ile Asn Gly Leu Met Val Ala Ile Ala Gly
 290 295 300

Leu Ile Leu Ser Ser Arg Leu Gly Ala Gly Ser Pro Ser Ala Gly Asn
 305 310 315 320

Ile Ala Glu Leu Asp Ala Ile Ala Ala Cys Val Ile Gly Gly Thr Ser
 325 330 335

Leu Ala Gly Gly Val Gly Ser Val Ala Gly Ala Val Met Gly Ala Phe
 340 345 350

Ile Met Ala Ser Leu Asp Asn Gly Met Ser Met Met Asp Val Pro Thr
 355 360 365

Phe Trp Gln Tyr Ile Val Lys Gly Ala Ile Leu Leu Leu Ala Val Trp
 370 375 380

Met Asp Ser Ala Thr Lys Arg Arg Thr
 385 390

<210> 55
 <211> 321

<212> PRT

<213> *Shigella dysenteriae*

<400> 55

```

Met Thr Lys Tyr Ala Leu Val Gly Asp Val Gly Gly Thr Asn Ala Arg
 1              5              10              15

Leu Ala Leu Cys Asp Ile Ala Ser Gly Glu Ile Ser Gln Ala Lys Thr
      20              25              30

Tyr Ser Gly Leu Asp Tyr Pro Ser Leu Glu Ala Val Ile Arg Val Tyr
      35              40              45

Leu Glu Glu His Lys Val Glu Val Lys Asp Gly Cys Ile Ala Ile Ala
 50              55              60

Cys Pro Ile Thr Gly Asp Trp Val Ala Met Thr Asn His Thr Trp Ala
65              70              75              80

Phe Ser Ile Ala Glu Met Lys Lys Asn Leu Gly Phe Ser His Leu Glu
      85              90              95

Ile Ile Asn Asp Phe Thr Ala Val Ser Met Ala Val Pro Met Leu Lys
      100              105              110

Lys Glu His Leu Ile Gln Phe Gly Gly Ala Glu Pro Ile Glu Gly Lys
      115              120              125

Pro Ile Ala Val Tyr Gly Ala Gly Thr Gly Leu Gly Val Ala His Leu
      130              135              140

Val His Val Asp Lys Arg Trp Val Ser Leu Pro Gly Glu Gly Gly His
      145              150              155              160

Val Asp Phe Ala Pro Asn Ser Glu Glu Glu Ala Ile Ile Leu Glu Ile
      165              170              175

Leu Arg Ala Glu Ile Gly His Val Ser Ala Glu Arg Val Leu Ser Gly
      180              185              190

Pro Gly Leu Val Asn Leu Tyr Arg Ala Ile Val Lys Ala Asp Asn Arg
      195              200              205

Leu Pro Glu Asn Leu Lys Pro Lys Asp Ile Thr Glu Arg Ala Leu Ala
      210              215              220

Asp Ser Cys Thr Asp Cys Arg Arg Ala Leu Ser Leu Phe Cys Val Ile
      225              230              235              240

