



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

**Generación y caracterización de una cepa de *Escherichia coli*
sobreproductora de shikimato mediante la implementación
de estrategias de ingeniería metabólica**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

M.C. José Alberto Rodríguez Ruiz

TUTOR PRINCIPAL:

Dr. Francisco Gonzalo Bolívar Zapata
[Instituto de Biotecnología - UNAM](#)

MIEMBROS DEL COMITÉ TUTOR:

Dr. José Luis Báez Viveros
[Centro de Investigación en Biotecnología - UAEM](#)
Dr. Baltazar Becerril Luján
[Instituto de Biotecnología - UNAM](#)

México, D.F., marzo de 2015



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Integrantes del jurado de examen de grado

Dra. Esperanza Martínez Romero

Departamento de Ecología Genómica, Centro de Ciencias Genómicas, UNAM

Dra. Laura Alicia Palomares Aguilera

Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología, UNAM

Dra. Marcela Ayala Aceves

Departamento de Ingeniería Celular y Biocatálisis, Instituto de Biotecnología, UNAM

Dr. Juan Carlos Sigala Alanís

Departamento de Procesos y Tecnología, UAM - Cuajimalpa

Dr. Daniel Genaro Segura González

Departamento de Microbiología Molecular, Instituto de Biotecnología, UNAM

Este trabajo fue realizado principalmente en las instalaciones del Instituto de Biotecnología - UNAM en Cuernavaca, México y complementado con una estancia de 5 meses en el Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés - INSA en Toulouse, Francia.

AGRADECIMIENTOS

A mis padres, por todo su apoyo y cariño.

Al Dr. Francisco Bolívar, por brindarme su confianza y guía.

A la Dra. Noemí Flores, por toda su ayuda durante mi formación profesional.

Al M.C. Juan Andrés Martínez por su valiosa colaboración en este proyecto.

A los Dres. José Luis Báez y Baltazar Becerril, por su apoyo en la planeación y evaluación del proyecto como miembros del comité tutorial.

Al Dr. Guillermo Gosset por aportar ideas y comentarios que contribuyeron de manera significativa al enriquecimiento de este trabajo.

Al Dr. Fabien Létisse por su hospitalidad al recibirme en la plataforma MetaToul del LISBP en el INSA-Toulouse como parte de una estancia académica.

A las Dras. Esperanza Martínez, Laura Palomares y Marcela Ayala, y a los Dres. Juan Carlos Sigala y Daniel Segura, por aceptar formar parte del jurado de examen de grado.

A Mercedes Enzaldo y Aurelia González, por el auxilio técnico, y a Paul Gaytán, Eugenio López y Jorge Yáñez, por su colaboración como integrantes de la unidad de síntesis y secuenciación de DNA del Instituto de Biotecnología - UNAM.

Al Programa de Apoyo a los Estudios de Posgrado (PAEP) de la UNAM y al CONACyT por los apoyos económicos otorgados durante mis estudios de doctorado.

A los integrantes del laboratorio de Ingeniería de Vías Metabólicas en el Instituto de Biotecnología - UNAM.

ÍNDICE

RESUMEN, **9**

SUMMARY, **10**

INTRODUCCIÓN, **11**

ANTECEDENTES, **15**

HIPÓTESIS, **19**

OBJETIVO GENERAL, **19**

OBJETIVOS PARTICULARES, **19**

RESULTADOS Y DISCUSIÓN, **20**

1. Construcción del vector y modificación cromosomal de la cepa PB12 para promover la acumulación de SA, **20**
 - 1.1. Construcción del plásmido pTrc327par para ser usado como vector de expresión de los genes de interés, **20**
 - 1.2. Ensamblaje de un operón sintético con los genes necesarios para favorecer la sobreproducción de SA bajo el control de un mismo promotor, **23**
 - 1.3. Generación de cepas derivadas de PB12 que propicien un aumento en el flujo de carbono hacia SA y su acumulación en el medio extracelular, **24**
2. Evaluación del crecimiento, consumo de Glc y producción de metabolitos de las cepas generadas, **28**
 - 2.1. Determinación de los efectos causados por la inactivación conjunta de PTS y *pykF* en cepas que expresan el operón Aro6, **28**
 - 2.2. Obtención de los perfiles de fermentación de la cepa AR36 con tres diferentes concentraciones de sustratos, **31**
3. Determinación de los efectos causados por el sistema de expresión sobre el metabolismo central mediante un análisis transcriptómico y metabolómico de la mejor cepa productora, **40**
 - 3.1. Perfiles transcripcionales de genes relevantes en la cepa AR36, **40**
 - 3.1.1. *Genes glucolíticos*, **42**
 - 3.1.2. *Genes del metabolismo de acetato*, **44**
 - 3.1.3. *Genes contenidos en el operón Aro6*, **45**
 - 3.1.4. *Genes del ciclo de los ácidos tricarboxílicos (TCA)*, **45**
 - 3.1.5. *Genes de la vía de las pentosas fosfato (PPP)*, **47**
 - 3.1.6. *Genes del nodo de PEP*, **48**
 - 3.2. Perfiles metabolómicos en respuesta a la presencia de plásmido y a la expresión del operón sintético Aro6, **49**

CONCLUSIONES, **56**

PERSPECTIVAS, **58**

MATERIALES Y MÉTODOS, **58**

REFERENCIAS, **67**

MATERIAL ADICIONAL 1. Secuencia nucleotídica del operón sintético construido en este trabajo y presente en el plásmido pTrcAro6, **77**

MATERIAL ADICIONAL 2. Esquema detallado de la construcción del plásmido pTrcAro6, **80**

MATERIAL ADICIONAL 3. Valores de transcripción relativa obtenidos por qPCR, correspondientes al promedio y desviación estándar de tres mediciones hechas para cada una de las dos réplicas biológicas, **81**

MATERIAL ADICIONAL 4. Oligonucleótidos utilizados en este trabajo, **82**

MATERIAL ADICIONAL 5. Perfil de producción de SA obtenido al cultivar a la cepa AR36 en un fermentador de 2 L en modo lote alimentado, suministrando simultáneamente Glc y YE durante 14 horas, **83**

ARTÍCULOS PUBLICADOS

ÍNDICE DE FIGURAS

Figura 1. Modificaciones genéticas empleadas en este trabajo para promover la producción de shikimato a partir de glucosa en *E. coli*, **13**

Figura 2. Concentraciones de SA y otros dos compuestos intermediarios de su vía biosintética, DAHP y DHS, obtenidas en PB12 y sus derivadas al ser transformadas con los plásmidos pJLBaroG^{br}tklA y pTOPOaroBaroE, **17**

Figura 3. Esquema simplificado de los pasos requeridos en la construcción del plásmido pTrcAro6, que contiene 6 genes sintéticos bajo el control del promotor Trc, **22**

Figura 4. Comprobación de la cepa AR3 transformada con el plásmido pTrc327par (AR3e) y con el plásmido conteniendo el operón sintético Aro6 (AR36) por visualización de DNA en geles de agarosa, **26**

Figura 5. Curvas de amplificación obtenidas por qPCR usando oligonucleótidos específicos para cada uno de los genes evaluados, **27**

Figura 6. Comportamiento de las cepas AR26, AR36 y sus derivadas con plásmido vacío, AR2e (*pykF*⁺) y AR3e (*pykF*⁻), usando matraces con 15 g/L de Glc y 5 g/L de YE y fermentadores de 1 L con 100 g/L de Glc y 15 g/L de YE, **29**

Figura 7. Perfiles de fermentación de la cepa AR36 cultivada en fermentadores de 1 L con tres diferentes concentraciones de sustratos, **33**

Figura 8. Subproductos aromáticos de la vía del SA detectados en las fermentaciones de 1 L con la cepa AR36 usando tres diferentes concentraciones de sustratos, **34**

Figura 9. Porcentaje molar de cada compuesto aromático producido en la cepa AR36 con respecto al total en cultivos en lote con las siguientes concentraciones iniciales de sustratos: 50 g/L de Glc y 15 g/L de YE; 100 g/L de Glc y 15 g/L de YE; 100 g/L de Glc y 30 g/L de YE, **37**

Figura 10. Representación de las reacciones catalizadas por las enzimas codificadas en los 40 genes evaluados en el análisis transcriptómico, **41**

Figura 11. Cambios transcripcionales ocasionados por la expresión del operón sintético Aro6 sobre genes glucolíticos y de metabolismo de acetato, **43**

Figura 12. Cambios transcripcionales ocasionados por la expresión del operón sintético Aro6 sobre genes del TCA, PPP y nodo de PEP, **46**

Figura 13. Perfiles de consumo de glucosa y parámetros de producción de SA obtenidos para las tres cepas en el sistema de células en reposo, **51**

Figura 14. Metabolómica comparativa de las tres cepas estudiadas, **53**

ÍNDICE DE TABLAS

Tabla 1. Plásmidos y cepas de *Escherichia coli* empleados en este trabajo, **25**

Tabla 2. Datos comparativos de las fermentaciones con las cepas AR26 y AR36, usando 100 g/L de Glc y 15 g/L de YE como sustratos, **31**

Tabla 3. Comparación de metabolitos medidos y parámetros cinéticos y estequiométricos calculados entre tres fermentaciones de la cepa AR36 con diferentes concentraciones de sustratos, **36**

Tabla 4. Cepas relevantes de *E. coli* que han sido modificadas para favorecer la producción de SA bajo diferentes condiciones de cultivo, **39**

NOMENCLATURA UTILIZADA

μ	velocidad específica de crecimiento
1,3BPG	1,3-bisfosfoglicerato
6-PG	gluconato 6-fosfato
AAA	aminoácidos aromáticos
AEC	carga energética de adenilatos
CHO	corismato
DAHP	3-desoxi-D-arabinoheptulosonato-7-fosfato
DCW	peso celular seco
DHQ	3-deshidroquinato
DHS	3-deshidros hikimato
DOT	tensión de oxígeno disuelto
E4P	eritrosa 4-fosfato
EE	fase exponencial temprana
F1,6BP	fructosa 1,6-bisfosfato
F6P	fructosa 6-fosfato
G1P	glucosa 1-fosfato

G3P	gliceraldehído 3-fosfato
GA	ácido gálico
Glc	glucosa
GlcNAc-1P	N-acetilglucosamina 1-fosfato
GLG	glucógeno
IDMS	dilución isotópica para espectrometría de masas
IPTG	isopropil- β -D-1-thiogalactopiranosido
LE	fase exponencial tardía
MCS	sitio múltiple de clonación
PEP	fosfoenolpiruvato
PPP	vía de las pentosas fosfato
PTS	sistema de fosfoenolpiruvato:carbohidrato fosfotransferasa
QA	ácido quínico
qp	velocidad específica de producción
Qp	velocidad volumétrica de producción
qs	velocidad específica de consumo
Qs	velocidad volumétrica de consumo
R1P	ribosa 1-fosfato
R5P	ribosa 5-fosfato
RBS	sitio de unión ribosomal
Ru5P	ribulosa 5-fosfato
S3P	shikimato 3-fosfato
S7P	sedoheptulosa 7-fosfato
SA	shikimato
ST	fase estacionaria
TAC	total de compuestos aromáticos
TCA	ciclo de los ácidos tricarboxílicos
X	biomasa
Xu5P	xilulosa 5-fosfato
Y	rendimiento
YE	extracto de levadura

RESUMEN

La producción microbiana de shikimato (SA) es una estrategia importante para alcanzar altas productividades volumétricas de este compuesto. A pesar de que la implementación de técnicas pertenecientes a la ingeniería metabólica y de bioprocesos han resultado en diversas cepas de *Escherichia coli* capaces de acumular SA, los rendimientos obtenidos hasta la fecha se encuentran todavía lejos del máximo teórico. Para promover una canalización de carbono más eficiente hacia la formación de SA, en este trabajo se desarrolló una plataforma alternativa basada en la expresión constitutiva en plásmido de seis genes en forma de operón, seleccionados de las vías de síntesis de aminoácidos aromáticos y pentosas fosfato. El plásmido construido, nombrado pTrcAro6, fue transformado en una cepa con inactivaciones de genes codificantes para componentes del sistema de fosfotransferasa (*ptsHIcrr*), las enzimas shikimato cinasas I y II (*aroK* y *aroL*), la enzima piruvato cinasa I (*pykF*) y el represor de lactosa (*lacI*). La cepa de producción resultante, llamada AR36, fue evaluada en fermentadores en modo lote de 1 L, usando medio mineral con concentraciones variables de glucosa y extracto de levadura. La medición del crecimiento y de las concentraciones extracelulares de SA, glucosa, acetato e intermediarios de la vía del SA, permitió una comparación de las productividades y rendimientos obtenidos en cada caso, resultando en una caracterización cinética y estequiométrica detallada. El siguiente paso involucró la combinación de esta información con perfiles transcriptómicos y metabolómicos de la cepa AR36, con el objetivo de alcanzar un mejor entendimiento de las consecuencias fisiológicas de emplear dicha combinación de plataforma de expresión, fondo genético y condiciones de cultivo. Los rendimientos de SA alcanzados sobrepasaron a todos los valores previamente reportados y postulan a la estrategia descrita en este trabajo como una alternativa interesante para la sobreproducción de SA.

SUMMARY

The microbial production of shikimate (SA) has been recognized as an important approach to achieve high volumetric productivities. Although the implementation of strain and process engineering strategies has resulted in diverse *Escherichia coli* strains capable of SA accumulation, the yields obtained to date are still far from the theoretical maximum. To promote a more efficient carbon channeling towards SA formation, we developed an alternative platform based on the plasmid-driven constitutive expression of six genes selected from the pentose phosphate and aromatic amino acid pathways, artificially arranged as an operon. The constructed plasmid, named pTrcAro6, was transformed into a modified strain carrying inactivated genes coding for phosphotransferase system components (*ptsHIcrr*), shikimate kinases I and II (*aroK* and *aroL*), pyruvate kinase I (*pykF*) and the lactose operon repressor (*lacI*). The resulting production strain, called AR36, was evaluated in 1 L batch fermentors using mineral media with variable concentrations of glucose and yeast extract. Measurements of growth and extracellular concentrations of SA, glucose, acetate and SA pathway intermediates, permitted a comparison of the productivities and yields attained for each case, resulting in a detailed kinetic and stoichiometric characterization. The next step involved the combination of this information with transcriptomic and metabolomic profilings of strain AR36, in order to achieve a better understanding of the physiological consequences of using such expression platform, genetic background and culture conditions. The attained SA yields surpassed all previously reported values and postulate the strategy described in this work as an interesting alternative for SA overproduction.

INTRODUCCIÓN

El ácido shikímico o shikimato (SA) es un compuesto intermediario en la vía biosintética de aminoácidos aromáticos (AAA) en plantas y bacterias. Este metabolito es utilizado principalmente como material de partida para la síntesis química de oseltamivir fosfato (Tamiflu[®]), empleado en el tratamiento contra la influenza [1]. Este fármaco actúa como inhibidor de la enzima neuraminidasa, presente en la cápside de los virus de influenza estacional, influenza aviar H5N1 e influenza humana H1N1, con lo que se impide la liberación de partículas virales de células infectadas [2]. Ante la latente amenaza de una pandemia de influenza y la posible aparición de nuevas cepas del virus, se ha reconocido la necesidad de aumentar las reservas mundiales de Tamiflu[®] [3]. Uno de los factores limitantes de la producción a gran escala de este fármaco es que el aislamiento tradicional de SA a partir de plantas del género *Illicium* resulta insuficiente para cubrir la demanda [4]. Considerando esta limitación, se ha planteado que la síntesis microbiana de SA a partir de recursos renovables como glucosa (Glc) constituye una excelente alternativa para obtener mayores productividades volumétricas [5]. Nuestro laboratorio ha incursionado en este campo mediante un proyecto multidisciplinario que involucra la generación y caracterización de cepas de *Escherichia coli* productoras de SA [6, 7]. Se está trabajando en conjunto con otros grupos para que la tecnología de producción de oseltamivir a partir de Glc pueda estar disponible en caso de una emergencia de salud en México [8]. Además de su importancia en síntesis orgánica, el SA es precursor de metabolitos secundarios y otros productos naturales derivados de la vía de AAA en diversos organismos. De esta manera, una cepa sobreproductora de SA puede emplearse como punto de partida para sintetizar otros compuestos de interés industrial, incluyendo anticancerígenos, suplementos alimenticios, analgésicos, antioxidantes y pigmentos [9]. Diversas estrategias se han llevado a cabo en las últimas dos décadas con el objetivo de mejorar los rendimientos y productividades de SA en *E. coli* [10, 11]. Los enfoques más utilizados contemplan incrementar la disponibilidad de los precursores metabólicos directos de los compuestos aromáticos, la eritrosa 4-fosfato (E4P) y el fosfoenolpiruvato (PEP), mediante la inactivación de enzimas que los consumen o la sobreexpresión de enzimas que los producen [12, 13]. Las modificaciones genéticas comúnmente aplicadas en la vía del SA

(Fig. 1) comprenden la sobreexpresión de las enzimas DAHP sintetas, shikimato deshidrogenasa y DHQ sintasa (codificadas por los genes *aroFGH*, *aroE* y *aroB*, respectivamente), así como el truncamiento de la vía justo después de la formación de SA mediante la delección de los genes codificantes para las isoenzimas shikimato cinasas (*aroK* y *aroL*). La reacción catalizada por las DAHP sintetas es el primer paso comprometido en la biosíntesis de AAA y controla la cantidad de carbono proveniente de intermediarios del metabolismo central que se dirige hacia esta vía [14]. Se sabe que las DAHP sintetas están sujetas a inhibición alostérica por cada uno de los tres AAA (L-tirosina, L-fenilalanina y L-triptófano) y por lo tanto, para sobreproducir compuestos derivados de esta vía es necesario utilizar enzimas con modificaciones que les confieran resistencia al efecto de retroinhibición (fbr: *feed-back resistant*) [15]. Por otro lado, en un intento de incrementar la disponibilidad intracelular de PEP y aumentar el flujo de carbono hacia la vía de AAA, también se han evaluado cepas con inactivaciones en los genes codificantes para las isoenzimas piruvato cinasas, *pykF* y *pykA*, o en el sistema fosfoenolpiruvato:carbohidrato fosfotransferasa (PTS) (Fig. 1) [6, 10, 16].

E. coli utiliza al PTS como sistema principal de internalización y fosforilación de Glc desde el periplasma hacia el citoplasma, convirtiendo una molécula de PEP a piruvato por cada molécula generada de glucosa 6-fosfato [17]. Sin embargo, se ha determinado que cuando *E. coli* crece en medio mineral con Glc como única fuente de carbono, el PTS consume aproximadamente 50% del PEP disponible, lo que ha promovido la implementación de estrategias para aumentar la disponibilidad de PEP y consecuentemente el rendimiento máximo teórico de producción de compuestos aromáticos [18]. Una estrategia común es el reemplazo de las capacidades de transporte y fosforilación de Glc del PTS por enzimas alternativas, como el facilitador de Glc y la glucocinasa de *Zymomonas mobilis* (codificados por *glf* y *glk*, respectivamente) [10, 19, 20], la permeasa de galactosa y la glucocinasa de *E. coli* (codificados por *galP* y *glk*, respectivamente) [21, 22], o el uso de un proceso de evolución adaptativa para seleccionar mutantes supresoras PTS⁻ que restauren el fenotipo y crezcan más rápidamente en Glc [13, 23]. Adicionalmente, se ha alcanzado mayor rendimiento de compuestos aromáticos al promover el reciclaje de PEP mediante la sobreexpresión de las enzimas PEP sintetasa [11, 24, 25] o PEP carboxicinasa [26, 27].

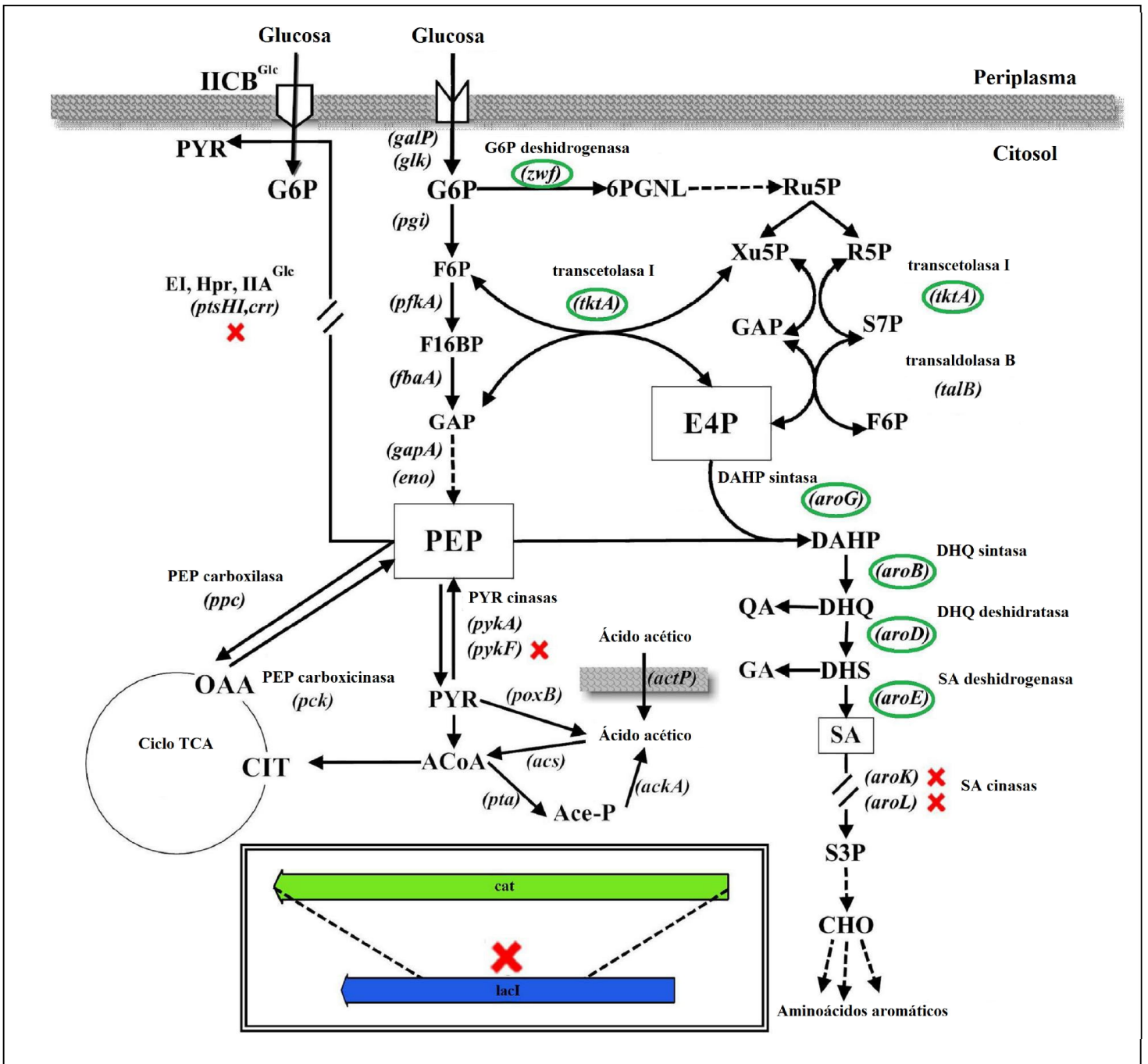


Figura 1. Modificaciones genéticas empleadas en este trabajo para promover la producción de shikimato a partir de glucosa en *E. coli*. Los genes inactivados están indicados con una cruz y los genes expresados en plásmido están circulados. Las flechas discontinuas indican más de un paso catalítico. G6P = glucosa 6-fosfato; F6P = fructosa 6-fosfato; F16BP = fructosa 1,6-bisfosfato; GAP = gliceraldehído 3-fosfato; 6PGNL = 6-fosfogluconolactona; Ru5P = ribulosa 5-fosfato; R5P = ribosa 5-fosfato; Xu5P = xilulosa 5-fosfato; S7P = sedoheptulosa 7-fosfato; E4P = eritrosa 4-fosfato; PEP = fosfoenolpiruvato; PYR = piruvato; ACoA = acetil-coenzima A; Ace-P = acetil fosfato; CIT = citrato; OAA = oxaloacetato; DAHP = 3-desoxi-D-arabinoheptuloseonato 7-fosfato; DHQ = 3-deshidroquinato; DHS = 3-deshidroschikimato; QA = ácido quínico; GA = ácido gálico; SA = shikimato; S3P = shikimato 3-fosfato; CHO = corimato; IICBGlc = componente de membrana de la permeasa PTS para glucosa; E1 = enzima 1 de PTS; Hpr = proteína de histidinas de PTS; IIAGlc = componente citosólico de la permeasa PTS para glucosa. Genes codificantes para enzimas no nombradas en la figura: *galP*, permeasa de galactosa; *glk*, glucocinasa; *pgi*, fosfogluconosa isomerasa; *pfkA*, 6-fosfofructocinasa I; *fbpA*, fructosa bisfosfato aldolasa; *gapA*, gliceraldehído 3-fosfato deshidrogenasa; *eno*, enolasa; *actP*, permeasa de acetato; *acs*, acetil-coenzima A sintetasa; *pta*, fosfato acetiltransferasa; *ackA*, acetato cinasa; *poxB*, piruvato oxidasa.

Por otro lado, la E4P es un metabolito que participa en reacciones reversibles presentes en la ruta no-oxidativa de la vía de las pentosas fosfato (PPP), así como sustrato de reacciones irreversibles que conducen a la producción de aminoácidos aromáticos o vitamina B₆ [28]. Se ha reportado que un incremento considerable en disponibilidad intracelular de E4P (inferido por el incremento en la producción de compuestos aromáticos) puede ser alcanzado mediante la sobreexpresión de genes que codifican para una transcetolasa (*tktA*) [24, 29, 30] o una transaldolasa (*talB*) [19, 31] (Fig. 1). Otras técnicas empleadas para incrementar el flujo de carbono hacia PPP y la producción de SA, incluyen el uso de cepas mutantes en la enzima fosfogluco isomerasa (codificada por *pgi*) [32], la sobreexpresión de la enzima glucosa 6-fosfato deshidrogenasa (codificada por *zwf*) [33], o el uso de fuentes de carbono múltiples, principalmente hexosas, pentosas y glicerol [34, 35].

Durante la última década se han caracterizado los efectos fisiológicos ocasionados por el empleo de las estrategias anteriores y se han reportado mejoras importantes en las capacidades de producción de SA y otros intermediarios aromáticos [6, 9, 12, 36]. No obstante, cabe resaltar que la mayoría de los trabajos reportados emplean sistemas de expresión génica controlados por mezclas de promotores nativos e inducibles, presentes en más de un tipo de plásmido. Estos factores imponen efectos adversos en los cultivos, como heterogeneidad en la intensidad y temporalidad de la expresión e incompatibilidad y segregación desigual de plásmidos, causando disminuciones en el rendimiento, productividad y crecimiento de las cepas [37–39]. La estabilidad y la calidad del DNA plasmídico se ven comprometidas cuando la maquinaria celular es incapaz de mantener un metabolismo que permita balancear la replicación, transcripción y traducción de los genes contenidos en los plásmidos con el resto de procesos celulares [40]. Además, los efectos adversos aumentan proporcionalmente con el tamaño y número de copias del plásmido y con la cantidad de proteínas recombinantes producidas [41, 42]. Algunos ejemplos de sistemas de expresión que han generado repercusiones favorables sobre el crecimiento y la producción de compuestos de interés incluyen el uso de plásmidos con bajo número de copias [37, 43], la integración y expresión cromosomal [44, 45], el incremento progresivo en el número de copias de un gen [46, 47] y el uso de bibliotecas de promotores [48].

Con respecto a la síntesis de SA, se ha reportado que una cepa silvestre de *E. coli* creciendo en medio mínimo con Glc destina únicamente 3 % de la poza interna de PEP hacia la

síntesis de AAA, como consecuencia de una gran regulación ejercida sobre esta vía mediante diferentes mecanismos de control [49, 50]. Por lo tanto, el empleo de modificaciones genéticas que incrementen artificialmente el flujo de carbono hacia la síntesis de SA y compuestos derivados podría afectar sustancialmente otros procesos celulares, dificultando el aislamiento de variantes con altos rendimientos. Esta consideración coincide con los relativamente bajos rendimientos de SA obtenidos hasta ahora usando cepas modificadas (10-30 %), que al ser comparados con el rendimiento máximo teórico calculado (86 %) sugiere que todavía hay mucho espacio para mejorar la distribución de flujos metabólicos de los organismos actualmente empleados [5]. En este contexto, debe tomarse en cuenta que una cepa que produzca SA con un rendimiento muy cercano al teórico podría estar muy limitada en sus capacidades de formación de biomasa [24], lo cual probablemente sería de poca utilidad para un proceso de producción. Una buena cepa productora será aquella que posea un fondo genético que le permita mantener un balance apropiado entre el crecimiento y el rendimiento de SA y que pueda prescindir de antibióticos, inductores o sustratos costosos sin comprometer la productividad del proceso. El presente trabajo tiene la finalidad de contribuir a la identificación de condiciones favorables para la producción de SA en *E. coli*, tomando en cuenta las particularidades que se describen a continuación.

ANTECEDENTES

Con el objetivo de producir aminoácidos aromáticos y compuestos derivados, en el laboratorio del Dr. Francisco Bolívar se han construido y caracterizado cepas de *E. coli* carentes de PTS [17, 18]. Una de esas cepas es PB11, la cual crece lentamente en Glc debido a la inactivación de este sistema. La cepa PB12, una derivada de PB11 con una velocidad de crecimiento incrementada 4 veces, fue aislada en un experimento de evolución adaptativa en laboratorio para promover la generación de mutantes con rápido crecimiento en Glc [51]. Se encontró que esta cepa internaliza Glc por el simportador GalP y puede canalizar el PEP no utilizado por PTS hacia la ruta de AAA (Fig. 1) [30, 52]. PB12 también puede utilizar Glc simultáneamente con otras fuentes de carbono (acetato, glicerol y varios carbohidratos) en medio mínimo debido a la ausencia de represión catabólica por PTS [53].

La secuenciación del genoma de la cepa PB12 sugirió que la eliminación de 12 genes, incluyendo a *rppH*, *galR* y *mutH*, es la principal causa de la recuperación del crecimiento rápido en Glc [54].

Se ha reportado que PB12 puede ser modificada para acumular SA en cultivos con Glc y extracto de levadura (YE) al portar dos plásmidos con los genes *tktA*, *aroB* y *aroE* controlados por sus propios promotores y el gen *aroG^{fbr}* bajo el promotor inducible lacUV5 [6]. Al evaluar la producción empleando cultivos en lote con concentraciones de sustratos de 25 g/L de Glc y 15 g/L de YE se observó que el título más alto de SA (7.1 g/L) y el rendimiento más alto de SA a partir de Glc (0.29 molSA/molGlc) fueron alcanzados por la cepa PB12 *aroK⁻ aroL⁻* (Fig. 2). Sin embargo, la cepa PB12 *aroK⁻ aroL⁻ pykF⁻* presentó el rendimiento más alto de compuestos aromáticos totales (0.50 molTAC/molGlc) como resultado de una mayor acumulación de subproductos intermediarios de la vía (Fig. 2). Lo anterior muestra que a pesar de que la eliminación del gen *pykF* incrementa el flujo de carbono hacia la vía biosintética de AAA, éste se está desviando a otros subproductos. Estos resultados concuerdan con otros estudios en los cuales se ha observado que la inactivación del gen *pykF* aumenta la concentración intracelular de enzimas de la ruta biosintética de aminoácidos aromáticos [55] y la producción de compuestos pertenecientes a la misma [56, 57]. Por otro lado, es evidente que tanto los vectores utilizados (2 plásmidos de diferente número de copias, portando cada uno 2 genes de la vía) como el sistema de expresión (uso de tres promotores con fuerzas no determinadas y un promotor inducible) pueden optimizarse con la finalidad de obtener una expresión más controlada y predecible de las enzimas requeridas y una canalización de los precursores hacia SA. Además, las cepas PTS⁻ *pykF⁻* que han sido empleadas en ese y otros trabajos para sobreproducir SA y compuestos derivados exhiben una menor velocidad de consumo de Glc y una menor generación de biomasa que sus contrapartes PTS⁻ *pykF⁺*, lo que ha limitado su aplicación en cultivos donde es esencial obtener altas productividades.

Este efecto ha sido parcialmente adjudicado a la acumulación de intermediarios glucolíticos en variantes *pykF⁻*, especialmente PEP, disminuyendo el flujo a través de la vía glucolítica y por ende las capacidades de formación de precursores esenciales para el crecimiento de la cepa [58]. Se ha determinado que el PEP puede inhibir alostéricamente a la enzima fosfofructocinasa I, PfkA [59], lo que ocasiona una disminución en el consumo de Glc al

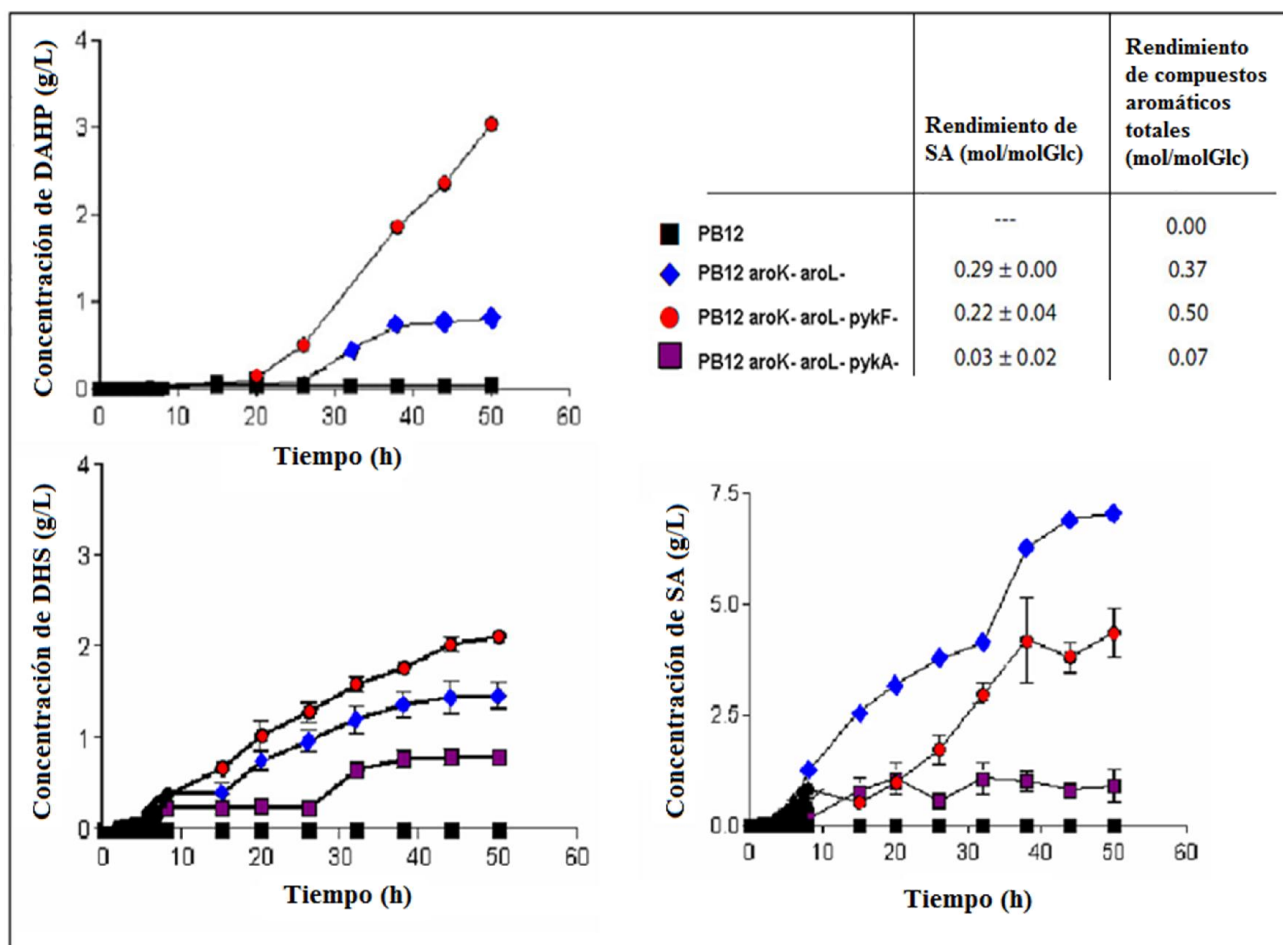


Figura 2. Concentraciones de SA y otros dos compuestos intermediarios de su vía biosintética, DAHP y DHS, obtenidas en PB12 y sus derivadas al ser transformadas con los plásmidos *pJLBaroG^{br}tklA* y *pTOPOaroBaroE*. Se muestra también el rendimiento de producción de shikimato y de compuestos aromáticos totales a partir de Glc para cada una de las cepas. Figura modificada de Escalante et al. 2010 [6].

ser esta enzima la responsable de catalizar la reacción que es considerada como el principal punto de control de la glucólisis [60]. Lo anterior sugiere que una estrategia factible para mejorar las capacidades de consumo de Glc y producción de SA y compuestos derivados en cepas *pykF* es evitar la acumulación de PEP y otros intermediarios glucolíticos, por ejemplo sobreexpresando de manera temprana enzimas que redirijan estos precursores hacia la vía de AAA.