```

Met Gly Arg Phe Gly Gly Asn Leu Ala Leu Asn Leu Gly Thr Phe Gly
245 250 255

Gly Val Phe Ile Ala Gly Gly Ile Val Pro Arg Phe Leu Glu Phe Phe
260 265 270

Lys Ala Ser Gly Phe Arg Ala Ala Phe Glu Asp Lys Gly Arg Phe Lys
275 280 285

Glu Tyr Val His Asp Ile Pro Val Tyr Leu Ile Val His Asp Asn Pro
290 295 300

Gly Leu Leu Gly Ser Gly Ala His Leu Arg Gln Thr Leu Gly His Ile
305 310 315 320

Leu

<210> 56

<211> 321

<212> PRT

<213> *Escherichia fergusonii*

<400> 56

Met Thr Lys Tyr Ala Leu Val Gly Asp Val Gly Gly Thr Asn Ala Arg
1 5 10 15

Leu Ala Leu Cys Asp Ile Ala Ser Gly Glu Ile Ser Gln Ala Lys Thr
20 25 30

Tyr Ser Gly Leu Asp Tyr Pro Ser Leu Glu Ala Val Ile Arg Val Tyr
35 40 45

Leu Glu Glu His Asn Val Glu Val Gln Asp Gly Cys Ile Ala Ile Ala
50 55 60

Cys Pro Ile Thr Gly Asp Trp Val Ala Met Thr Asn His Thr Trp Ala
65 70 75 80

Phe Ser Ile Ala Glu Met Lys Lys Asn Leu Gly Phe Ser His Leu Glu
85 90 95

Ile Ile Asn Asp Phe Thr Ala Val Ser Met Ala Ile Pro Met Leu Lys
100 105 110

Lys Glu His Leu Ile Gln Phe Gly Gly Ala Glu Pro Val Glu Gly Lys
115 120 125

5

10

Pro Ile Ala Val Tyr Gly Ala Gly Thr Gly Leu Gly Val Ala His Leu
130 135 140

Val His Val Asp Lys Arg Trp Val Ser Leu Pro Gly Glu Gly Gly His
145 150 155 160

Val Asp Phe Ala Pro Asn Ser Glu Glu Glu Gly Ile Ile Leu Glu Ile
165 170 175

Leu Arg Ala Glu Ile Gly His Val Ser Ala Glu Arg Val Leu Ser Gly
180 185 190

Pro Gly Leu Val Asn Leu Tyr Arg Ala Ile Val Lys Ala Asp Asn Arg
195 200 205

Leu Pro Glu Asn Leu Lys Pro Lys Asp Ile Thr Glu Arg Ala Leu Ala
210 215 220

Asp Ser Cys Thr Asp Cys Arg Arg Ala Leu Ser Leu Phe Cys Val Ile
225 230 235 240

Met Gly Arg Phe Gly Gly Asn Leu Ala Leu Thr Leu Gly Thr Phe Gly
245 250 255

Gly Val Tyr Ile Ala Gly Gly Ile Val Pro Arg Phe Leu Glu Phe Phe
260 265 270

Lys Ala Ser Gly Phe Arg Ala Ala Phe Glu Asp Lys Gly Arg Phe Lys
275 280 285

Glu Tyr Val His Asp Ile Pro Val Tyr Leu Ile Val His Asp Asn Pro
290 295 300

Gly Leu Leu Gly Ser Gly Ala His Leu Arg Gln Thr Leu Gly His Ile
305 310 315 320

Leu

<210> 57
<211> 371
5 <212> PRT
<213> *Shigella dysenteriae*
<400> 57

Met Tyr Lys Asn Gly Gly Arg Tyr Gln Arg Ser Val Arg Asn Pro Lys
1 5 10 15

Gln	Ser	Arg	Leu	Ala	Ile	Thr	Val	Leu	Ser	Ala	Val	Ala	Leu	Thr	Leu	20	25	30	
Arg	Arg	Asp	Ile	Ile	Arg	Ala	Asp	Arg	Leu	His	Pro	Leu	Thr	Ser	Cys	35	40	45	
Leu	Lys	Arg	Asn	Thr	Ile	Met	Thr	Asp	Lys	Leu	Thr	Ser	Leu	Arg	Gln	50	55	60	
Tyr	Thr	Thr	Val	Val	Ala	Asp	Thr	Gly	Asp	Ile	Ala	Ala	Met	Lys	Leu	65	70	75	80
Tyr	Gln	Pro	Gln	Asp	Ala	Thr	Thr	Asn	Pro	Ser	Leu	Ile	Leu	Asn	Ala	85	90	95	
Ala	Gln	Ile	Pro	Glu	Tyr	Arg	Lys	Leu	Ile	Asp	Asp	Ala	Val	Ala	Trp	100	105	110	
Ala	Lys	Gln	Gln	Ser	Asn	Asp	Arg	Ala	Gln	Gln	Ile	Val	Asp	Ala	Thr	115	120	125	
Asp	Lys	Leu	Ala	Val	Asn	Ile	Gly	Leu	Glu	Ile	Leu	Lys	Leu	Val	Pro	130	135	140	
Gly	Arg	Ile	Ser	Thr	Glu	Val	Asp	Ala	Arg	Leu	Ser	Tyr	Asp	Thr	Glu	145	150	155	160
Ala	Ser	Ile	Ala	Lys	Ala	Lys	Arg	Leu	Ile	Lys	Leu	Tyr	Asn	Asp	Ala	165	170	175	
Gly	Ile	Ser	Asn	Asp	Arg	Ile	Leu	Ile	Lys	Leu	Ala	Ser	Thr	Trp	Gln	180	185	190	
Gly	Ile	Arg	Ala	Ala	Glu	Gln	Leu	Glu	Lys	Glu	Gly	Ile	Asn	Cys	Asn	195	200	205	
Leu	Thr	Leu	Leu	Phe	Ser	Phe	Ala	Gln	Ala	Arg	Ala	Cys	Ala	Glu	Ala	210	215	220	
Gly	Val	Phe	Leu	Ile	Ser	Pro	Phe	Val	Gly	Arg	Ile	Leu	Asp	Trp	Tyr	225	230	235	240
Lys	Ala	Asn	Thr	Asp	Lys	Lys	Glu	Tyr	Ala	Pro	Ala	Glu	Asp	Pro	Gly	245	250	255	
Val	Val	Ser	Val	Ser	Glu	Ile	Tyr	Gln	Tyr	Tyr	Lys	Glu	His	Gly	Tyr	260	265	270	

Glu Thr Val Val Met Gly Ala Ser Phe Arg Asn Ile Gly Glu Ile Leu
275 280 285

Glu Leu Ala Gly Cys Asp Arg Leu Thr Ile Ala Pro Ala Leu Leu Lys
290 295 300

Glu Leu Ala Glu Ser Glu Gly Ala Ile Glu Arg Lys Leu Ser Tyr Thr
305 310 315 320

Gly Glu Val Lys Ala Arg Pro Ala Arg Ile Thr Glu Ser Glu Phe Leu
325 330 335

Trp Gln His Asn Gln Asp Pro Met Ala Val Asp Lys Leu Ala Glu Gly
340 345 350

Ile Arg Lys Phe Ala Ile Asp Gln Glu Lys Leu Glu Lys Met Ile Gly
355 360 365

Asp Leu Leu
370

<210> 58

<211> 317

<212> PRT

<213> *Escherichia fergusonii*

<400> 58

Met Thr Asp Lys Leu Thr Ser Leu Arg Gln Tyr Thr Thr Val Val Ala
1 5 10 15

Asp Thr Gly Asp Ile Ala Ala Met Lys Leu Tyr Gln Pro Gln Asp Ala
20 25 30

Thr Thr Asn Pro Ser Leu Ile Leu Asn Ala Ala Gln Ile Pro Glu Tyr
35 40 45

Arg Lys Leu Ile Asp Asp Ala Val Ala Trp Ala Lys Gln Gln Ser Ser
50 55 60

Asp Arg Ala Gln Gln Ile Val Asp Ala Thr Asp Lys Leu Ala Val Asn
65 70 75 80

Ile Gly Leu Glu Ile Leu Lys Leu Val Pro Gly Arg Ile Ser Thr Glu
85 90 95

Val Asp Ala Arg Leu Ser Tyr Asp Thr Glu Ala Ser Ile Ala Lys Ala
100 105 110

5

10

Lys Arg Leu Ile Lys Leu Tyr Asn Asp Ala Gly Ile Ser Asn Asp Arg
 115 120 125
 Ile Leu Ile Lys Leu Ala Ser Thr Trp Gln Gly Ile Arg Ala Ala Glu
 130 135 140
 Gln Leu Glu Lys Glu Gly Ile Asn Cys Asn Leu Thr Leu Leu Phe Ser
 145 150 155 160
 Phe Ala Gln Ala Arg Ala Cys Ala Glu Ala Gly Val Tyr Leu Ile Ser
 165 170 175
 Pro Phe Val Gly Arg Ile Leu Asp Trp Tyr Lys Ala Asn Thr Asp Lys
 180 185 190
 Lys Glu Tyr Ala Pro Ala Glu Asp Pro Gly Val Val Ser Val Ser Glu
 195 200 205
 Ile Tyr Gln Tyr Tyr Lys Glu His Gly Tyr Glu Thr Val Val Met Gly
 210 215 220
 Ala Ser Phe Arg Asn Met Gly Glu Ile Leu Glu Leu Ala Gly Cys Asp
 225 230 235 240
 Arg Leu Thr Ile Ala Pro Ala Leu Leu Lys Glu Leu Ala Glu Ser Glu
 245 250 255
 Gly Ala Val Glu Arg Lys Leu Ser Phe Ser Gly Glu Val Lys Ala Arg
 260 265 270
 Pro Glu Arg Ile Thr Glu Ser Glu Phe Leu Trp Gln His Asn Gln Asp
 275 280 285
 Pro Met Ala Val Asp Lys Leu Ala Asp Gly Ile Arg Lys Phe Ala Val
 290 295 300
 Asp Gln Glu Lys Leu Glu Lys Met Ile Gly Glu Leu Leu
 305 310 315

<210> 59

<211> 673

<212> PRT

<213> *Shigella dysenteriae*

<400> 59

Met Pro Phe Ile Ile Arg Ser Gly Val Lys Met Ser Ser Arg Lys Glu
 1 5 10 15

ES 2 696 235 T3

Leu Ala Asn Ala Ile Arg Ala Leu Ser Met Asp Ala Val Gln Lys Ala
 20 25 30
 Lys Ser Gly His Pro Gly Ala Pro Met Gly Met Ala Asp Ile Ala Glu
 35 40 45
 Val Leu Trp Arg Asp Phe Leu Lys His Asn Pro Gln Asn Pro Ser Trp
 50 55 60
 Ala Asp Arg Asp Arg Phe Val Leu Ser Asn Gly His Gly Ser Met Leu
 65 70 75 80
 Ile Tyr Ser Leu Leu His Leu Thr Gly Tyr Asp Leu Pro Met Glu Glu
 85 90 95
 Leu Lys Asn Phe Arg Gln Leu His Ser Lys Thr Pro Gly His Pro Glu
 100 105 110
 Val Gly Tyr Thr Ala Gly Val Glu Thr Thr Thr Gly Pro Leu Gly Gln
 115 120 125
 Gly Ile Ala Asn Ala Val Gly Met Ala Ile Ala Glu Lys Thr Leu Ala
 130 135 140
 Ala Gln Phe Asn Arg Pro Gly His Asp Ile Val Asp His Tyr Thr Tyr
 145 150 155 160
 Ala Phe Met Gly Asp Gly Cys Met Met Glu Gly Ile Ser His Glu Val
 165 170 175
 Cys Ser Leu Ala Gly Thr Leu Lys Leu Gly Lys Leu Ile Ala Phe Tyr
 180 185 190
 Asp Asp Asn Gly Ile Ser Ile Asp Gly His Val Glu Gly Trp Phe Thr
 195 200 205
 Asp Asp Thr Ala Met Arg Phe Glu Ala Tyr Gly Trp His Val Ile Arg
 210 215 220
 Asp Ile Asp Gly His Asp Ala Ala Ser Ile Lys Arg Ala Val Glu Glu
 225 230 235 240
 Ala Arg Ala Val Thr Asp Lys Pro Ser Leu Leu Met Cys Lys Thr Ile
 245 250 255
 Ile Gly Phe Gly Ser Pro Asn Lys Ala Gly Thr His Asp Ser His Gly