En ese sentido, se debe hacer una consideración adicional, pues las cepas *aroK* *aroL* que son usadas para acumular eficientemente SA no pueden producir corismato (CHO), el

último intermediario común en las vías de síntesis de AAA y precursor de aminoácidos y vitaminas aromáticas esenciales para el crecimiento (Fig. 1). Por ello, es necesario suplementar estos componentes al medio de cultivo y se ha determinado que una opción viable para contrarrestar la auxotrofia de AAA es el uso de YE [6, 10]. El YE contiene predominantemente aminoácidos y carbohidratos con diferentes puntos de entrada al metabolismo celular, por lo que la presencia simultánea de Glc y YE seguramente propiciará un mejor crecimiento y generación de biomasa.

Con base en las consideraciones anteriores, en este trabajo se propone que una cepa PTS⁻ *pykF*⁻ tiene el potencial de incrementar las capacidades de producción de SA al compararse contra una cepa isogénica *pykF*⁺, siempre que el sistema de expresión permita una coordinación temporal adecuada en la síntesis de las enzimas requeridas. Esta situación promovería un consumo constante de Glc y un redireccionamiento eficiente de precursores glucolíticos hacia la síntesis de SA desde el inicio de la fermentación. Además de proveer los AAA, la suplementación de YE al medio de cultivo puede aminorar el desbalance de precursores biosintéticos celulares causado por la presencia de un plásmido y la sobreexpresión de proteínas de interés codificadas en el mismo [41]. De este modo, una mayor proporción de Glc podría redirigirse hacia la síntesis de SA sin afectar de manera tan severa la formación de biomasa, comparado con lo que se esperaría al emplear un medio más simple. Aunado a esto, si se permite que exista una expresión simultánea de las enzimas involucradas en la biosíntesis de SA podrían disminuirse los cuellos de botella de la vía, lo que puede incrementar el título y rendimiento de este compuesto y reducir la acumulación de subproductos.

Para evaluar esta hipótesis, se decidió construir un operón sintético con las secuencias codificantes de seis genes seleccionados de las vías de AAA y de PPP, controlados por un mismo promotor constitutivo e insertado en un plásmido de alto número de copias con estabilidad segregacional. El plásmido resultante se transformó en una cepa PB12 portadora de inactivaciones en los genes *aroK*, *aroL*, *pykF* y *lacI* (Fig. 1 y Tabla 1), la cual fue inicialmente evaluada en cultivos en modo lote con altas concentraciones de Glc y YE. En términos generales, la estrategia propuesta permitió la sobreproducción de SA desde el inicio de las fermentaciones, resultando en una alta concentración y rendimiento de SA con relativamente baja acumulación de subproductos. También se efectuó una caracterización

cinética, metabolómica y transcripcional de la mejor cepa de producción para elucidar algunos efectos fisiológicos ocasionados por la expresión constitutiva del operón sintético en este fondo genético y sistema de cultivo. Los resultados obtenidos permitieron identificar nuevos blancos de modificación y proponer estrategias para el mejoramiento metabólico de la cepa de producción.

HIPÓTESIS

La expresión constitutiva y simultánea de seis enzimas pertenecientes a las vías de las pentosas y de aminoácidos aromáticos incrementará el rendimiento de SA y la capacidad de consumo de Glc en una cepa con un fondo genético *pykF⁻*, comparados contra el fondo *pykF⁺*, mediante una canalización temprana y constante del flujo de carbono para evitar la acumulación de intermediarios metabólicos y la excreción de subproductos.

OBJETIVO GENERAL

Generar cepas de *E. coli* sobreproductoras de SA mediante la incorporación de modificaciones cromosomales y la construcción de un plásmido que permita expresar constitutivamente un operón sintético, así como realizar una caracterización de los perfiles cinéticos, metabolómicos y transcripcionales de la mejor cepa resultante para evaluar algunas implicaciones del uso este sistema de expresión.

OBJETIVOS PARTICULARES

1. Construcción del vector y modificación cromosomal de la cepa PB12 para promover la acumulación de SA: se empleará el plásmido pTrc327par para expresar a los genes *aroB*, *tktA*, *aroG^{fbr}*, *aroE*, *aroD* y *zwf*, ensamblados en forma de operón, en una derivada de PB12 con inactivaciones en los genes *aroK*, *aroL*, *pykF* y *lacI*.
2. Evaluación del crecimiento, consumo de Glc y producción de metabolitos de las cepas generadas: mediante cultivos en matraces y fermentadores se determinarán las diferencias en rendimiento, productividad y título de SA, consumo de Glc,

crecimiento y producción de acetato e intermediarios aromáticos, resultantes de la inactivación del gen *pykF* y del uso de diferentes concentraciones de sustratos en modo lote.

3. Determinación de los efectos causados por el sistema de expresión sobre el metabolismo central mediante un análisis transcriptómico y metabolómico de la mejor cepa productora: se obtendrán los valores de expresión relativa de genes relevantes por RT-qPCR y los cambios en las concentraciones relativas de metabolitos intracelulares por LC-MS.

RESULTADOS Y DISCUSIÓN

1. Construcción del vector y modificación cromosomal de la cepa PB12 para promover la acumulación de SA

1.1. Construcción del plásmido pTrc327par para ser usado como vector de expresión de los genes de interés

Considerando que las cepas sobreproductoras de SA que se han construido hasta el momento están todavía lejos de los parámetros de producción deseables [5, 6, 10, 61], la primera estrategia de este trabajo consistió en diseñar un vector apropiado para la sobreexpresión de los genes necesarios para acumular este metabolito.

En primer lugar, se generó el plásmido pTrc327par, que posee las siguientes características: a) el promotor Trc, el sitio múltiple de clonación (MCS) y terminadores de la transcripción del plásmido pTrc99A [62]; y b) el origen de replicación y alta estabilidad del pBR327par [63]. Esta construcción resultó en un vector de 4.4 Kb (Fig. 3) que, al igual que su parental, debería estar presente en 50 copias por célula [63, 64]. Además, el plásmido resultante pTrc327par contiene la región *par*, que confiere estabilidad segregacional mediante un sistema que posiciona el plásmido en la membrana, lo agrega con proteínas unidas al centrómero para formar un segregosoma y finalmente lo distribuye a las células hijas durante la formación del septo en el proceso de división celular [65, 66]. La funcionalidad

de esta región ha sido verificada por presencia del plásmido en más del 99% de las colonias formadas, después de realizar cultivos en medio LB por 100 generaciones sin antibiótico [63].

Por otro lado, la proteína codificada por el gen *lacI* (LacI) se une con alta afinidad a la secuencia de DNA correspondiente al operador *lac* en ausencia de lactosa, inhibiendo la transcripción de los genes posicionados río abajo [67]. Una herramienta vital para la ingeniería genética ha sido la generación de promotores híbridos con mayores fuerzas de expresión que incluyan la secuencia del operador *lac*, en conjunto con la presencia celular de una o más copias del gen *lacI* y el uso del inductor no-metabolizable isopropil- β -D-1-thiogalactopiranosido (IPTG) [68]. En cepas poseedoras del gen *lacI* en el cromosoma, el promotor híbrido Trc presente en el pTrc327par es inducible por IPTG, lo que le permite ser usado en un sistema de dos fases temporalmente separadas: una fase de proliferación celular sin inducción, y una fase de producción. Esto permite maximizar la generación de biomasa antes de inducir la sobreproducción de proteínas, estrategia empleada para reducir la competencia de precursores necesarios para la producción y el crecimiento [48, 62].

Sin embargo, experimentos anteriores realizados en nuestro laboratorio que emplean a la cepa PB12 para la producción de SA y compuestos aromáticos, han mostrado que es benéfico agregar al inicio del cultivo el inductor de la transcripción de genes requeridos inicialmente [6, 53, 57]. Esto se ha observado también para cepas PTS⁻ con otro fondo genético, donde la mejor producción de SA se alcanza con un cultivo en una sola fase [10]. Por esta razón, se decidió interrumpir la secuencia codificante del represor LacI en el cromosoma, de manera que el promotor Trc pueda expresarse de manera constitutiva. A diferencia de otros vectores de expresión que emplean promotores derivados de Lac como pTrc99A, el plásmido construido pTrc327par no incluye la secuencia codificante de LacI, volviendo constitutiva la expresión de genes clonados en este plásmido al ser transformado en cepas *lacI*⁻. Por otro lado, es probable que la presencia de un gen clonado en el MCS del pTrc327par pueda generar un nivel de expresión suficiente para otras aplicaciones al introducirlo en cepas que, a diferencia de PB12, no posean la variante sobreproductora *lacI*⁺. Esto responde a que el alto número de copias del vector puede ocasionar una expresión de escape debido a la baja titulación del represor sintetizado a partir de una sola copia cromosomal del alelo silvestre *lacI*. De cualquier manera, la eliminación del

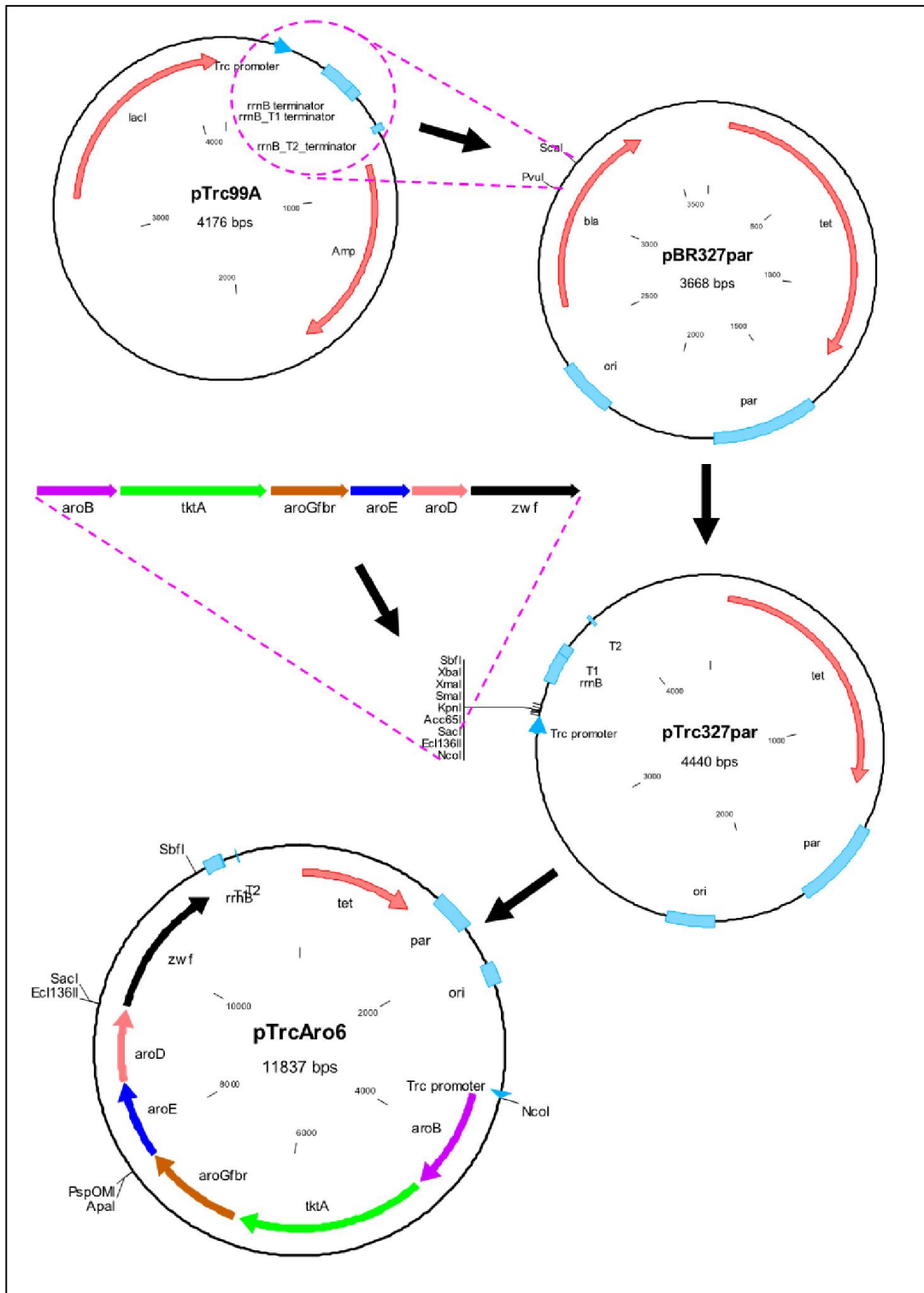


Figura 3. Esquema simplificado de los pasos requeridos en la construcción del plásmido pTrcAro6, que contiene 6 genes sintéticos bajo el control del promotor Trc. Primero, un segmento del plásmido pTrc99A fue amplificado y ligado en pBR327par, creando el plásmido pTrc327par. Un operón sintético con los genes requeridos (*aroB*, *tktA*, *aroG^{fbr}*, *aroE*, *aroD* y *zwf*) fue ensamblado por separado y transferido al pTrc327par, generando el plásmido pTrcAro6. Las líneas punteadas indican el sitio y orientación de algunas de las reacciones de ligación efectuadas. Sólo se muestran algunos sitios de restricción relevantes. Un esquema detallado de las construcciones efectuadas se muestra en Material adicional 2.

requerimiento de IPTG, además de evitar problemas de heterogeneidad en la inducción, también es ventajosa desde el punto de vista económico, pues el alto costo del IPTG impide su uso en fermentaciones de escala industrial.

1.2. Ensamblaje de un operón sintético con los genes necesarios para favorecer la sobreproducción de SA bajo el control de un mismo promotor

Se construyó un operón sintético con 6 genes requeridos para la producción de SA mediante la clonación en serie de sus secuencias codificantes e insertando sitios de unión ribosomal (RBS) con secuencia consenso (AGGAGG) y espaciamiento óptimo (8 bp) al sitio de inicio de la traducción de cada uno [7]. Los genes fueron amplificados por PCR empleando oligonucleótidos que generaron sitios únicos de restricción en los extremos del amplicón para facilitar la inserción en el MCS del plásmido pBRINT-Ts Cm [69], usado como plataforma de clonación (ver Materiales y métodos). El ordenamiento de los genes dentro del operón fue designado por la facilidad en la clonación. Este operón de 6 genes, denominado Aro6 (aproximadamente 8 Kb) fue subclonado en el plásmido pTrc327par, donde quedó flanqueado por el promotor Trc y los terminadores de la transcripción para generar el pTrcAro6 (Fig. 3). El operón sintético fue secuenciado para verificar la integridad del fragmento resultante (Material adicional 1).

La selección de los genes que constituyen el operón se efectuó de la siguiente manera: se decidió incluir a los genes *aroG^{fbr}* y *tktA*, porque se ha demostrado que la sobreexpresión de ambos es necesaria para canalizar el flujo de carbono hacia la vía de AAA [38]; se sabe que la reacción catalizada por la enzima codificada por *aroB* es un paso limitante del flujo en la vía [70]; la reacción catalizada por AroE es inhibida por su producto [71]; en condiciones de sobreexpresión de los genes anteriores se ha sugerido que *aroD* puede ser limitante cuando sólo hay una copia presente en el cromosoma [72]; y aumentar la expresión de la primera enzima de la vía de las pentosas, *Zwf*, podría incrementar la disponibilidad de NADPH (requerido por AroE como cofactor) y también redirigir una mayor parte de la glucosa hacia el precursor E4P por la vía oxidativa de PPP [73].

Algunas combinaciones de los 6 genes considerados aquí ya han sido ensambladas y expresadas bajo diferentes esquemas de control en otros trabajos con la finalidad de

aumentar la producción de compuestos aromáticos, pero este es el primer reporte de la expresión combinada de todos ellos. Además, la contribución específica de *aroD* y *zwf* a la producción de SA todavía no ha sido determinada [5, 70]. No obstante, en el presente trabajo no se buscó caracterizar la contribución de cada uno de los genes involucrados sino evaluar los efectos producidos por su expresión de manera conjunta en forma de operón.

1.3. Generación de cepas derivadas de PB12 que propicien un aumento en el flujo de carbono hacia SA y su acumulación en el medio extracelular

El tercer paso en la construcción de cepas sobreproductoras de SA contempló la modificación cromosomal de la cepa base PB12 mediante inactivaciones en los genes codificantes para el represor de lactosa (*lacI*) y las enzimas shikimato cinasas (*aroK* y *aroL*) y piruvato cinasa I (*pykF*). La construcción de cada una de las mutantes se llevó a cabo mediante recombinación homóloga de productos de PCR [36], seleccionando integrantes con la incorporación pasajera de un gen que confiere resistencia a antibiótico para después comprobar las inactivaciones cromosomales por PCR (Materiales y métodos). Las derivadas fueron transformadas con los plásmidos pTrc327par y pTrcAro6 y sembradas en cajas con medio selectivo, a partir de donde se aisló, verificó y almacenó una colonia de cada una de las cepas. Los genotipos resultantes y los nombres asignados a las cepas se muestran en la Tabla 1. La comprobación de la presencia de los dos plásmidos transformados a la cepa AR3 se muestra en las Figuras 4 y 5.