ES 2 696 235 T3

260					265					270					
Ala	Pro	Leu	Gly	Asp	Ala	Glu	Ile	Ala	Leu	Thr	Arg	Glu	Gln	Leu	Gly
		275					280					285			
Trp	Lys	Tyr	Ala	Pro	Phe	Glu	Ile	Pro	Ser	Glu	Ile	Tyr	Ala	Gln	Trp
	290					295					300				
Asp	Ala	Lys	Glu	Ala	Gly	Gln	Ala	Lys	Glu	Ser	Ala	Trp	Asn	Glu	Lys
305					310					315					320
Phe	Ala	Ala	Tyr	Ala	Lys	Ala	Tyr	Pro	Gln	Glu	Ala	Ala	Glu	Phe	Thr
				325					330					335	
Arg	Arg	Met	Lys	Gly	Glu	Met	Pro	Ser	Asp	Phe	Asp	Ala	Lys	Ala	Lys
			340					345					350		
Glu	Phe	Ile	Ala	Lys	Leu	Gln	Ala	Asn	Pro	Ala	Lys	Ile	Ala	Ser	Arg
		355					360					365			
Lys	Ala	Ser	Gln	Asn	Ala	Ile	Glu	Ala	Phe	Gly	Pro	Leu	Leu	Pro	Glu
	370					375					380				
Phe	Leu	Gly	Gly	Ser	Ala	Asp	Leu	Ala	Pro	Ser	Asn	Leu	Thr	Leu	Trp
385					390					395					400
Ser	Gly	Ser	Lys	Ala	Ile	Asn	Glu	Asp	Ala	Ala	Gly	Asn	Tyr	Ile	His
				405					410					415	
Tyr	Gly	Val	Arg	Glu	Phe	Gly	Met	Thr	Ala	Ile	Ala	Asn	Gly	Ile	Ser
			420					425					430		
Leu	His	Gly	Gly	Phe	Leu	Pro	Tyr	Thr	Ser	Thr	Phe	Leu	Met	Phe	Val
		435					440					445			
Glu	Tyr	Ala	Arg	Asn	Ala	Val	Arg	Met	Ala	Ala	Leu	Met	Lys	Gln	Arg
	450					455					460				
Gln	Val	Met	Val	Tyr	Thr	His	Asp	Ser	Ile	Gly	Leu	Gly	Glu	Asp	Gly
465					470					475					480
Pro	Thr	His	Gln	Pro	Val	Glu	Gln	Val	Ala	Ser	Leu	Arg	Val	Thr	Pro
				485					490					495	
Asn	Met	Ser	Thr	Trp	Arg	Pro	Cys	Asp	Gln	Val	Glu	Ser	Ala	Val	Ala
			500					505					510		

Trp Lys Tyr Gly Val Glu Arg Gln Asp Gly Pro Thr Ala Leu Ile Leu
515 520 525

Ser Arg Gln Asn Leu Ala Gln Gln Glu Arg Thr Glu Glu Gln Leu Ala
530 535 540

Asn Ile Ala Arg Gly Gly Tyr Val Leu Lys Asp Cys Ala Gly Gln Pro
545 550 555 560

Glu Leu Ile Phe Ile Ala Thr Gly Ser Glu Val Glu Leu Ala Val Ala
565 570 575

Ala Tyr Glu Lys Leu Thr Ala Glu Gly Val Lys Ala Arg Val Val Ser
580 585 590

Met Pro Ser Thr Asp Ala Phe Asp Lys Gln Asp Ala Ala Tyr Arg Glu
595 600 605

Ser Val Leu Pro Lys Ala Val Thr Ala Arg Val Ala Val Glu Ala Gly
610 615 620

Ile Ala Asp Tyr Trp Tyr Lys Tyr Val Gly Leu Asn Gly Ala Ile Val
625 630 635 640

Gly Met Thr Thr Phe Gly Glu Ser Ala Pro Ala Glu Leu Leu Phe Glu
645 650 655

Glu Phe Gly Phe Thr Val Asp Asn Val Val Ala Lys Ala Lys Glu Leu
660 665 670

Leu

<210> 60

<211> 663

5 <212> PRT

<213> *Escherichia fergusonii*

<400> 60

Met Ser Ser Arg Lys Glu Leu Ala Asn Ala Ile Arg Ala Leu Ser Met
1 5 10 15

Asp Ala Val Gln Lys Ala Lys Ser Gly His Pro Gly Ala Pro Met Gly
20 25 30

Met Ala Asp Ile Ala Glu Val Leu Trp Arg Asp Phe Leu Lys His Asn
35 40 45

ES 2 696 235 T3

Pro	Gln	Asn	Pro	Ser	Trp	Ala	Asp	Arg	Asp	Arg	Phe	Val	Leu	Ser	Asn	50	55	60	
Gly	His	Gly	Ser	Met	Leu	Ile	Tyr	Ser	Leu	Leu	His	Leu	Thr	Gly	Tyr	65	70	75	80
Asp	Leu	Pro	Met	Glu	Glu	Leu	Lys	Asn	Phe	Arg	Gln	Leu	His	Ser	Lys	85	90	95	
Thr	Pro	Gly	His	Pro	Glu	Val	Gly	Tyr	Thr	Ala	Gly	Val	Glu	Thr	Thr	100	105	110	
Thr	Gly	Pro	Leu	Gly	Gln	Gly	Ile	Ala	Asn	Ala	Val	Gly	Met	Ala	Ile	115	120	125	
Ala	Glu	Lys	Thr	Leu	Ala	Ala	Gln	Phe	Asn	Arg	Pro	Gly	His	Asp	Ile	130	135	140	
Val	Asp	His	Tyr	Thr	Tyr	Ala	Phe	Met	Gly	Asp	Gly	Cys	Met	Met	Glu	145	150	155	160
Gly	Ile	Ser	His	Glu	Val	Cys	Ser	Leu	Ala	Gly	Thr	Leu	Lys	Leu	Gly	165	170	175	
Lys	Leu	Ile	Ala	Phe	Tyr	Asp	Asp	Asn	Gly	Ile	Ser	Ile	Asp	Gly	His	180	185	190	
Val	Glu	Gly	Trp	Phe	Thr	Asp	Asp	Thr	Ala	Met	Arg	Phe	Glu	Ala	Tyr	195	200	205	
Gly	Trp	His	Val	Ile	Arg	Gly	Ile	Asp	Gly	His	Asp	Ala	Ala	Ser	Ile	210	215	220	
Lys	Arg	Ala	Val	Glu	Glu	Ala	Arg	Ala	Val	Thr	Asp	Lys	Pro	Ser	Leu	225	230	235	240
Leu	Met	Cys	Lys	Thr	Ile	Ile	Gly	Phe	Gly	Ser	Pro	Asn	Lys	Ala	Gly	245	250	255	
Thr	His	Asp	Ser	His	Gly	Ala	Pro	Leu	Gly	Asp	Ala	Glu	Ile	Ala	Leu	260	265	270	
Thr	Arg	Glu	Gln	Leu	Gly	Trp	Lys	Tyr	Ala	Pro	Phe	Glu	Ile	Pro	Ser	275	280	285	
Glu	Ile	Tyr	Ala	Gln	Trp	Asp	Ala	Lys	Glu	Ala	Gly	Gln	Ala	Lys	Glu	290	295	300	

Ser	Ala	Trp	Asn	Glu	Lys	Phe	Ala	Ala	Tyr	Ala	Lys	Ala	Tyr	Pro	Gln	305	310	315	320
Glu	Ala	Ala	Glu	Phe	Thr	Arg	Arg	Met	Lys	Gly	Glu	Met	Pro	Ser	Asp	325	330	335	
Phe	Asp	Ala	Lys	Ala	Lys	Glu	Phe	Ile	Ala	Lys	Leu	Gln	Ala	Asn	Pro	340	345	350	
Ala	Lys	Ile	Ala	Ser	Arg	Lys	Ala	Ser	Gln	Asn	Ala	Ile	Glu	Ala	Phe	355	360	365	
Gly	Pro	Leu	Leu	Pro	Glu	Phe	Leu	Gly	Gly	Ser	Ala	Asp	Leu	Ala	Pro	370	375	380	
Ser	Asn	Leu	Thr	Leu	Trp	Ser	Gly	Ser	Lys	Ala	Ile	Asn	Glu	Asp	Ala	385	390	395	400
Ala	Gly	Asn	Tyr	Ile	His	Tyr	Gly	Val	Arg	Glu	Phe	Gly	Met	Thr	Ala	405	410	415	
Ile	Ala	Asn	Gly	Ile	Ser	Leu	His	Gly	Gly	Phe	Leu	Pro	Tyr	Thr	Ser	420	425	430	
Thr	Phe	Leu	Met	Phe	Val	Glu	Tyr	Ala	Arg	Asn	Ala	Val	Arg	Met	Ala	435	440	445	
Ala	Leu	Met	Lys	Gln	Arg	Gln	Val	Met	Val	Tyr	Thr	His	Asp	Ser	Ile	450	455	460	
Gly	Leu	Gly	Glu	Asp	Gly	Pro	Thr	His	Gln	Pro	Val	Glu	Gln	Val	Ala	465	470	475	480
Ser	Leu	Arg	Val	Thr	Pro	Asn	Met	Ser	Thr	Trp	Arg	Pro	Cys	Asp	Gln	485	490	495	
Val	Glu	Ser	Ala	Val	Ala	Trp	Lys	Tyr	Gly	Val	Glu	Arg	Gln	Asp	Gly	500	505	510	
Pro	Thr	Ala	Leu	Ile	Leu	Ser	Arg	Gln	Asn	Leu	Ala	Gln	Gln	Glu	Arg	515	520	525	
Thr	Glu	Glu	Gln	Leu	Ala	Asn	Ile	Ala	Arg	Gly	Gly	Tyr	Val	Leu	Lys	530	535	540	
Asp	Cys	Ala	Gly	Gln	Pro	Glu	Leu	Ile	Phe	Ile	Ala	Thr	Gly	Ser	Glu	545	550	555	560

Val Glu Leu Ala Val Ala Ala Tyr Glu Lys Leu Thr Ala Glu Gly Val
565 570 575

Lys Ala Arg Val Val Ser Met Pro Ser Thr Asp Ala Phe Asp Lys Gln
580 585 590

Asp Ala Ala Tyr Arg Glu Ser Val Leu Pro Lys Ala Val Thr Ala Arg
595 600 605

Val Ala Val Glu Ala Gly Ile Ala Asp Tyr Trp Tyr Lys Tyr Val Gly
610 615 620

Leu Asn Gly Ala Ile Ile Gly Met Asn Thr Phe Gly Glu Ser Ala Pro
625 630 635 640

Ala Glu Gln Leu Phe Glu Glu Phe Gly Phe Thr Val Asp Asn Val Val
645 650 655

Ala Lys Ala Lys Ala Leu Leu
660

5 <210> 61
<211> 51
<212> ADN
<213> Artificial

10 <220>
<223> oligonucleótido

<400> 61
ggtaccaagg aaggactgat catgagttct gaaagtagtc agggcttagt c 51

15 <210> 62
<211> 28
<212> ADN
<213> Artificial

20 <220>
<223> oligonucleótido

<400> 62
ctgcagctac ttctgggagc gccacatc 28

25 <210> 63
<211> 46
<212> ADN
<213> Artificial

30 <220>
<223> oligonucleótido

<400> 63
ctgcagaagg aaggactgat cgtgcgtata ggtatcgatt taggcg 46

35 <210> 64
<211> 30

<212> ADN
 <213> Artificial

 <220>
 5 <223> oligonucleótido

 <400> 64
 tctagattac tctgtggcc ataaccacgc 30

 10 <210> 65
 <211> 51
 <212> ADN
 <213> Artificial

 15 <220>
 <223> oligonucleótido

 <400> 65
 20 tctagaaagg aaggtcgact catgcaaaaa ctcattaact cagtgcaaaa c 51

 <210> 66
 <211> 35
 <212> ADN
 <213> Artificial
 25
 <220>
 <223> oligonucleótido

 <400> 66
 30 gagctcttac agcttgtgt aaacacgcgc taaac 35

 <210> 67
 <211> 28
 <212> ADN
 35 <213> Artificial

 <220>
 <223> oligonucleótido

 <400> 67
 40 gttctatcga ttgggtaat gccagtgg 28

 <210> 68
 <211> 26
 45 <212> ADN
 <213> Artificial

 <220>
 <223> oligonucleótido
 50
 <400> 68
 ggaacatctg cggcgcataa tacagc 26

 <210> 69
 55 <211> 21
 <212> ADN
 <213> Artificial

 <220>
 60 <223> oligonucleótido

 <400> 69
 gcgaggtgc agcgggaagt g 21

 65 <210> 70
 <211> 27

	<212> ADN		
	<213> Artificial		
	<220>		
5	<223> oligonucleótido		
	<400> 70		
	aatttcactg ccttcagcg catgtcc	27	
10	<210> 71		
	<211> 22		
	<212> ADN		
	<213> Artificial		
15	<220>		
	<223> oligonucleótido		
	<400> 71		
20	cgctgacgcc ttccttgcg at	22	
	<210> 72		
	<211> 24		
	<212> ADN		
	<213> Artificial		
25	<220>		
	<223> oligonucleótido		
	<400> 72		
30	gcggtgtttc agcgaacagg aaca	24	
	<210> 73		
	<211> 23		
	<212> ADN		
35	<213> Artificial		
	<220>		
	<223> oligonucleótido		
40	<400> 73		
	cgactcacta tagggagagc ggc	23	
	<210> 74		
	<211> 24		
45	<212> ADN		
	<213> Artificial		
	<220>		
	<223> oligonucleótido		
50	<400> 74		
	aagaacatcg atttccatg gcag	24	
	<210> 75		
55	<211> 17		
	<212> ADN		
	<213> Artificial		
	<220>		
60	<223> oligonucleótido		
	<400> 75		
	gtaaaacgac ggccagt	17	
65	<210> 76		
	<211> 25		

	<212> ADN		
	<213> Artificial		
	<220>		
5	<223> oligonucleótido		
	<400> 76		
	ctcgagccgg aattgccagc tgggg	25	
10	<210> 77		
	<211> 30		
	<212> ADN		
	<213> Artificial		
15	<220>		
	<223> oligonucleótido		
	<400> 77		
20	catatgaaac gatcctcatc ctgtctcttg	30	
	<210> 78		
	<211> 35		
	<212> ADN		
	<213> Artificial		
25	<220>		
	<223> oligonucleótido		
	<400> 78		
30	catatgacaa agtatgcatt agtcggtgat gtggg	35	
	<210> 79		
	<211> 42		
	<212> ADN		
35	<213> Artificial		
	<220>		
	<223> oligonucleótido		
40	<400> 79		
	gtcgacttac agaattgtgac ctaaggtctg gcgtaaatgt gc	42	
	<210> 80		
	<211> 52		
45	<212> ADN		
	<213> Artificial		
	<220>		
	<223> oligonucleótido		
50	<400> 80		
	ggatccatga ggatcgtttc gcatgagttc tgaaagtagt cagggtctag tc	52	
	<210> 81		
55	<211> 34		
	<212> ADN		
	<213> Artificial		
	<220>		
60	<223> oligonucleótido		
	<400> 81		
	gagctcctac ttctgggagc gccacatctc ctcg	34	
65	<210> 82		
	<211> 45		