Tomando como molde a los plásmidos extraídos de cultivos saturados con las cepas transformadas, se observaron los perfiles de digestión y de amplificación por PCR al cargar el DNA resultante en geles de agarosa. El plásmido pTrcAro6 mide aproximadamente 12 Kb y posee dos sitios de corte para *KpnI*, lo que genera una banda de 8 Kb y otra de 4 Kb. Por el contrario, el plásmido pTrc327par sólo tiene un sitio de corte para *KpnI*, lo que lo lineariza en un fragmento de 4.4 Kb (Fig. 4a). Asimismo, pueden observarse diferencias entre los amplicones obtenidos al efectuarse un PCR con oligonucleótidos que flanquean el MCS del pTrc327par, generando un producto de 625 bp con el plásmido vacío y un producto de 8 Kb cuando está clonado el operón sintético (Fig. 4b y Material adicional 2). Para confirmar la presencia del operón sintético en el fondo genético esperado y descartar

Tabla 1. Plásmidos y cepas de *Escherichia coli* empleados en este trabajo.

Cepas		
Nombre	Características	Referencias
JM101	<i>F' traD36 proA⁺ proB⁺ lacI^h lacZΔM15/supE thi Δ(lac-proAB) rpoS(33 am)</i>	Messing (1979)
PB11	JM101 $\Delta(ptsH, ptsI, crr)::kan$	Flores et al. (2005); Flores et al. (2007)
PB12	PB11, PTS-Glc ⁺ ; cepa evolucionada en laboratorio	Flores et al. (2005); Flores et al. (2007)
AR2	PB12 <i>lacI aroK⁻ aroL⁻</i>	Este trabajo
AR3	PB12 <i>lacI aroK⁻ aroL⁻ pykF⁻</i>	Este trabajo
AR26	AR2 + pTrcAro6 (<i>Trc/aroB⁺ tktA⁺ aroG^{br+} aroE⁺ aroD⁺ zwf⁺</i>)	Este trabajo
AR36	AR3 + pTrcAro6 (<i>Trc/aroB⁺ tktA⁺ aroG^{br+} aroE⁺ aroD⁺ zwf⁺</i>)	Este trabajo
AR2e	AR2 + pTrc327par (plásmido sin operón sintético)	Este trabajo
AR3e	AR3 + pTrc327par (plásmido sin operón sintético)	Este trabajo
Plásmidos		
Nombre	Características	Referencias
pKD3	Templado de PCR para la amplificación del gen de resistencia a cloranfenicol flanqueado por secuencias homólogas para la recombinación con el cromosoma	Datsenko y Wanner (2000)
pKD46	Plásmido de expresión de la recombinasa λ -Red con origen de replicación termosensible	Datsenko y Wanner (2000)
pCP20	Plásmido de expresión de la recombinasa FLP	Cherepanov y Wackernagel (1995)
pBR327par	Derivado de pBR322 con mayor número de copias y estabilidad segregacional	Zurita et al. (1984)
pTrc99A	Plásmido de expresión que contiene al gen <i>lacI</i> y un polilinker en frente del promotor Trc	Amann et al. (1988)
pTrc327par	Contiene el promotor, polilinker y terminadores de la transcripción del pTrc99A y las regiones <i>par</i> y <i>ori</i> del pBR327par	Este trabajo
pTrcAro6	pTrc327par conteniendo un operón de seis genes para promover la producción de shikimato	Este trabajo

contaminación durante el aislamiento de la cepa, se observaron por qPCR los perfiles de amplificación de los transcritos correspondientes a los seis genes presentes en el operón y tres de los genes inactivados en el cromosoma. Al realizar el experimento con una concentración fija de DNA molde, puede compararse el número de ciclos de PCR que son necesarios para alcanzar una concentración determinada de DNA (definida por un umbral de fluorescencia) y con ello inferir el número de copias de cada gen. Para determinar la

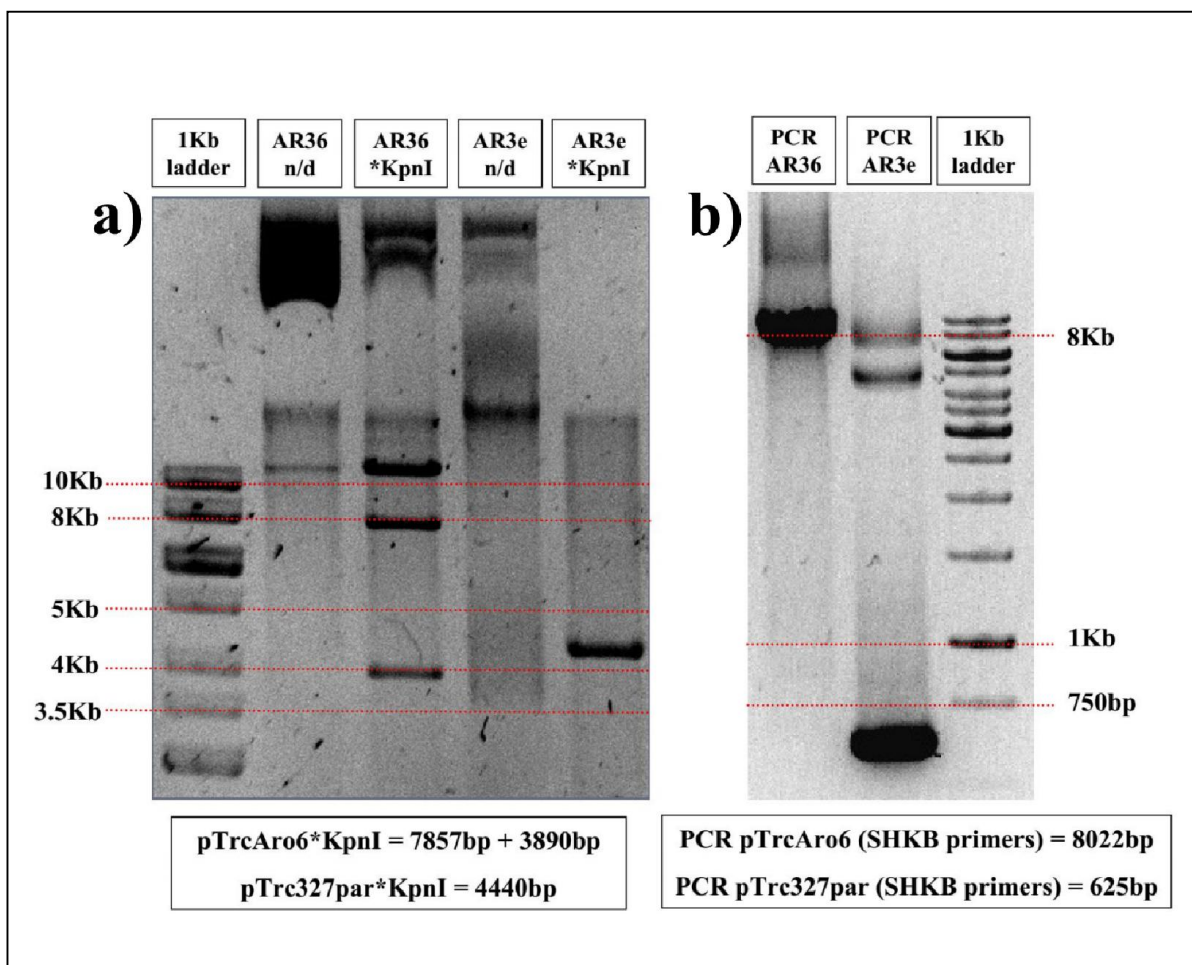


Figura 4. Comprobación de la cepa AR3 transformada con el plásmido pTrc327par (AR3e) y con el plásmido conteniendo el operón sintético Aro6 (AR36) por visualización de DNA en geles de agarosa. **a)** Plásmidos extraídos y digeridos con *KpnI*; **b)** PCR de los plásmidos extraídos empleando oligonucleótidos SHKB que aparecen en sitios que flanquean el MCS del pTrc327par, generando un producto mayor cuando está clonado el operón sintético. n/d: no digerido.

dosis relativa de los seis genes incluidos en el operón (*aroB*, *tktA*, *aroG^{flbr}*, *aroE*, *aroD* y *zwf*) se usaron como molde los plásmidos extraídos de las cepas AR3e y AR36. Puede observarse que los genes del operón alcanzan el umbral de fluorescencia (1 unidad relativa) después de 10 ciclos de PCR con DNA extraído de la cepa AR36, mientras que se necesitan casi 20 ciclos para que el DNA extraído de la cepa AR3e alcance la misma concentración (Fig. 5a). El hecho de que se hayan observado también curvas de amplificación para los

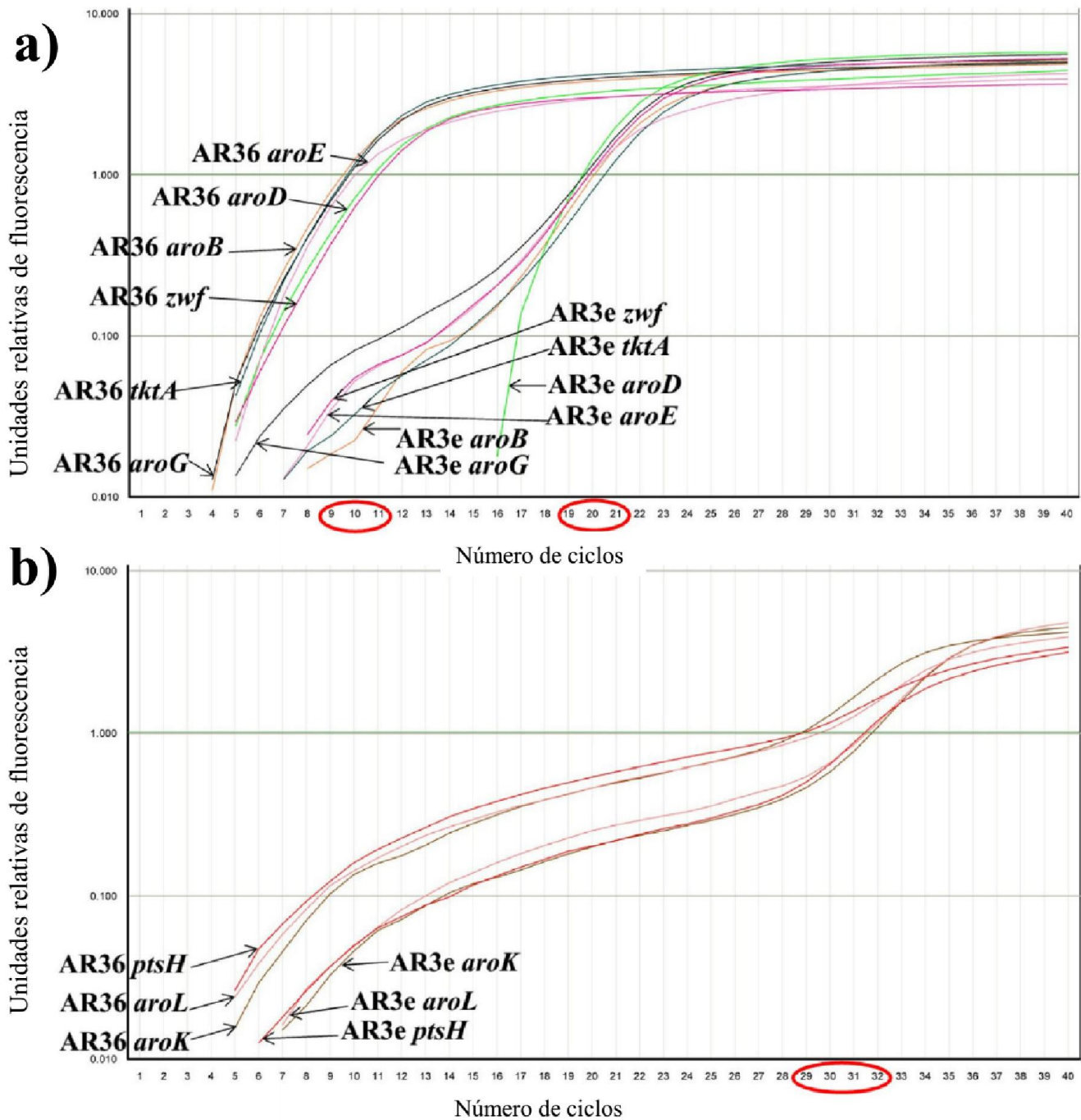


Figura 5. Curvas de amplificación obtenidas por qPCR usando oligonucleótidos específicos para cada uno de los genes evaluados. **a)** Amplificación de genes del operón sintético usando como molde a los plásmidos presentes en las cepas AR36 y AR3e; **b)** amplificación de genes interrumpidos en el cromosoma usando como molde al DNA cromosomal de ambas cepas. Los óvalos rojos indican el número aproximado de ciclos que fueron necesarios para alcanzar el umbral de fluorescencia.

genes del operón al usar plásmido extraído de AR3e (aunque en una cantidad mucho menor) se debe al inevitable aislamiento de pequeñas cantidades de DNA cromosomal al purificar plásmidos por lisis alcalina, lo cual es suficiente para funcionar como molde de PCR. Complementariamente, se llevaron a cabo pruebas de amplificación de tres genes interrumpidos usando como molde DNA cromosomal de ambas cepas. Como se esperaba para estos genes, ningún perfil de amplificación es positivo pues las curvas alcanzan el umbral solamente después de 30 ciclos y con una pendiente irregular (Fig. 5b), lo cual es un comportamiento típico de muestras sin DNA molde [52], confirmando las inactivaciones cromosomales.

2. Evaluación del crecimiento, consumo de Glc y producción de metabolitos de las cepas generadas

2.1. Determinación de los efectos causados por la inactivación conjunta de PTS y *pykF* en cepas que expresan el operón Aro6

Para evaluar los efectos causados por la inactivación de *pykF* sobre la producción de SA se comparó el desempeño de las cepas transformadas con el sistema de expresión, AR26 (*pykF*⁺) y AR36 (*pykF*⁻), usando matraces agitados con 15 g/L de Glc y 5 g/L de YE. Como control, también se incluyeron las derivadas que portan el plásmido pTrc327par (sin el operón Aro6), AR2e y AR3e (Tabla 1).

Aun cuando el SA se acumuló en todos los casos, como es esperado para mutantes en *aroK* y *aroL*, las cepas que contienen el pTrcAro6 (AR26 y AR36) alcanzaron concentraciones de SA más altas que las que contienen el plásmido vacío (Fig. 6b). Además, el título de SA fue casi dos veces mayor en AR36 que en AR26 (6.1 g/L contra 3.3 g/L). Se observó un decremento en el consumo de Glc en la cepa AR26 después de aproximadamente 18 h de cultivo, correlacionando con una alta concentración de acetato y un cese en la producción de SA. En contraste, la cepa AR36 exhibió un consumo de Glc constante y acumuló concentraciones muy bajas de acetato (Fig. 6c, 6d). Estos resultados sugieren que los genes presentes en el operón artificial son funcionales y promueven la producción de SA desde el inicio del cultivo. La probable expresión constitutiva de los seis genes del operón Aro6

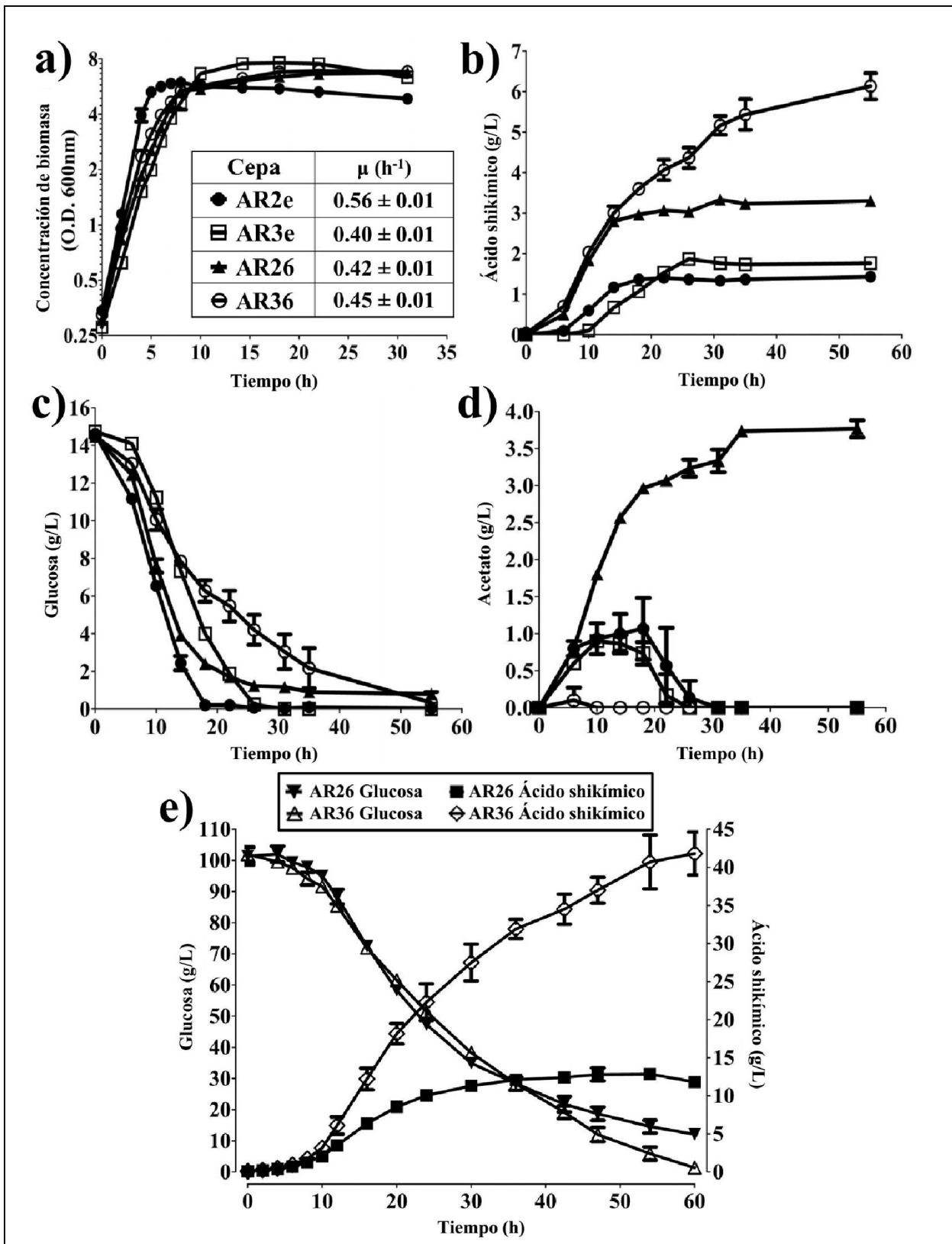


Figura 6. Comportamiento de las cepas AR26, AR36 y sus derivadas con plásmido vacío, AR2e (*pykF*⁺) y AR3e (*pykF*⁻), usando matraces con 15 g/L de Glc y 5 g/L de YE (a,b,c,d) y fermentadores de 1 L con 100 g/L de Glc y 15 g/L de YE (e). a) Crecimiento; b) producción de SA; c) consumo de Glc; d) producción de acetato; e) consumo de Glc y producción de SA de las cepas AR26 y AR36 en fermentadores.

disminuyó la velocidad de crecimiento (μ) en 25 % en el fondo *pykF*⁺, y la incrementó marginalmente en el fondo *pykF*⁻, pero no ocasionó cambios significativos a la biomasa máxima (X_{\max}) comparado con cepas con plásmido vacío (Fig. 6a). Importantemente, en estas condiciones de crecimiento la inactivación del gen *pykF* incrementó la producción de SA, eliminó la acumulación de acetato y permitió consumir completamente la Glc.