<212> ADN
 <213> Artificial

 <220>
 5 <223> oligonucleótido

 <400> 82
 attagggccc aaggaggta cagcatgcaa gcctatttg accag 45

 10 <210> 83
 <211> 31
 <212> ADN
 <213> Artificial

 15 <220>
 <223> oligonucleótido

 <400> 83
 20 attgccggg ttattgtcg aacagataat g 31

 <210> 84
 <211> 45
 <212> ADN
 <213> Artificial
 25
 <220>
 <223> oligonucleótido

 <400> 84
 30 attgccggg aaggaggta cagcatgtat atcgggatag atctt 45

 <210> 85
 <211> 31
 <212> ADN
 35 <213> Artificial

 <220>
 <223> oligonucleótido

 <400> 85
 40 ccgactagt ttacgcatt aatggcagaa g 31

 <210> 86
 <211> 44
 45 <212> ADN
 <213> Artificial

 <220>
 <223> oligonucleótido
 50
 <400> 86
 ccgactagt aaggaggta cagcatgaat acccagtata attc 44

 <210> 87
 55 <211> 31
 <212> ADN
 <213> Artificial

 <220>
 60 <223> oligonucleótido

 <400> 87
 attgccggg ttacagcgtg gcagttgtt g 31

 65 <210> 88
 <211> 46

<212> ADN
 <213> Artificial

 <220>
 5 <223> oligonucleótido

 <400> 88
 cccgactagt aaggaggcta cagcatgaaa ataaagaaca ttctac 46

 10 <210> 89
 <211> 30
 <212> ADN
 <213> Artificial

 15 <220>
 <223> oligonucleótido

 <400> 89
 20 attgccgcgg tcaagaacgg cgtttggtg 30

 <210> 90
 <211> 24
 <212> ADN
 <213> Artificial
 25
 <220>
 <223> oligonucleótido

 <400> 90
 30 aaaaccgcgg cggctcggaa gtcg 24

 <210> 91
 <211> 23
 <212> ADN
 35 <213> Artificial

 <220>
 <223> oligonucleótido

 <400> 91
 40 cgactcacta tagggagagc ggc 23

 <210> 92
 <211> 24
 45 <212> ADN
 <213> Artificial

 <220>
 <223> oligonucleótido
 50
 <400> 92
 aagaacatcg attttcatg gcag 24

 <210> 93
 55 <211> 17
 <212> ADN
 <213> Artificial

 <220>
 60 <223> oligonucleótido

 <400> 93
 gtaaaacgac ggccagt 17

 65 <210> 94
 <211> 17

	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> oligonucleótido	
	<400> 94	
	gtaaaacgac ggccagt	17
10	<210> 95	
	<211> 16	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> oligonucleótido	
	<400> 95	
20	aacagctatg accatg	16

REIVINDICACIONES

1. Un microorganismo de los géneros *Cupriavidus* o *Ralstonia* modificado genéticamente para expresar:
 - 5 xilosa isomerasa (EC 5.3.1.5), xiluloquinasa (EC 2.7.1.17) y Xyl G (EC 3.6.3.17), Xyl F y Xyl H, en donde dicho microorganismo es capaz de crecer en arabinosa y, opcionalmente, también en glucosa y/o galactosa como fuente de carbono.
 2. El microorganismo de acuerdo con la reivindicación 1, en donde el microorganismo está modificado genéticamente para expresar un cotransportador unidireccional E de protones de xilosa.
 3. El microorganismo de acuerdo con las reivindicaciones 1 o 2, en donde el microorganismo está modificado genéticamente para expresar: fosfomanosa isomerasa (EC 5.3.1.8) y proteína de difusión facilitada para la absorción de manosa (EC 1.3.1.74) y, opcionalmente, manofructoquinasa (EC 2.7.1.4), en donde dicho microorganismo es capaz de crecer en manosa y, opcionalmente, también en glucosa como fuente de carbono.
 4. El microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en donde el microorganismo está modificado genéticamente para expresar transaldolasa (EC 2.2.1.2) y/o transcetolasa (EC 2.2.1.1).
 - 20 5. El microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en donde el microorganismo está modificado genéticamente para expresar glucoquinasa (EC 2.7.1.2).
 6. El microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en donde el microorganismo pertenece a la especie *R. eutropha* (*C. necator*).
 - 25 7. Un vector que comprende una secuencia de nucleótidos que codifica xilosa isomerasa (EC 5.3.1.5), una secuencia de nucleótidos que codifica xiluloquinasa (EC 2.7.1.17) y una secuencia de nucleótidos que codifica Xyl G (EC 3.6.3.17), Xyl F y Xyl H, y elementos reguladores para regular la expresión de dichos genes en *Cupriavidus* o *Ralstonia*.
 - 30 8. El vector de acuerdo con la reivindicación 7 que comprende una secuencia de nucleótidos que codifica un cotransportador unidireccional E de protones de xilosa.
 9. El vector de acuerdo con las reivindicaciones 7 u 8, que comprende una secuencia de nucleótidos que codifica fosfomanosa isomerasa (EC 5.3.1.8) y una secuencia de nucleótidos que codifica una proteína de difusión facilitada para la absorción de manosa (EC 1.3.1.74) y, opcionalmente, una secuencia de nucleótidos que codifica manofructoquinasa (EC 2.7.1.4), y elementos reguladores para regular la expresión de dichos genes en *Cupriavidus* o *Ralstonia*.
 - 35 10. El vector de acuerdo con una cualquiera de las reivindicaciones 7 a 9 que comprende una secuencia de nucleótidos que codifica transaldolasa (EC 2.2.1.2) y/o transcetolasa (EC 2.2.1.1).
 - 40 11. El vector de acuerdo con una cualquiera de las reivindicaciones 7 a 10 que comprende un gen que codifica glucoquinasa (EC 2.7.1.2).
 - 45 12. El vector de acuerdo con una cualquiera de las reivindicaciones 7 a 11, en el que el elemento regulador comprende un potente promotor constitutivo.
 13. El vector de acuerdo con una cualquiera de las reivindicaciones 7 a 12, en donde el vector es un vector de amplia gama de hospedadores.
 - 50 14. El vector de acuerdo con una cualquiera de las reivindicaciones 7 a 13, en el que el promotor es el promotor de neokanamicina, el promotor lac, el promotor de gliceraldehído-3-fosfato deshidrogenasa, el P2, el promotor hidrogenasa (SH) o el PL.
 - 55 15. Un sistema vectorial que comprende
 - (i) el vector de las reivindicaciones 7 u 8,
 - (ii) un vector que comprende una secuencia de nucleótidos que codifica fosfomanosa isomerasa (EC 5.3.1.8) y una secuencia de nucleótidos que codifica una proteína de difusión facilitada para la absorción de manosa (EC 1.3.1.74) y, opcionalmente, una secuencia de nucleótidos que codifica manofructoquinasa (EC 2.7.1.4) y elementos reguladores para regular la expresión de dichos genes en *Cupriavidus* o *Ralstonia*
 - 60 (iii) y, opcionalmente, un vector que comprende una secuencia de nucleótidos que codifica transaldolasa (EC 2.2.1.2) y/o transcetolasa (EC 2.2.1.1) y, opcionalmente, un gen que codifica glucoquinasa (EC 2.7.1.2).
 - 65 16. Un método para cultivar el microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 6, o un microorganismo *Cupriavidus* o *Ralstonia* que comprende el vector de acuerdo con una cualquiera de las

reivindicaciones 7 a 14 o el sistema vectorial de acuerdo con la reivindicación 15, que comprende que el medio de cultivo comprende manosa, glucosa, xilosa, arabinosa o galactosa o una combinación de las mismas como fuente de carbono.

5 17. El método de acuerdo con la reivindicación 16, que comprende que el microorganismo puede sintetizar, opcionalmente después de la modificación genética, un producto químico, como un lípido, y el producto químico se recupera del medio de cultivo o de las células.

10 18. El método de acuerdo con las reivindicaciones 16 o 17, que comprende que el microorganismo puede sintetizar polihidroxialcanoatos, tales como PHB o ácidos grasos y/o derivados de ácidos grasos.

15 19. Uso del microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 6, posiblemente después de una modificación genética, para producir un producto químico seleccionado de lípidos, polihidroxialcanoatos, ácidos grasos y derivados de ácidos grasos.

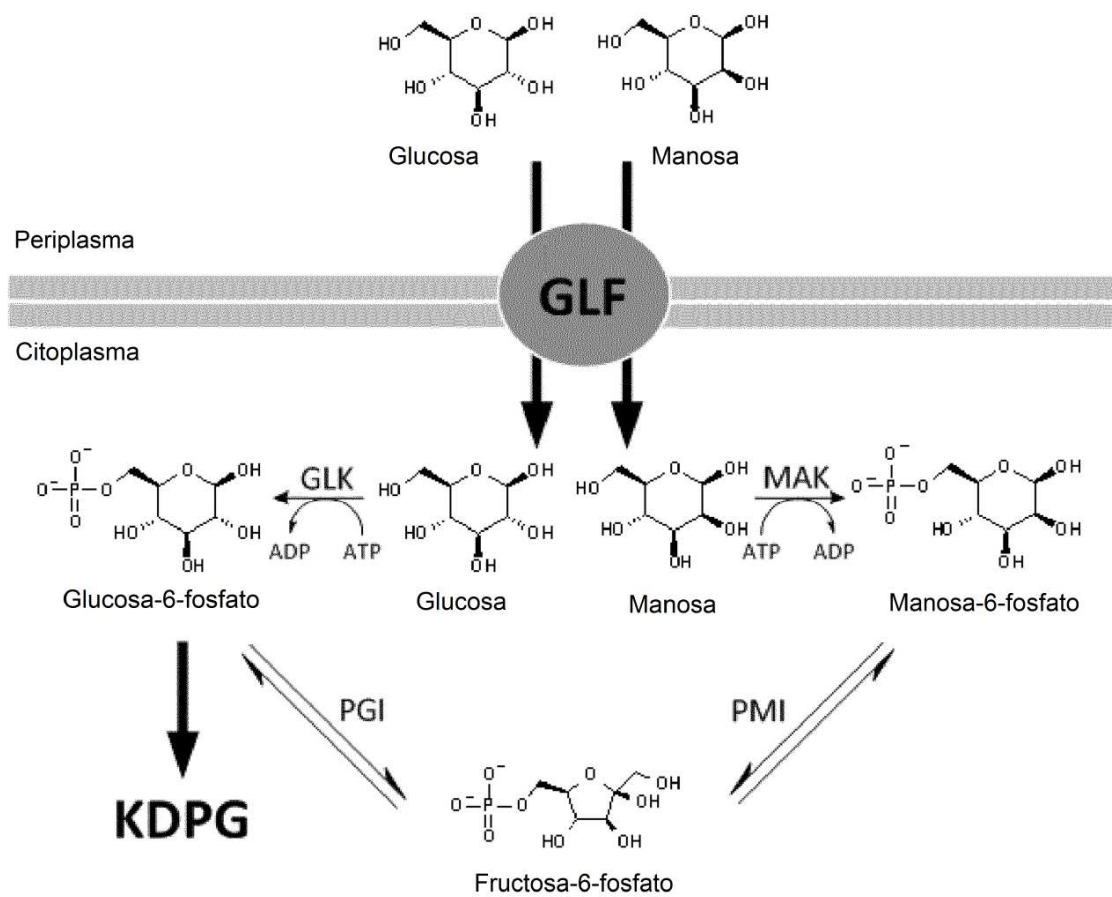


Fig. 1