Para determinar si la mayor producción de acetato y menor producción de SA en AR26 comparada con AR36 es una consecuencia de la inherente baja disponibilidad de oxígeno y acidificación del medio de cultivo en matraces, ambas cepas fueron cultivadas en fermentadores de 1 L en modo lote bajo condiciones controladas de pH y tensión de oxígeno disuelto (DOT). Con el objetivo de incrementar el título de SA, la concentración inicial de Glc en estos experimentos se elevó a 100 g/L y la concentración de YE se incrementó a 15 g/L para permitir mayor formación de biomasa. Bajo estas condiciones, la cepa AR36 produjo 42 g/L de SA en 60 h, consumiendo toda la glucosa y acumulando 12 g/L de acetato. En contraste, después de 47 h la cepa AR26 produjo un máximo de 13 g/L de SA, no consumió toda la glucosa y acumuló 29 g/L de acetato (Fig. 6e y Tabla 2). Independientemente de las condiciones controladas en fermentadores, donde el pH se mantuvo en 7 y la DOT fue mayor al 20% en todo momento, los perfiles de producción de ambas cepas fueron parecidos al comportamiento observado en matraces, con AR26 produciendo más acetato y menos SA. Aun cuando la μ , X_{\max} y la velocidad volumétrica global de consumo de Glc ($Q_{S_{\text{global}}}$) fueron similares en ambas cepas, la productividad, el rendimiento y el título fueron más de dos veces mayores en AR36 que en AR26 (Tabla 2). Es sorprendente que las diferencias tan grandes observadas en la producción de SA y acetato se deban a la inactivación de un solo gen, lo que muestra las ventajas de la inactivación en el gen *pykF* sobre la producción de SA al usar un sistema de expresión constitutivo. Es importante considerar que es probable que si se usaran otros sistemas de expresión que no produjeran las enzimas que consumen intermediarios glucolíticos desde el inóculo, los efectos positivos de la inactivación en *pykF* podrían verse enmascarados por desventajas impuestas por una excesiva acumulación de PEP, tales como un flujo glucolítico reducido. Sin embargo, en ausencia de experimentos complementarios la consideración anterior es únicamente una hipótesis.

Tabla 2. Datos comparativos de las fermentaciones con las cepas AR26 y AR36, usando 100 g/L de Glc y 15 g/L de YE como sustratos.

Cepa	AR26	AR36
Título de SA (g/L)	12.95 ± 0.64	41.80 ± 2.83
Glc consumida (g/L)	82.65 ± 4.88	103.70 ± 6.79
Duración del cultivo (h)	47	60
Y_{SA/Glc} (mol/mol)	0.161 ± 0.020	0.417 ± 0.001
Título de acetato (g/L)	29.35 ± 0.21	11.90 ± 0.14
X_{max} (g/L)	6.18 ± 0.10	6.54 ± 0.09
μ (h⁻¹)	0.45 ± 0.01	0.45 ± 0.02
Q_pglobal (gSA/L*h)	0.27 ± 0.02	0.75 ± 0.07
Q_sglobal (gGlc/L*h)	-1.76 ± 0.10	-1.73 ± 0.11

Y_{SA/Glc} = rendimiento de SA a partir de Glc; X_{max} = biomasa máxima alcanzada expresada como peso seco; Q_pglobal = productividad volumétrica de SA para la fermentación completa; Q_sglobal = velocidad volumétrica de consumo de Glc para la fermentación completa; μ = velocidad específica de crecimiento.

2.2. Obtención de los perfiles de fermentación de la cepa AR36 con tres diferentes concentraciones de sustratos

Tomando en cuenta los resultados anteriores, se seleccionó a la cepa AR36 para efectuar una caracterización más a fondo de su comportamiento cinético y estequiométrico en fermentadores de 1 L. Para llevar a cabo este objetivo, se evaluó la producción de SA con tres diferentes concentraciones de sustratos. Para cada caso se determinó el crecimiento y las concentraciones de Glc y metabolitos extracelulares, lo que permitió efectuar una comparación de las productividades y rendimientos.

Primero, se utilizaron 50 g/L de Glc y 15 g/L de YE (Fig. 7a). La cepa presentó crecimiento durante las primeras 10 h, generando 6.3 g/L de peso seco (DCW) con una μ de 0.53 h⁻¹. Bajo esta condición, se produjeron 24 g/L de SA en 32 h. Se observó consumo de Glc y producción de SA desde el inicio de la fermentación y hasta finalizarse la Glc, a pesar de

que las velocidades específicas de producción de SA y de consumo de sustrato (q_p y q_s , respectivamente) fueron mayores en fase exponencial (Tabla 3). El rendimiento de SA a partir de Glc ($Y_{SA/Glc}$) fue de 0.47 mol/mol y la productividad volumétrica de SA para la fermentación completa ($Q_{p_{global}}$) fue de 0.74 gSA/L*h (Tabla 3). Con respecto a la acumulación de subproductos en la vía del SA, se detectaron concentraciones de 2.4 g/L de DAHP, 2.1 g/L de DHS, 1.4 g/L de QA, 0.4 g/L de GA y 0.3 g/L de DHQ al final de la fermentación (Fig. 8a). Bajo estas condiciones, la producción de acetato fue casi nula durante el curso de la fermentación, alcanzando una concentración máxima de 1.5 g/L después de 32 h (Fig. 7a).

Considerando que los 50 g/L de Glc fueron consumidos completamente, se inició un segundo experimento en lote con 100 g/L de Glc y 15 g/L de YE. Como se describió en la comparación con AR26 en la sección previa, AR36 crecida en estas condiciones produce aproximadamente 42 g/L de SA en 60 h (Fig. 7b). En este caso, después de consumir cerca de 100 g/L de Glc y alcanzar una concentración de 41.8 g/L de SA, la cepa produjo 12 g/L de acetato. Los valores obtenidos para $Y_{SA/Glc}$, $Q_{p_{global}}$, $Q_{s_{global}}$, X_{max} y μ , fueron similares a los obtenidos con 50 g/L de Glc y 15 g/L de YE (Tabla 3). Estos experimentos muestran que cuando se usa la misma concentración de YE, el doble de Glc es consumida en el doble de tiempo, indicando que la tasa promedio de consumo de Glc se mantiene entre ambas condiciones de cultivo. De manera semejante, la Q_p calculada es muy similar en ambos experimentos porque se produce aproximadamente el doble de SA en el doble de tiempo, es decir, el cultivo presenta la misma productividad volumétrica sin importar si se usaron 50 o 100 g/L de Glc iniciales (Tabla 3). Se detectaron concentraciones de 4.8 g/L de DAHP, 2.8 g/L de DHS, 3.4 g/L de QA, 0.7 g/L de GA y 0.9 g/L de DHQ en el sobrenadante después de 60 h (Fig. 8b). Interesantemente, cuando se duplica la concentración de Glc los productos intermedios de la vía de AAA incrementaron de manera casi proporcional con el SA, indicando que el consumo de 100 g/L de Glc aparentemente no generó nuevos cuellos de botella en la vía. Como resultado, la cantidad de SA formado con respecto a los compuestos aromáticos totales fue cercana al 80% para ambos experimentos (Fig. 9).

Se investigó el efecto de incrementar el YE sobre la productividad de SA con un tercer grupo de experimentos, usando 100 g/L de Glc y 30 g/L de YE. A pesar de que la biomasa

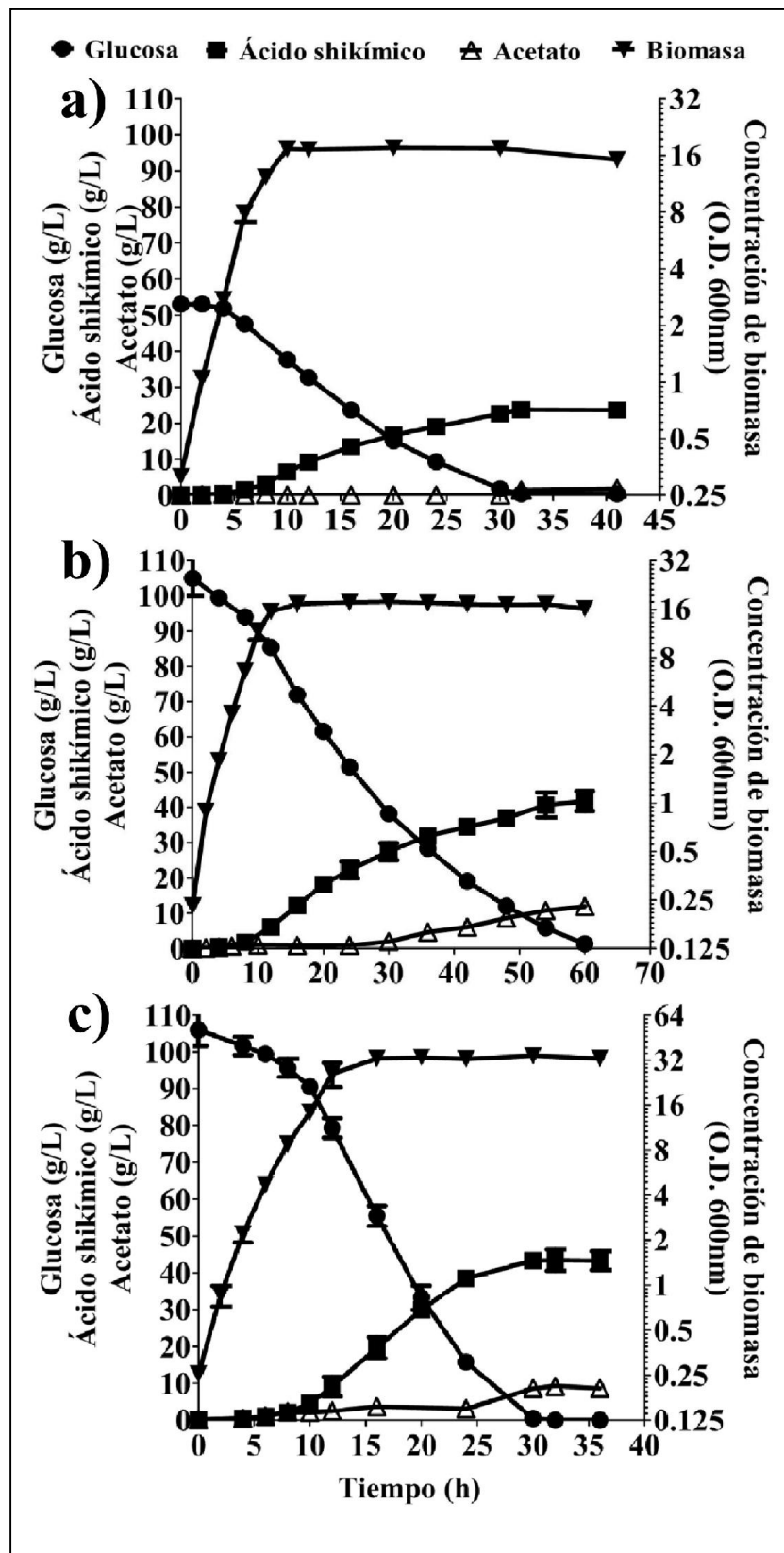


Figura 7. Perfiles de fermentación de la cepa AR36 cultivada en fermentadores de 1 L con tres diferentes concentraciones de sustratos. a) 50 g/L de Glc y 15 g/L de YE; b) 100 g/L de Glc y 15 g/L de YE; c) 100 g/L de Glc y 30 g/L de YE. Glc: círculos; SA: cuadros; acetato: triángulos abiertos; concentración de biomasa: triángulos invertidos.

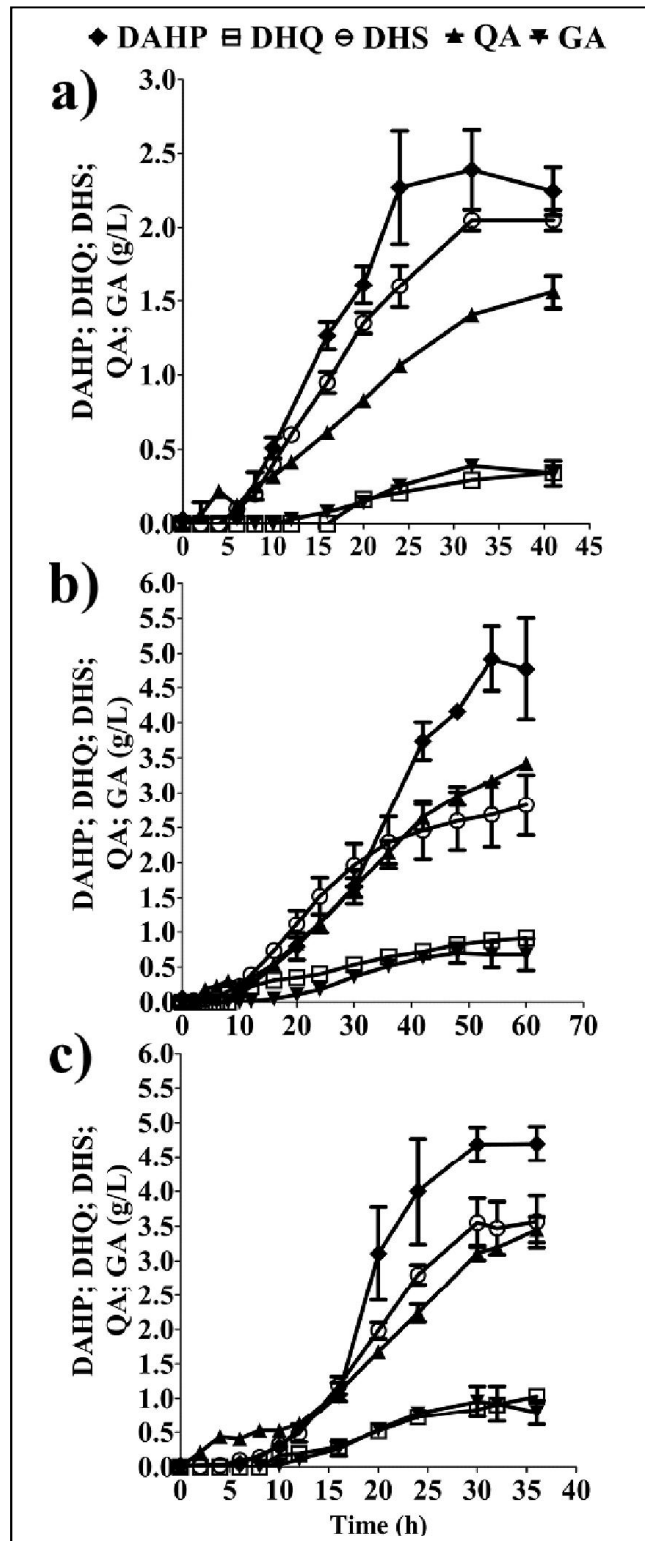


Figura 8. Subproductos aromáticos de la vía del SA detectados en las fermentaciones de 1 L con la cepa AR36 usando tres diferentes concentraciones de sustratos. **a)** 50g/L de Glc y 15g/L de YE; **b)** 100g/L de Glc y 15g/L de YE; **c)** 100g/L de Glc y 30g/L de YE. Diamantes: DAHP (3-desoxi-D-arabinoheptuloseonato 7-fosfato); cuadros: DHQ (ácido 3-deshidroquinico); círculos: DHS (ácido 3-deshidroshikimico); triángulos: QA (ácido quinico); triángulos invertidos: GA (ácido gálico).

se duplicó al usar el doble de YE, el título de SA, la μ y el $Y_{SA/Glc}$ fueron muy similares a los obtenidos en el cultivo con 100 g/L de Glc y 15 g/L de YE (Fig. 7b y 7c). En conjunto con datos obtenidos de las otras dos condiciones, estos hallazgos sugieren que la cantidad de YE principalmente determina la biomasa máxima que se puede alcanzar. Además, un incremento en la concentración inicial de YE no alteró el título de SA y fortalece la suposición de que el SA está siendo producido principalmente a partir de la Glc. La relación directa entre el YE y la biomasa máxima generada, sugiere que uno o más nutrientes limitantes están siendo proporcionados exclusivamente por el YE. Es de esperarse que los aminoácidos presentes en el YE, necesarios para contrarrestar la auxotrofia de AR36, se vuelvan limitantes; sin embargo, otros compuestos presentes en este medio complejo podrían también jugar un papel en la limitación del crecimiento.

Para una concentración inicial de 30 g/L de YE, 106 g/L de Glc totales fueron consumidos y 43 g/L de SA fueron producidos en aproximadamente la mitad del tiempo que la fermentación con 15 g/L de YE, duplicando la productividad volumétrica, aun cuando el título de SA no cambió (Tabla 3). Como la biomasa también se incrementó al doble, el consumo y productividad específica (q_s y q_p) fueron similares entre los tres experimentos, tanto en fase exponencial como estacionaria.

Además, los resultados muestran que un incremento en la concentración de YE no alteró considerablemente la concentración de intermediarios de la vía (Fig. 8c). Con respecto a este resultado, se ha reconocido que la presencia de altas cantidades de intermediarios de la vía de AAA tiene un efecto negativo sobre la purificación de SA del medio de cultivo. Esta situación ha propiciado esfuerzos de algunos grupos para evaluar condiciones de cultivo, fondos genéticos y el uso de análogos de glucosa no metabolizables, como intentos para minimizar la formación de subproductos [12].

En este trabajo se ha encontrado una alta proporción de SA relativo a subproductos sin aplicar ninguna otra modificación a la cepa o al proceso. La concentración de cada intermediario de la vía fue comparada contra la suma de todos los intermediarios aromáticos y sus porcentajes fueron usados para calcular la proporción molar de SA a cada subproducto al final de las fermentaciones (Fig. 9). La proporción de SA resultó ser mayor a 10 para DHS, QA o DAHP, y mayor a 40 para GA o DHQ en todas las concentraciones

Tabla 3. Comparación de metabolitos medidos y parámetros cinéticos y estequiométricos calculados entre tres fermentaciones de la cepa AR36 con diferentes concentraciones de sustratos.

Cepa	AR36	AR36	AR36
Condiciones de cultivo	Lote 50g/L Glc + 15g/L YE	Lote 100g/L Glc + 15g/L YE	Lote 100g/L Glc + 30g/L YE
Título de SA (g/L)	23.810 ± 0.014	41.800 ± 2.828	43.300 ± 0.566
Glc consumida (g/L)	52.650 ± 1.202	103.700 ± 6.788	105.550 ± 4.455
Duración del cultivo (h)	32	60	30
Y_{SA/Glc} (mol/mol)	0.468 ± 0.011	0.417 ± 0.001	0.424 ± 0.012
Título de acetato (g/L)	1.445 ± 0.007	11.900 ± 0.141	8.650 ± 0.919
X_{max} (g/L)	6.299 ± 0.086	6.543 ± 0.086	12.540 ± 0.056
μ (h⁻¹)	0.531 ± 0.028	0.448 ± 0.016	0.448 ± 0.002
Q_pglobal (gSA/L*h)	0.751 ± 0.004	0.754 ± 0.065	1.443 ± 0.019
Q_sglobal (gGlc/L*h)	-1.645 ± 0.038	-1.728 ± 0.113	-3.518 ± 0.148
q_pexp (gSA/gDCW*h)	0.381 ± 0.002	0.335 ± 0.026	0.461 ± 0.059
q_sexp (gGlc/gDCW*h)	-1.138 ± 0.153	-1.110 ± 0.005	-1.183 ± 0.030
q_psta (gSA/gDCW*h)	0.160 ± 0.006	0.193 ± 0.016	0.197 ± 0.013
q_ssta (gGlc/gDCW*h)	-0.359 ± 0.007	-0.420 ± 0.007	-0.416 ± 0.028

Y_{SA/Glc} = rendimiento de SA a partir de Glc; X_{max} = biomasa máxima alcanzada expresada como peso seco; Q_pglobal = productividad volumétrica de SA para la fermentación completa; Q_sglobal = velocidad volumétrica de consumo de Glc para la fermentación completa; q_pexp = productividad específica de SA en fase exponencial; q_sexp = velocidad específica de consumo de Glc en fase exponencial; q_psta = productividad específica de SA en fase estacionaria; q_ssta = velocidad específica de consumo de Glc en fase estacionaria; μ = velocidad específica de crecimiento.

de sustrato evaluadas. Importantemente, en todas las condiciones los rendimientos de SA estuvieron cerca del 50 % del máximo teórico y los rendimientos de compuestos aromáticos totales (TAC) fueron superiores al 60 % del máximo teórico (Fig. 9). Lo anterior refleja un redireccionamiento eficiente de Glc hacia la vía de AAA en la cepa AR36, incluso cuando

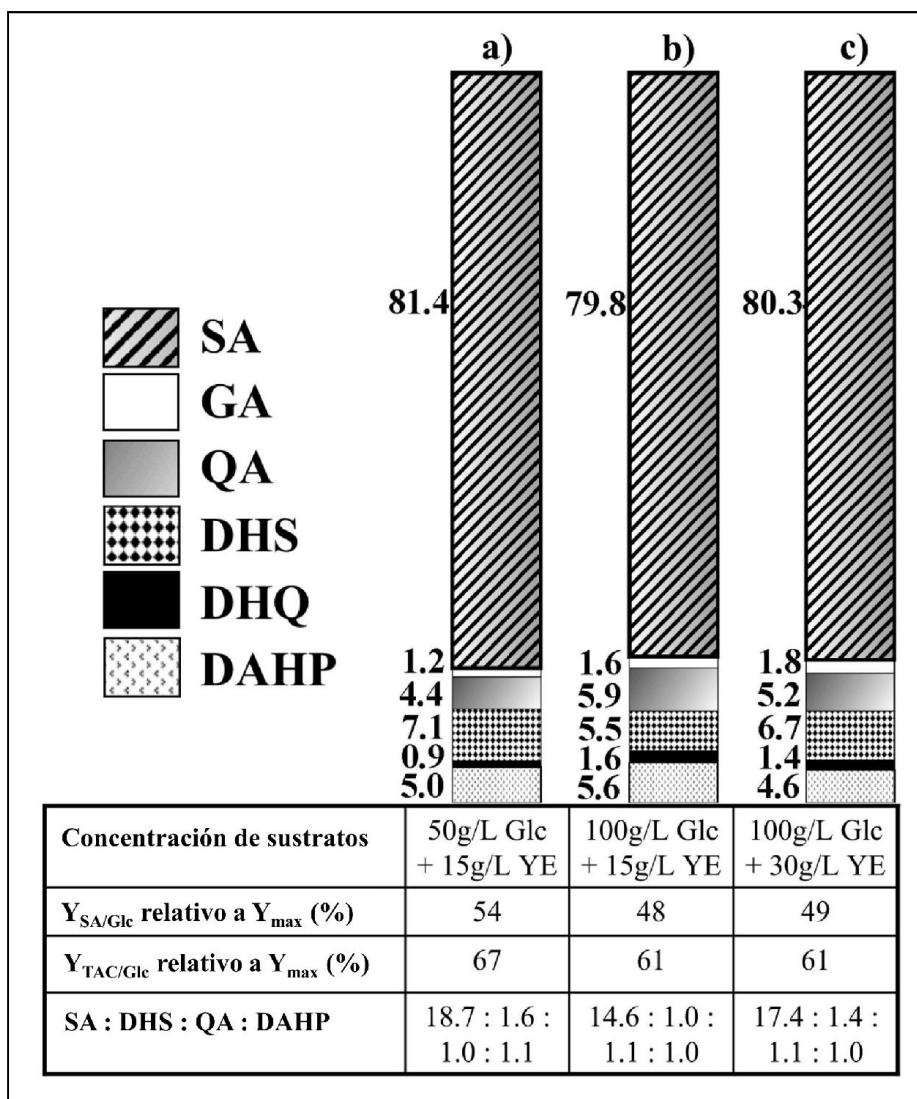


Figura 9. Porcentaje molar de cada compuesto aromático producido en la cepa AR36 con respecto al total en cultivos en lote con las siguientes concentraciones iniciales de sustratos: **a)** 50 g/L de Glc y 15 g/L de YE; **b)** 100 g/L de Glc y 15 g/L de YE; **c)** 100 g/L de Glc y 30 g/L de YE. Los rendimientos calculados y las proporciones molares de los compuestos aromáticos producidos se muestran debajo de cada barra. Las comparaciones se realizaron tomando en cuenta las concentraciones medidas en el sobrenadante al final de las fermentaciones. $Y_{SA/Glc}$ = rendimiento de SA a partir de Glc; $Y_{TAC/Glc}$ = rendimiento de compuestos aromáticos totales a partir de Glc; Y_{max} = rendimiento máximo teórico de compuestos aromáticos.

se utilizaron cultivos en lote con altas concentraciones de Glc. En este punto vale la pena mencionar que el título de SA más alto reportado en la literatura es de 84 g/L, el cual se ha alcanzado cultivando una cepa modificada de *E. coli* en modo lote alimentado [10]. Este modo de cultivo permite alcanzar altas densidades celulares como resultado del suministro de grandes cantidades de Glc, manteniendo una baja velocidad de adición y una baja concentración de Glc en el fermentador. Dicha estrategia es utilizada para controlar la q_s y

la μ de la cepa, evitando un excesivo flujo glucolítico y los problemas asociados al mismo, tales como bajos rendimientos producto/sustrato y acumulación de acetato [74]. Sin embargo, un cultivo en modo lote alimentado también puede representar un incremento en los costos de operación, comparado con un cultivo en lote, y una q_s más baja puede hacer más largo el tiempo de fermentación y disminuir la productividad [74, 75]. En el presente trabajo se muestra que la cepa AR36 puede ser cultivada en lote con altas concentraciones de Glc y YE y presentar parámetros de producción que compitan con los obtenidos por otras cepas y en otras condiciones de cultivo. Por ejemplo, el título de SA (43 g/L), aunque todavía lejano a los 84 g/L, es el más alto que se ha obtenido con un cultivo en lote y la productividad volumétrica de SA (1.44 g/L*h) es cercana a la máxima productividad reportada de 2.05 g/L*h (Tabla 4). Además, el rendimiento de SA fue constante en todas las condiciones evaluadas y representa el valor más alto para cualquier condición de cultivo (Tabla 4).

Por otro lado, no en todos los trabajos se reportan las mediciones del total de subproductos de la vía del SA, lo que dificulta hacer una comparación de la acumulación de estos compuestos entre las distintas cepas. No obstante, en algunos casos es posible determinar la proporción molar de SA al principal subproducto aromático obtenido, como una medida de la eficiencia en la canalización de los intermediarios hacia el producto final. De esta manera, puede observarse que la cepa AR36 ocupa el segundo lugar en acumulación de SA con respecto al principal subproducto, alcanzando 12 moles de SA por cada mol de DHS al final de la fermentación (Tabla 4). La proporción más alta de SA con respecto a un intermediario reportada en la literatura corresponde a un proceso en lote alimentado con una cepa llamada SA5 [76], que si bien presenta una productividad 6 veces menor que los cultivos con AR36, logra acumular 21 moles de SA por mol de DHQ (Tabla 4). Esta ventaja puede ser resultado de la inactivación del gen *ydiB* en la cepa SA5, el cual codifica para una enzima shikimato/quinato deshidrogenasa [77]. Se ha sugerido que YdiB puede contribuir a la producción de DHS, DHQ y QA y que su inactivación puede evitar la conversión de SA a estos intermediarios [72, 77]. Sin embargo, hasta la fecha no hay evidencia contundente de que la inactivación cromosomal de *ydiB* ocasione una disminución en la acumulación de intermediarios aromáticos, pues aparentemente el resultado depende del contexto evaluado [70, 76, 78, 79].

Tabla 4. Cepas de *E. coli* que han sido modificadas para favorecer la producción de SA bajo diferentes condiciones de cultivo.

Cepa	Características relevantes	(Título ^a , rendimiento ^b , productividad ^c). Condiciones de cultivo. [Proporción molar de SA al principal subproducto aromático]	Referencia
SP1.1pts/pSC6.090B (derivada de RB791)	$\Delta ptsHIcrr \Delta aroK \Delta aroL$ <i>serA::aroB</i> / (plásmido) <i>aroF^{fbr}</i> <i>tktA</i> , P _{tac} <i>aroE serA</i> , P _{tac} <i>glf^d glk^d</i>	(71, 0.27, 1.18). Fermentadores 1 L modo lote alimentado con Glc y AAA. [5:1] ^e	[10]
SP1.1pts/pSC6.090B (derivada de RB791)	$\Delta ptsHIcrr \Delta aroK \Delta aroL$ <i>serA::aroB</i> / (plásmido) <i>aroF^{fbr}</i> <i>tktA</i> , P _{tac} <i>aroE serA</i> , P _{tac} <i>glf^d glk^d</i>	(84, 0.33, 2.05). Fermentadores 10 L modo lote alimentado con Glc, AAA y 15 g/L de YE. [8:1] ^e	[10]
KPM1 SA1 (derivada de K-12)	$\Delta aroK \Delta aroL$ / (plásmido) <i>aroF^{fbr}</i> , P _R -P _L <i>aroE</i>	(7, 0.11, 0.20). Fermentadores 5 L modo lote con 60 g/L de Glc, AAA y 5 g/L de YE. [NR]	[32]
SA116 (derivada de BW25113)	$\Delta aroK \Delta aroL$ P _{pps::P_{lacQ1}} , P _{csrB::P_{lacQ1}} / (cromosoma) <i>aroG^{fbr} tktA aroB aroE</i> , P _{T5} <i>ppsA</i> <i>csrB</i> , 5P _{tac} <i>tktA nadK</i>	(3, 0.33, 0.04). Matraces modo lote con 10 g/L de Glc, 1 g/L de peptona y 1 g/L de prolina. [NR]	[47]
DHPYAAS-T7 (derivada de DH5 α)	$\Delta ptsHIcrr \Delta aroK \Delta aroL \Delta ydiB$ / (plásmido) P _{T7} <i>aroB aroE glk</i> <i>tktA aroF^{fbr}</i>	(1, 0.04 ^f , 0.02). Matraces modo lote con 25 g/L de glicerol, 10 g/L de peptona y 15 g/L de YE. [NR]	[79]
MG1655	(plásmido) P _{lac-UV5} <i>aroE aroD</i> <i>aroB^{op}</i> , P _{L-tetO1} <i>aroG^{fbr} ppsA tktA</i>	(0.8, 0.14, 0.03). Matraces modo lote con 5 g/L de Glc. [6:1] ^e	[70]
SA5 (derivada de B0013)	$\Delta ptsG \Delta aroK \Delta aroL \Delta ydiB$ $\Delta ackA-ptA$ / (plásmido) P _{lac} <i>aroG^{fbr} ppsA tktA</i>	(15, 0.29, 0.26). Fermentador 7 L modo lote alimentado con Glc y AAA. [21:1] ^g	[76]
PB12.SA22 (derivada de JM101)	$\Delta ptsHIcrr \Delta aroK \Delta aroL$ / (plásmido) P _{lacUV5} <i>aroG^{fbr}</i> , <i>tktA</i> , (plásmido) <i>aroE aroB</i>	(7, 0.29, 0.16). Fermentadores 1 L modo lote con 25 g/L de Glc y 15 g/L de YE. [5:1] ^e	[6]
PB12.SA31 (derivada de JM101)	$\Delta ptsHIcrr \Delta aroK \Delta aroL \Delta pykF$ / (plásmido) P _{lacUV5} <i>aroG^{fbr}</i> , <i>tktA</i> , (plásmido) <i>aroE aroB</i>	(4, 0.22, 0.11). Fermentadores 1 L modo lote con 25 g/L de Glc y 15 g/L de YE. [2:1] ^h	[6]
AR36 (derivada de JM101)	$\Delta ptsHIcrr \Delta aroK \Delta aroL \Delta lacI$ $\Delta pykF$ / (plásmido) P _{trc} <i>aroB</i> <i>tktA aroG^{fbr} aroE aroD zwf</i>	(43, 0.42, 1.44). Fermentadores 1 L modo lote con 100 g/L de Glc y 30 g/L de YE. [12:1] ^e	Este trabajo
AR36 (derivada de JM101)	$\Delta ptsHIcrr \Delta aroK \Delta aroL \Delta lacI$ $\Delta pykF$ / (plásmido) P _{trc} <i>aroB</i> <i>tktA aroG^{fbr} aroE aroD zwf</i>	(59, 0.42, 2.46). Fermentadores 2 L modo lote alimentado con Glc y 40 g/L de YE. [NR]	Este trabajo

^agSA/L; ^bmol SA/mol Glc; ^cgSA/(L*h); ^dgenes de *Zymomonas mobilis*; ^emol SA/mol DHS; ^fmol SA/mol glicerol; ^gmol SA/mol DHQ; ^hmol SA/mol DAHP. NR = no reportado.

De cualquier manera, los buenos parámetros de producción obtenidos con la cepa AR36 pudieran atribuirse a que la plataforma desarrollada en este trabajo dicta una expresión constitutiva y simultánea de las enzimas necesarias, en un fondo genético que puede propiciar un mejor redireccionamiento de carbono hacia SA. Esto contrasta con otros sistemas de expresión en los cuales los genes requeridos han sido expresados a partir de plásmidos separados, bajo diferentes promotores o en cepas no optimizadas para la asimilación eficiente de altos niveles de Glc. Además de las ventajas conferidas a la dinámica de la expresión génica, el hecho de que no se necesite IPTG para inducir al operón Aro6 representa un importante beneficio económico para el proceso de producción. Hasta el momento, AR36 es la única cepa productora de SA que no requiere de la adición de un inductor para sobreexpresar a los genes biosintéticos.