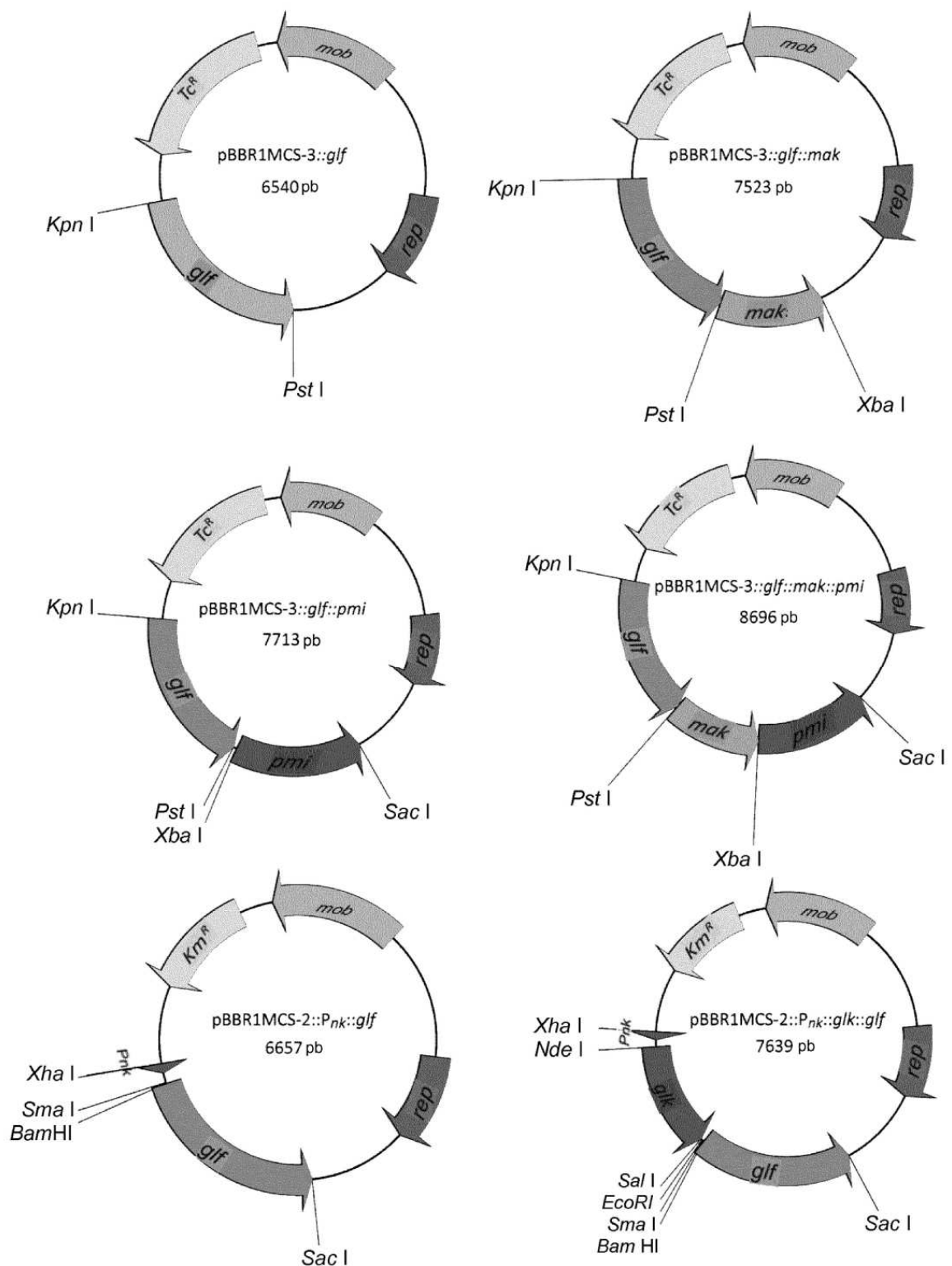


Fig. 2

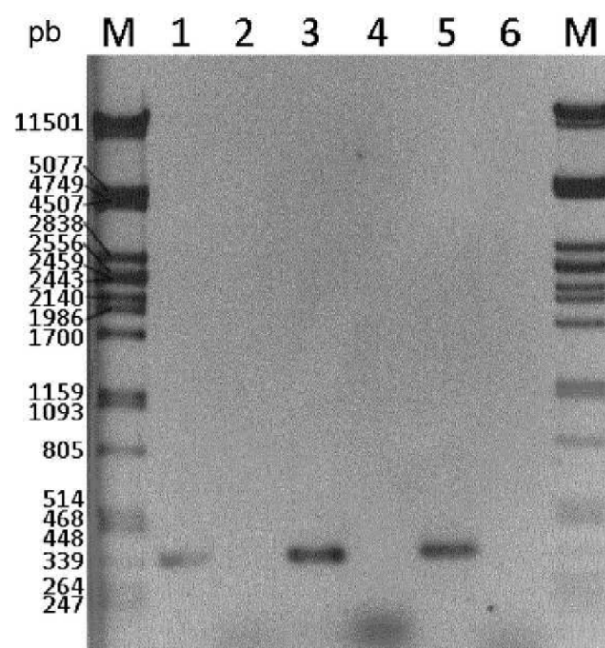


Fig. 3

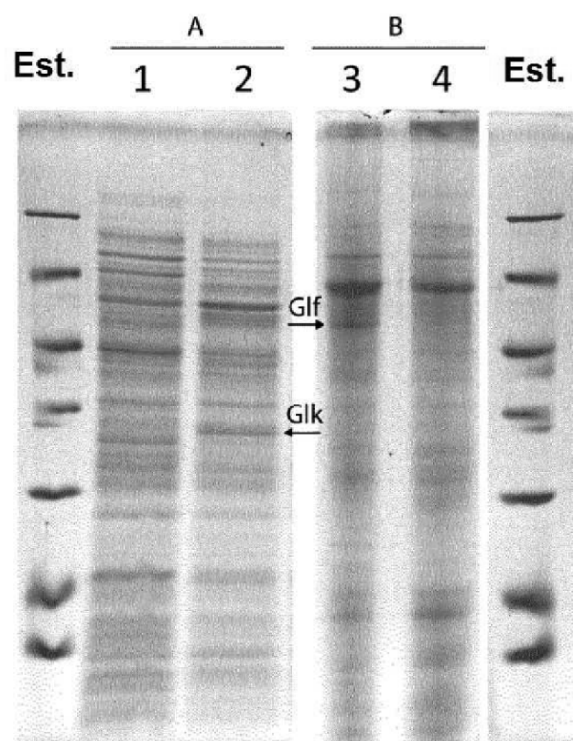


Fig. 4

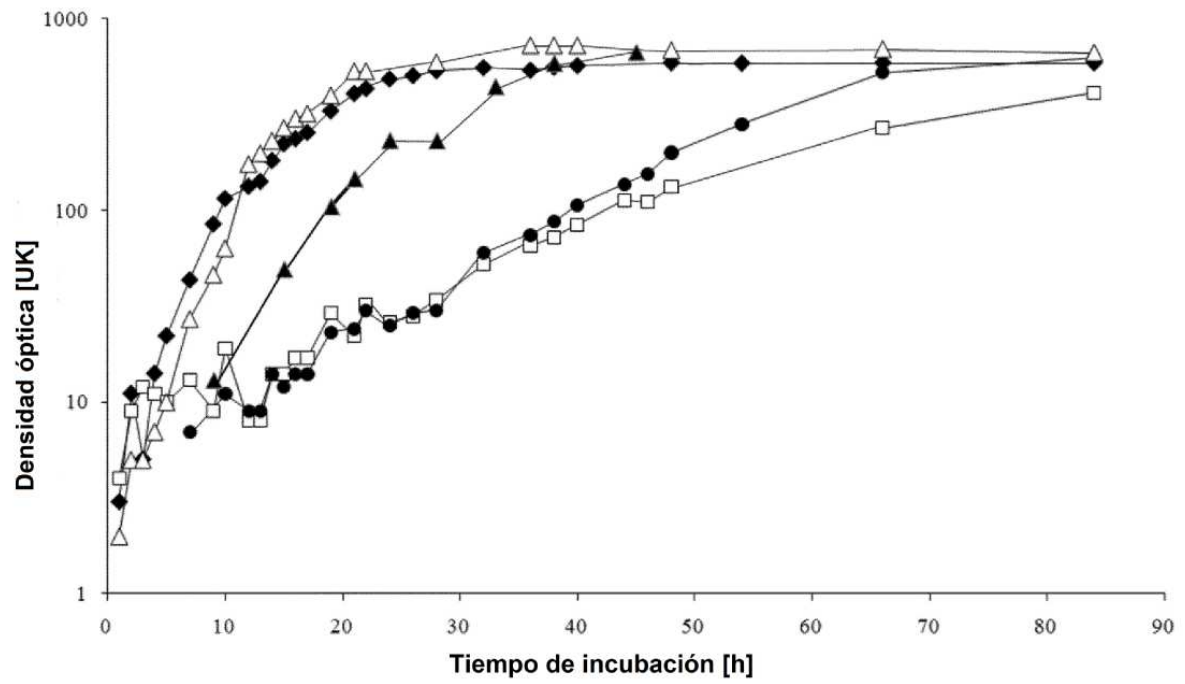


Fig. 5

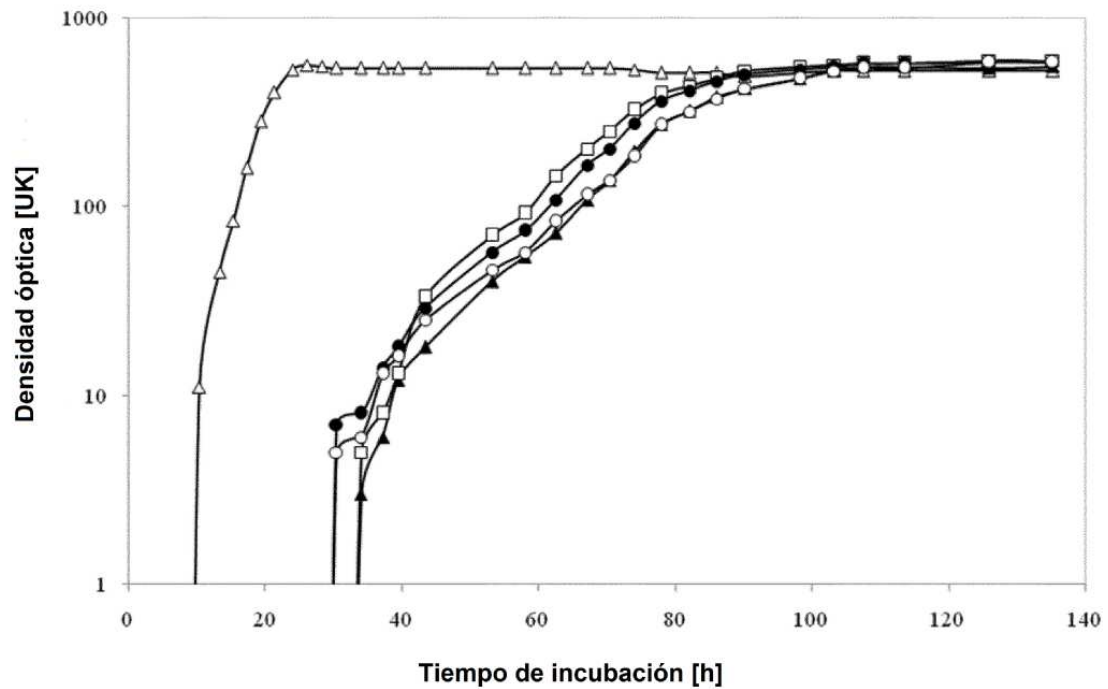


Fig. 6

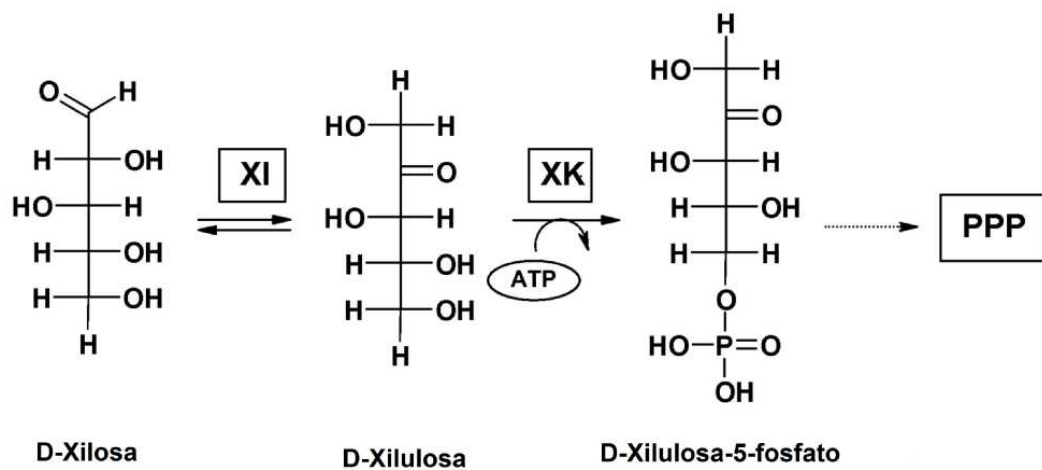


Fig. 7

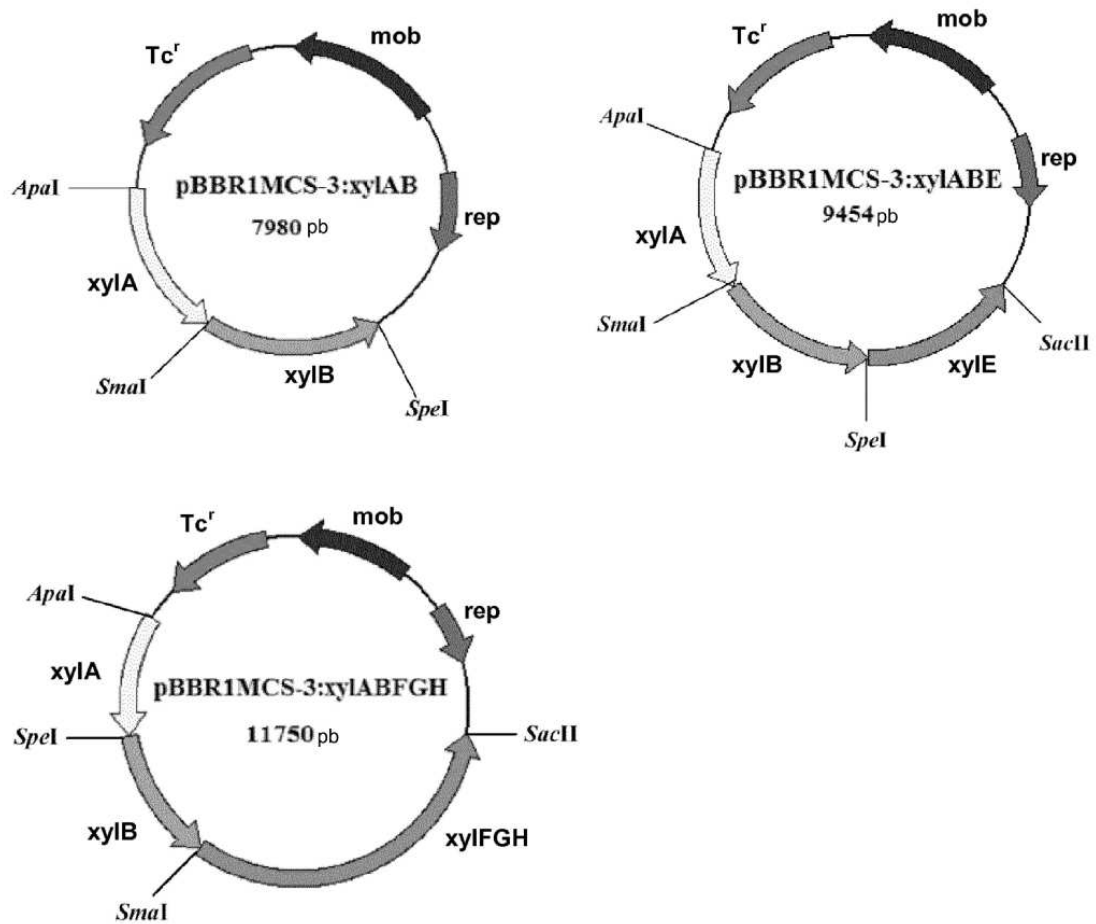
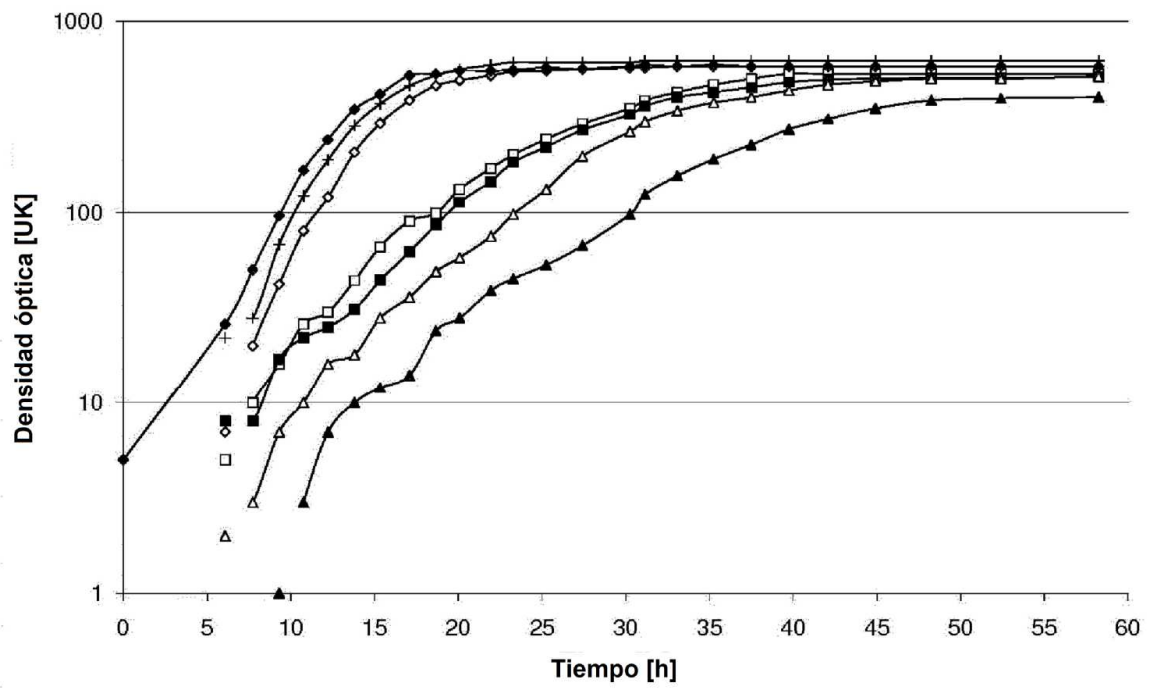


Fig. 8

**Fig. 9**