3. Determinación de los efectos causados por el sistema de expresión sobre el metabolismo central mediante un análisis transcriptómico y metabolómico de la mejor cepa productora

3.1. Perfiles transcripcionales de genes relevantes en la cepa AR36

Para obtener un mejor panorama de los cambios metabólicos ocasionados por la expresión constitutiva del operón sintético Aro6 en condiciones de producción, los niveles de transcrito de 40 genes codificantes para enzimas de metabolismo central en AR36 (Fig. 10) fueron comparados contra los de la cepa que contiene el plásmido vacío, AR3e. Simultáneamente, se decidió evaluar el cambio en los niveles de transcripción de los mismos 40 genes en la cepa productora a lo largo del cultivo. Específicamente, se compararon los niveles de transcrito de todos los genes evaluados, determinados a partir de muestras celulares provenientes de cultivos en fermentadores con 50 g/L de Glc y 15 g/L de YE y colectadas en fase exponencial temprana (EE), exponencial tardía (LE), y estacionaria (ST). Como control, todos los datos fueron normalizados a los valores obtenidos para la cepa AR3e, crecida en las mismas condiciones y de la cual se extrajo RNA en fase EE. De esta manera, con una sola muestra control puede analizarse de manera independiente el efecto de expresar constitutivamente el operón (comparando AR36 y AR3e) y la influencia

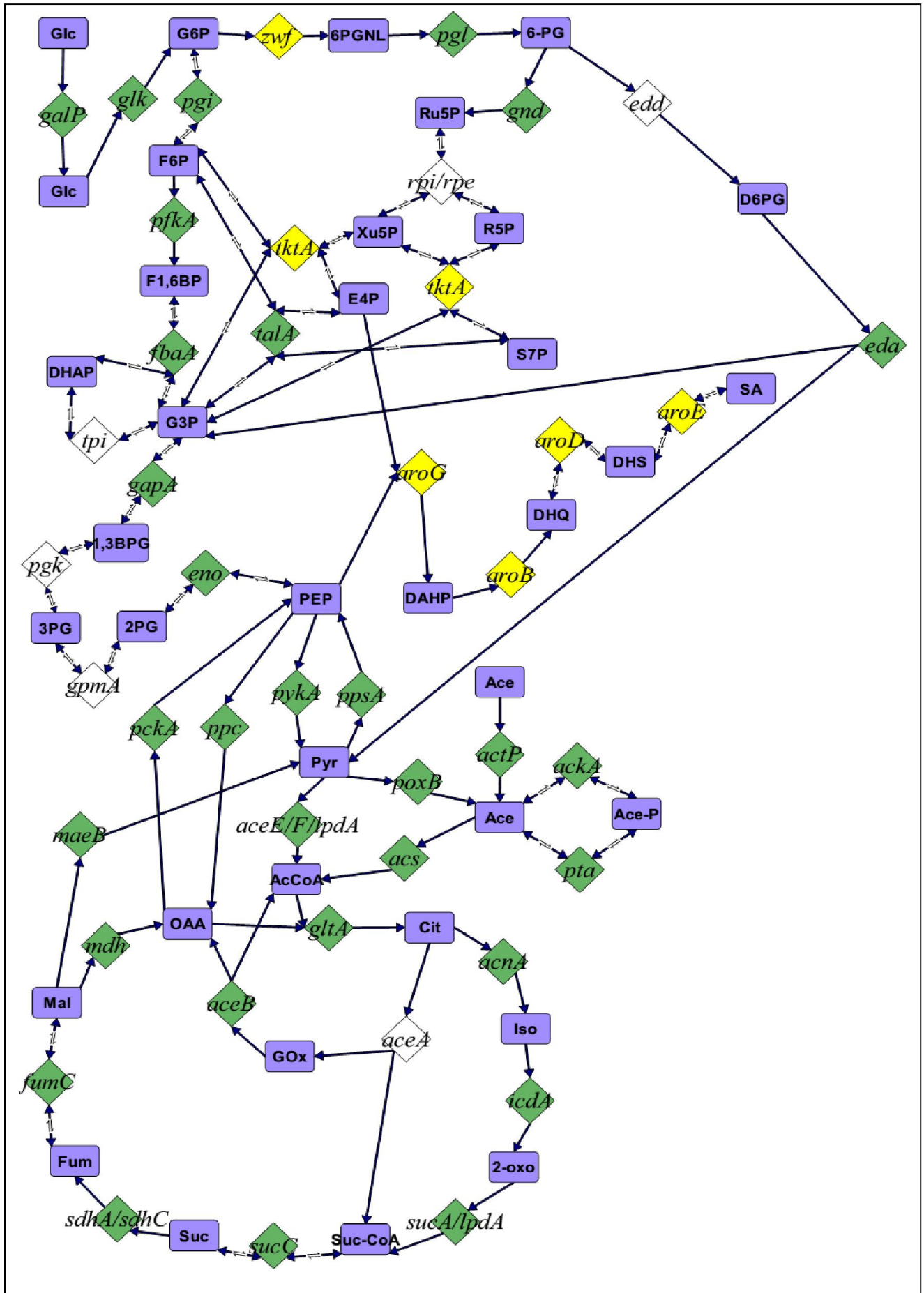


Figura 10. Representación de las reacciones catalizadas por las enzimas codificadas en los 40 genes evaluados en el análisis transcriptómico. Se muestran en rombos verdes los genes cuantificados en este estudio y en rombos blancos los genes no cuantificados. Los genes en rombos amarillos también fueron cuantificados y corresponden a los incluidos en el operón Aro6. Como referencia, los metabolitos producidos en cada reacción se indican en rectángulos azules y las abreviaturas utilizadas son las mismas a las descritas en las Figuras 1 y 14, excepto: D6PG, 2-deshidro-3-desoxi-gluconato 6-fostato; Iso, isocitrato; 2-oxo, 2-oxoglutarato; GOx, glioxilato; Suc-CoA, succinil-CoA. Productos de los genes señalados en la figura: *galP*, permeasa de galactosa; *glk*, glucocinasa; *pgi*, fosfoglucoosa isomerasa; *pfkA*, 6-fosfofructocinasa I; *tpi*, triosa fosfato isomerasa; *fbaA*, fructosa bisfosfato aldolasa; *gapA*, gliceraldehído 3-fosfato deshidrogenasa; *eno*, enolasa; *pykA*, piruvato cinasa II; *edd*, fosfogluconato deshidratasa; *rpi*, ribosa 5-fosfato isomerasa; *rpe*, ribulosa 5-fosfato epimerasa; *pgk*, fosfoglicerato cinasa; *gpmA*, fosfoglicerato mutasa; *actP*, permeasa de acetato; *acs*, acetil-coenzima A sintetasa; *pta*, fosfato acetiltransferasa; *ackA*, acetato cinasa; *poxB*, piruvato oxidasa; *aceE/F/lpdA*, subunidades del complejo piruvato deshidrogenasa; *gltA*, citrato sintasa; *acnA*, aconitasa I; *icdA*, isocitrato deshidrogenasa; *aceA*, isocitrato liasa; *aceB*, malato sintasa; *sucA/lpdA*, subunidades del complejo 2-oxoglutarato deshidrogenasa; *sucC*, subunidad del complejo succinil-coA sintetasa; *sdhA/sdhC*, succinato deshidrogenasa; *fumC*, fumarasa C; *mdh*, malato deshidrogenasa; *maeB*, enzima málica dependiente de NADP; *pckA*, fosfoenolpiruvato carboxicinas; *ppc*, fosfoenolpiruvato carboxilasa; *ppsA*, fosfoenolpiruvato sintetasa; *zwf*, glucosa 6-fosfato deshidrogenasa; *pgl*, 6-fosfogluconolactonasa; *gnd*, 6-fosfogluconato deshidrogenasa; *tktA*, transcetolasa I; *talA*, transaldolasa A; *eda*, 2-ceto-3-desoxigluconato 6-fosfato aldolasa; *aroG*, DAHP sintasa; *aroB*, DHQ sintasa; *aroD*, DHQ deshidratasa; *aroE*, shikimato deshidrogenasa.

de la fase de crecimiento en cuestión (comparando dos muestras de AR36 en distintas fases de crecimiento). Para una mejor discusión, los genes evaluados fueron separados en seis rutas o nodos metabólicos y los resultados se muestran en la Fig. 11 y Fig. 12. Cada gen fue medido por triplicado a partir de RNA extraído de dos experimentos independientes (ver Materiales y métodos) y se graficó el promedio y la desviación estándar obtenidos en cada caso. Los datos que aparecen en las gráficas se muestran en Material adicional 3.

3.1.1. Genes glucolíticos

Tras comparar AR36 contra AR3e, puede observarse que la presencia y expresión constitutiva del operón incrementa los niveles transcripcionales de varios genes codificantes para enzimas glucolíticas durante las fases EE y LE (Fig. 11a). El aumento en la expresión de *galP* y *glk* es particularmente interesante porque se ha reportado que sus productos controlan el importe y fosforilación de Glc en PB12, la cepa parental de AR36 [51]. También hay un incremento significativo en los niveles de *pgi* y *eno*, pero no *pykA*. Estos cambios podrían traducirse en una disponibilidad más alta de PEP y de fructosa 6-fosfato (la cual puede ser convertida en E4P por la transcetolasa codificada en plásmido), incrementando el rendimiento de compuestos aromáticos. Por otro lado, la expresión de

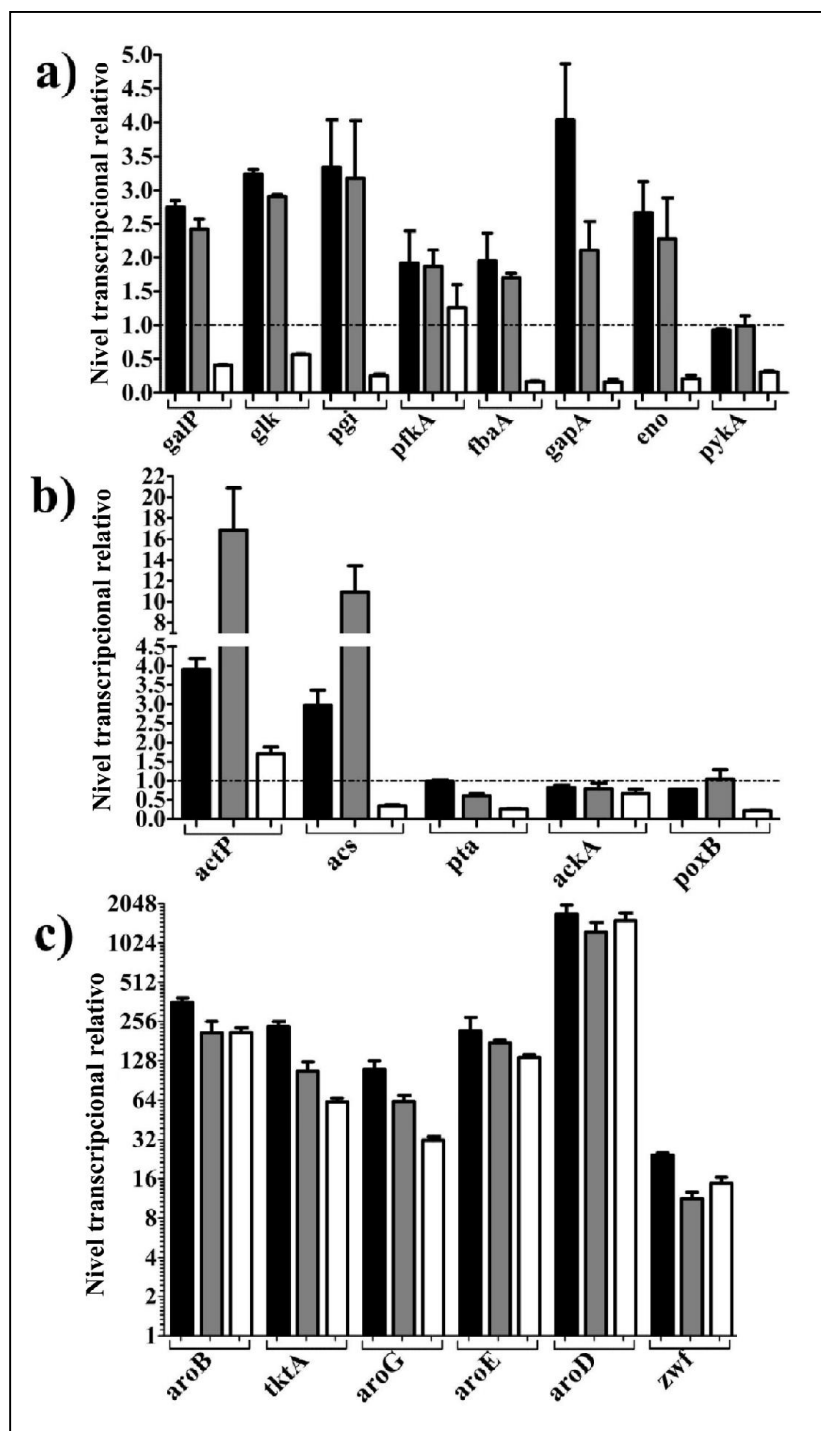


Figura 11. Cambios transcripcionales ocasionados por la expresión del operón sintético Aro6 sobre genes glucolíticos y de metabolismo de acetato. La cuantificación se efectuó en tres puntos diferentes en la curva de crecimiento de la cepa AR36, crecida con 50 g/L de Glc y 15 g/L de YE (ver Fig. 7a). Todos los datos fueron normalizados contra los valores obtenidos de la cepa AR3e en fase de crecimiento exponencial temprana, crecida en las mismas condiciones. **a)** Genes codificantes para enzimas glucolíticas; **b)** genes involucrados en la asimilación y biosíntesis de acetato; **c)** genes incluidos en el operón sintético Aro6 (ver Fig. 10 y Fig. 3). Barras negras: fase exponencial temprana; barras grises: fase exponencial tardía; barras blancas: fase estacionaria. Las barras de error representan la desviación estándar de los valores obtenidos por triplicado para cada una de las dos réplicas biológicas (Material adicional 3).

casi todos los genes cae notablemente en la fase ST, coincidiendo con una baja de 3 veces en el valor de q_s al entrar a fase estacionaria (Tabla 3). Estos resultados difieren de los reportados recientemente con la cepa PB12.SA22, donde no se observaron diferencias significativas en la expresión de genes de metabolismo central o de la vía de SA al pasar de fase exponencial a estacionaria [80]. Pareciera que la expresión constitutiva del operón en la cepa AR36 genera una situación metabólica difícil de mantener al entrar a fase estacionaria, teniendo como respuesta una baja drástica en la expresión de genes de metabolismo central.

3.1.2. Genes del metabolismo de acetato

Contrario a los genes glucolíticos, los niveles transcripcionales de genes codificantes para enzimas involucradas en la biosíntesis de acetato (*poxB*, *ackA* y *pta*) no se vieron modificados considerablemente por la presencia del operón sintético. Por otro lado, se detectó una fuerte sobreexpresión de los genes *actP* y *acs* en fases EE y LE, codificantes para enzimas involucradas en la asimilación de acetato (Fig. 10 y Fig. 11b). Cabe mencionar que la sobreexpresión de *actP* y *acs* en fase exponencial ha sido detectada anteriormente para la cepa parental PB12, la cual es capaz de co-utilizar Glc y acetato en medio mínimo [52, 57, 81]. Estos hallazgos correlacionan con los bajos niveles de acetato producidos por AR36 en esta condición de cultivo y nos permiten sugerir que AR36 podría utilizar simultáneamente Glc y acetato, incluso en presencia de YE (Fig. 7a). En contraste, los valores transcripcionales de estos genes fueron relativamente bajos en la fase ST (Fig. 11b). Este comportamiento podría explicar parcialmente la acumulación de acetato observada en fermentaciones con 100 g/L de Glc, las cuales consumen mayores cantidades de Glc durante la fase ST (Fig. 7b y Fig. 7c). Contrario a la estrategia comúnmente usada de inactivar a los genes de producción de acetato, los resultados resaltan a *actP* y *acs* como blancos potenciales para incrementar artificialmente su expresión en etapas tardías del cultivo, tomando ventaja de las capacidades esperadas de la cepa AR36 para co-utilizar ambas fuentes de carbono.

3.1.3. Genes contenidos en el operón *Aro6*

Los genes presentes en el operón sintético mostraron niveles de expresión muy altos, incluso en fase estacionaria, reflejando la naturaleza constitutiva del promotor y el alto número de copias del plásmido (Fig. 11c). Estos resultados correlacionan con los perfiles ininterrumpidos de consumo de Glc y producción de SA observados durante la fermentación (Fig. 7a), sugiriendo que las enzimas codificadas por los genes en el operón se encuentran presentes a lo largo de todo el cultivo. Puede observarse que los niveles de transcrito de *aroD* y *zwf* son comparativamente mayores y menores, respectivamente, a los otros cuatro genes en el operón. Esta observación debe abordarse con cuidado, ya que los seis genes en el operón están siendo comparados con los equivalentes en el cromosoma de la cepa de referencia AR3e. Tomando en cuenta que los valores de los seis genes no están normalizados entre ellos, variaciones entre su expresión cromosomal en la cepa AR3e pueden alterar las comparaciones relativas con la cepa AR36. No obstante, los datos transcriptómicos son consistentes con la alta proporción de SA con respecto a intermediarios aromáticos, lo que es de esperarse si todos los genes en el operón se han expresado adecuadamente. Junto con datos cinéticos y estequiométricos, estos resultados resaltan los beneficios de emplear un operón sintético expresado constitutivamente como una estrategia alternativa para incrementar el rendimiento de SA en una cepa que pueda metabolizar altas concentraciones de Glc.

3.1.4. Genes del ciclo de los ácidos tricarboxílicos (TCA)

De manera similar a los genes glucolíticos, los genes pertenecientes al TCA se vieron sobreexpresados como respuesta a la presencia del operón sintético, con excepción de *gltA*, *fumC* y *aceB* (Fig. 12a). Específicamente, se observaron incrementos en la expresión de *sucA* y *sucC*, codificantes para enzimas implicadas en el catabolismo de 2-oxoglutarato. Esta respuesta tiene sentido al considerar que los aminoácidos provenientes del YE producen intermediarios de TCA (2-oxoglutarato, succinil-CoA y oxaloacetato) mediante vías anapleróticas, generando un flujo de carbono importante a través de la parte baja del

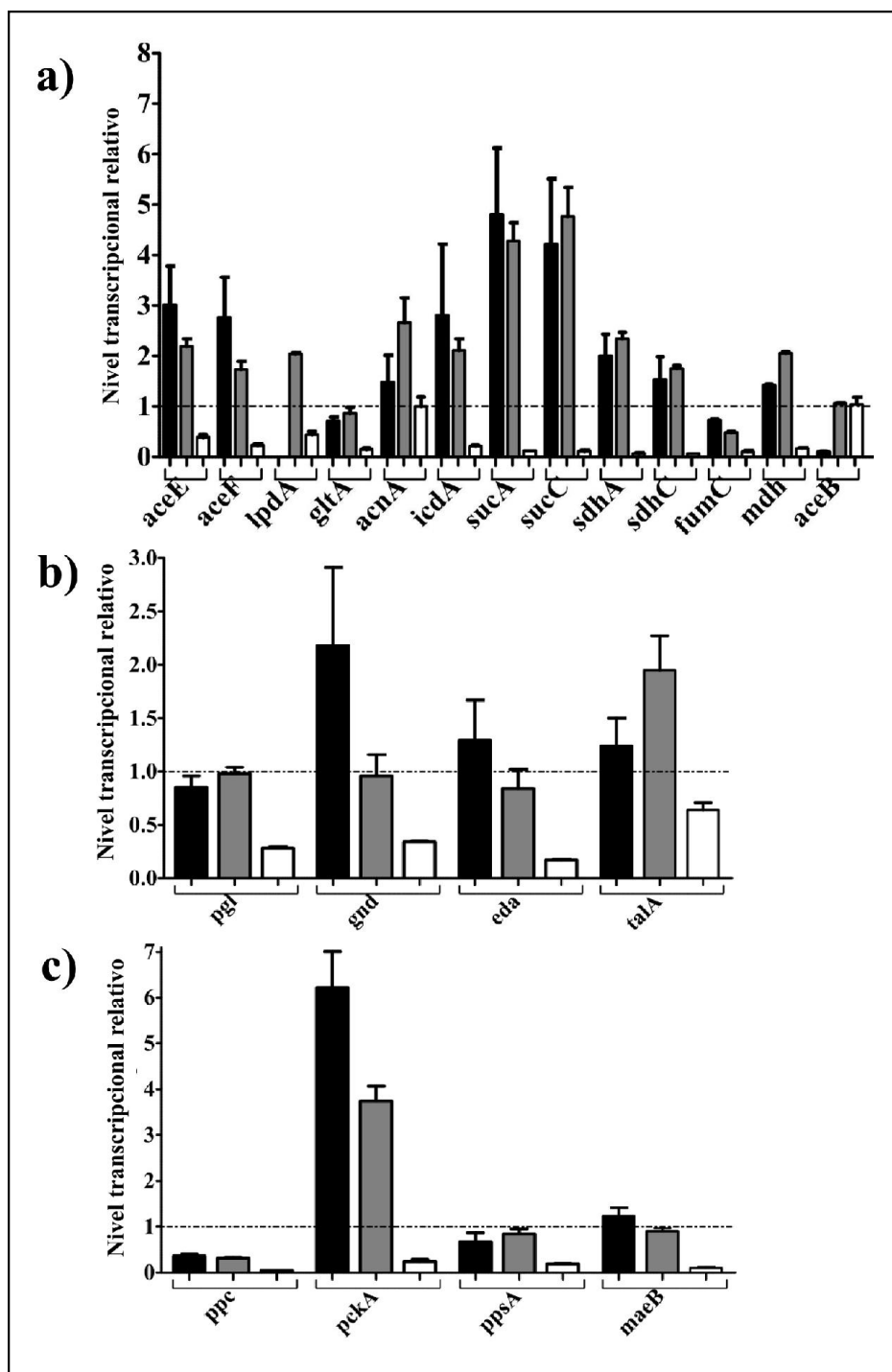


Figura 12. Cambios transcripcionales ocasionados por la expresión del operón sintético Aro6 sobre genes de TCA, PPP y nodo de PEP. La cuantificación se efectuó en tres puntos diferentes en la curva de crecimiento de la cepa AR36, crecida con 50 g/L de Glc y 15 g/L de YE (ver Fig. 7a). Todos los datos fueron normalizados contra los valores obtenidos de la cepa AR3e en fase de crecimiento exponencial temprana, crecida en las mismas condiciones. **a)** Genes codificantes de enzimas del ciclo de ácidos tricarbóxicos; **b)** genes codificantes para algunos genes de la vía de las pentosas; **c)** genes pertenecientes al nodo de PEP. Barras negras: fase exponencial temprana; barras grises: fase exponencial tardía; barras blancas: fase estacionaria. Las barras de error representan la desviación estándar de los valores obtenidos por triplicado para cada una de las dos réplicas biológicas (Material adicional 3).

ciclo. La sobreexpresión de *aceE*, *aceF* y *lpdA* (codificantes para el complejo piruvato deshidrogenasa), así como *acnA* e *icdA* (codificantes para las enzimas que controlan la oxidación de citrato), sugieren que a pesar de que la cepa tiene inactivado el gen *pykF*, existe todavía flujo de carbono de glucólisis a TCA. Estos resultados concuerdan con los obtenidos a partir de otros estudios enzimáticos, transcriptómicos y de análisis de flujos metabólicos empleando derivadas de PB12 con inactivaciones en *pykF* [16, 57, 82]. La conversión de PEP a piruvato debe depender completamente de la otra piruvato cinasa, *pykA*, y la entrada de carbono a TCA coincide con un requerimiento constante de oxígeno en las fermentaciones, lo que indica que el ciclo está activo durante todo el cultivo. No obstante, los niveles de transcripción relativa de todos los genes de TCA bajan en fase ST, sugiriendo que hay una menor oxidación de sustratos y menor generación de energía en esta fase, lo cual correlaciona con una baja en la *qs* y *qp* comparada con los valores en fase exponencial (Tabla 3).

3.1.5. Genes de la vía de las pentosas fosfato (PPP)

Se decidió medir el nivel de transcrito de cuatro genes de la vía de las pentosas para determinar si presentan cambios en respuesta a la sobreexpresión de los genes *tktA* y *zwf* contenidos en el operón Aro6. Los resultados muestran que *gnd* y *eda* presentan una ligera sobreexpresión al inicio del cultivo pero regresan a niveles similares a los de la cepa control en etapa LE, mientras que el gen *pgl* no presenta cambios. El gen *talA*, involucrado en la generación de E4P, también se observa ligeramente sobreexpresado en etapas tempranas pero su expresión relativa disminuye conforme progresa el cultivo (Fig. 12b). Se puede concluir que la sobreexpresión de *zwf* no genera sobreexpresión del resto de la vía oxidativa, por lo que la E4P podría ser principalmente generada a partir de intermediarios glucolíticos (F6P y G3P) mediante la vía no-oxidativa (empleando a la enzima TktA codificada en plásmido). La sobreexpresión conjunta de *zwf* y *tktA* podría aumentar simultáneamente la disponibilidad de NADPH en la primera reacción de la vía oxidativa (necesario en cantidades catalíticas por la enzima shikimato deshidrogenasa, AroE) y de E4P a través de la vía no-oxidativa. De cualquier manera, resulta interesante determinar el efecto de la presencia del operón sobre la acumulación de intermediarios de PPP, pues es

posible que incluso sin haber cambios en la expresión de genes que consumen a la 6-fosfogluconolactona, la actividad enzimática presente sea suficiente para convertir este metabolito de manera eficiente en precursores biosintéticos como ribosa 1-fosfato y E4P.

3.1.6. Genes del nodo de PEP

Uno de los nodos más interesantes con respecto a la producción de SA es el nodo de PEP, encargado de regular el flujo de carbono entre PEP, piruvato y oxaloacetato, así como las reacciones que canalizan intermediarios de TCA hacia gluconeogénesis [83]. La cuantificación relativa muestra que la expresión constitutiva del operón genera un aumento en la expresión durante la fase exponencial de 4 a 6 veces para el gen *pckA* y un decremento de 3 a 5 veces para el gen *ppc* (Fig. 12c). Este perfil llama la atención al considerar que la respuesta de una cepa silvestre a un incremento en el flujo glucolítico es precisamente la inversa, la sobreexpresión de *ppc* y la subexpresión de *pckA* [84–86]. Cuando hay un alto consumo de Glc como única fuente de carbono en condiciones aerobias, el flujo a través de TCA es insuficiente para oxidar todo el piruvato producido por glucólisis y resulta en el agotamiento de oxaloacetato, lo que arresta la entrada de acetil-CoA al ciclo y provoca acumulación de acetato [87]. Este fenómeno incide también sobre la acumulación de PEP y es por ello que la sobreexpresión de *ppc* es una respuesta celular empleada para reestablecer la disponibilidad de oxaloacetato en la célula y permitir que se siga consumiendo Glc. El hecho de que los datos indiquen una respuesta metabólica opuesta probablemente obedece a que el YE provee suficientes intermediarios de TCA durante la fase exponencial, por lo que no sería necesario generar oxaloacetato a partir de PEP. Por otro lado, el conjunto de reacciones catalizadas por las enzimas codificadas por *maeA*, *maeB* y *ppsA* conforman la principal vía gluconeogénica empleada por *E. coli* [83]; sin embargo, se observa que la expresión de *maeB* y *ppsA* no cambió considerablemente, en contraste con la alta expresión de *pckA* (Fig. 12c), proponiendo que esta ruta secundaria de generación de PEP es más activa en AR36. Además, los datos anteriores sugieren que los intermediarios de TCA obtenidos a partir del YE podrían contribuir a la generación de oxaloacetato, PEP e incluso SA, y una consecuencia interesante es que el agotamiento de oxaloacetato (o de sus precursores aspartato o glutamato) en el YE provocaría disminución

en el consumo de Glc. En esta situación la cepa tendría dificultad para regenerar oxaloacetato a partir de Glc, lo que probablemente afectaría a la fosforilación oxidativa y al balance de acarreadores de electrones, pudiendo ocasionar una acumulación de NADH incluso en condiciones de suficiencia de oxígeno. De esta manera, la deficiencia de oxaloacetato podría ser una de las causas de la mayor acumulación de acetato y de las disminuciones observadas en qs y qp durante la fase estacionaria.

Se debe tener presente que la discusión de las respuestas metabólicas presentadas aquí se da suponiendo que los cambios transcriptómicos generan cambios en los flujos de carbono. Aunque es cierto que el principal control de la actividad de una enzima en *E. coli* se da a nivel transcripcional [88], la actividad también depende de los factores reguladores propios de cada vía y de las características bioquímicas de las enzimas implicadas. Es por ello que los datos transcriptómicos deben complementarse con la medición de actividades enzimáticas, la determinación de flujos metabólicos o con la obtención de las concentraciones de compuestos intermediarios relevantes, como se describe a continuación.

3.2. Perfiles metabolómicos en respuesta a la presencia del plásmido y a la expresión del operón sintético Aro6

Los estudios de metabolómica permiten determinar de forma precisa las concentraciones de metabolitos relevantes en una situación fisiológica definida y se han vuelto una herramienta muy importante para evaluar el impacto de algunas estrategias de ingeniería metabólica [89, 90]. Es por ello que se decidió complementar los estudios transcripcionales con una cuantificación de compuestos de metabolismo central y la vía de AAA como respuesta a la presencia del plásmido y a la expresión del operón Aro6. Sin embargo, las condiciones de cultivo empleadas en este trabajo (cultivos en lote con altas concentraciones de YE) no son compatibles con los análisis finos de detección por espectrometría de masas. Este problema se agrava porque la auxotrofia de AAA que presentan las cepas desarrolladas en este trabajo hace que sea difícil crecerlas de manera eficiente sin usar YE. Por estas razones se decidió efectuar los experimentos con células en reposo, donde los inóculos fueron crecidos en medio rico bajo las condiciones normalmente empleadas y después las células fueron lavadas y transferidas a matraces que contienen medio mínimo con 3 g/L de Glc como

única fuente de carbono. De esta manera, se pretendió determinar la concentración de intermediarios metabólicos provenientes de la Glc en condiciones de nulo crecimiento pero con producción de SA, equivalente a la producción en fase estacionaria. Aunque las células en reposo no brindan información sobre la generación de biomasa y omiten la contribución del YE al metabolismo celular, en este caso pueden ser útiles para analizar los cambios ocasionados por la presencia del plásmido y por el redireccionamiento artificial de carbono hacia SA proveniente de Glc. Los experimentos se hicieron comparando tres cepas: la cepa de producción AR36, la cepa con plásmido vacío AR3e, y la cepa sin plásmido AR3, las cuales mostraron un consumo de Glc constante desde el inicio del experimento, sugiriendo que cada una se encuentra en un estado metabólico estable a pesar de exhibir valores diferenciales de q_s (Fig. 13).

Un resultado interesante es que la presencia del plásmido vacío incrementó en 75% el consumo de glucosa y que la cepa sin plásmido presenta la q_s más baja (Fig. 13). Aunque raro, este efecto ha sido observado para algunas combinaciones plásmido-cepa, donde se ha reportado que el tamaño y número de plásmidos usados pueden incrementar proporcionalmente la q_s , la demanda de oxígeno y las concentraciones de productos de la fermentación ácido mixta [42, 91, 92]. Una manera en que se ha explicado este efecto es que el plásmido aumenta el transporte dependiente de PTS mediante de una regulación transcripcional positiva de *ptsG* por parte el complejo cAMP-CRP, lo que coincide con un incremento de cAMP [42]. Interesantemente, este es el primer trabajo donde se observa que la presencia de plásmido promueve el transporte y catabolismo de glucosa en una cepa carente del sistema PTS. Es posible que en las cepas estudiadas derivadas de PB12 el complejo cAMP-CRP induzca la expresión del transportador GalP y de otros genes glucolíticos, los cuales también están controlados por CRP [93]. Por otro lado, la cepa que expresa las proteínas del operón sintético produce SA con un $Y_{SA/Glc}$, q_s y q_p similares a los obtenidos en fase estacionaria de los cultivos en fermentadores con altas concentraciones de Glc y YE (ver Tabla 3). Esto ayuda a validar la utilidad de estos experimentos al suponer que la distribución de flujos provenientes de Glc en la parte alta del metabolismo es similar a las condiciones de producción.

Para obtener los perfiles metabolómicos, se tomaron muestras de medio con células (*whole broth*) por triplicado de cada punto en el tiempo y se efectuó un lisado celular y apagado de

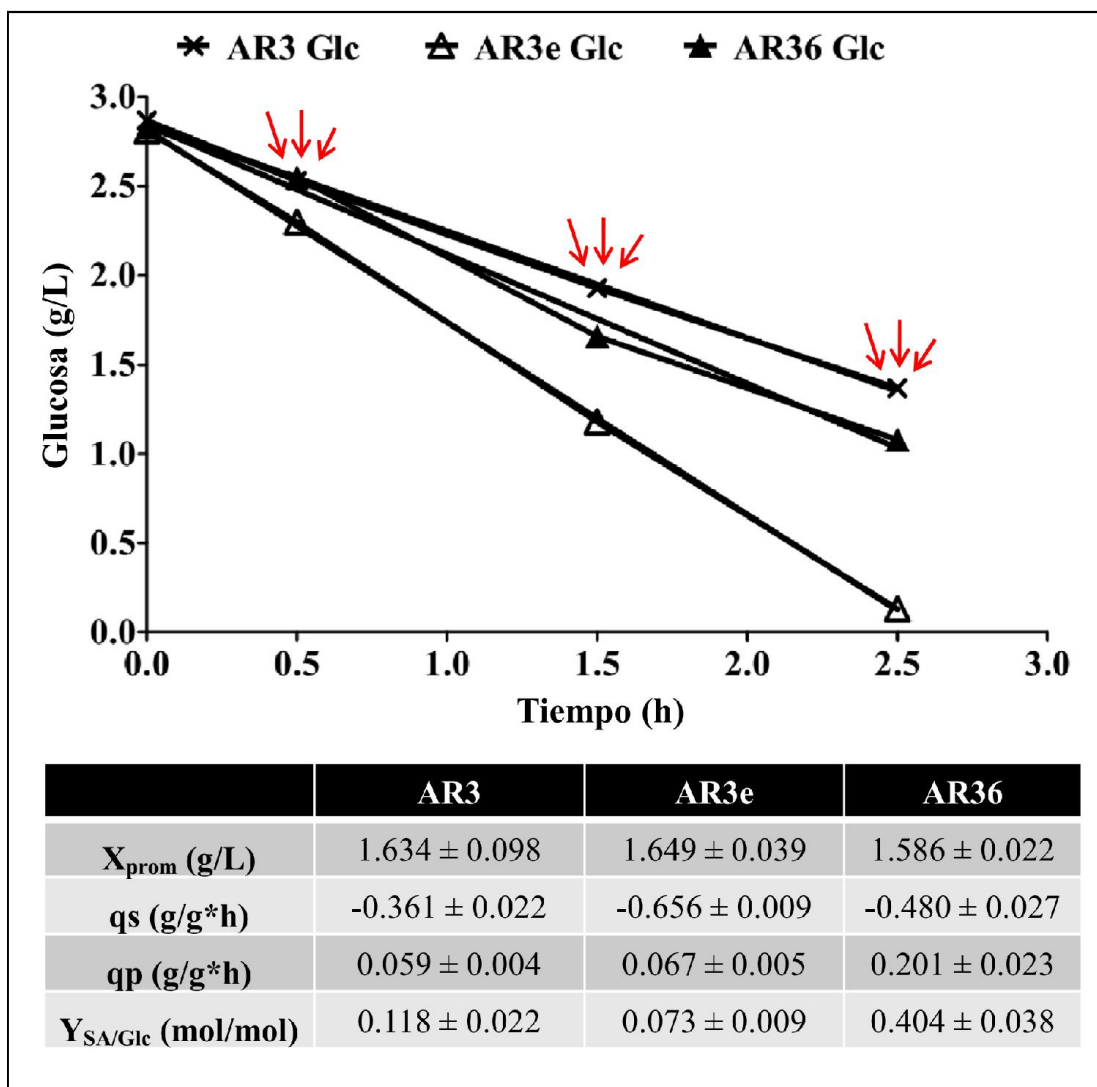


Figura 13. Perfiles de consumo de glucosa y parámetros de producción de SA obtenidos para las tres cepas en el sistema de células en reposo. Las flechas rojas indican los tiempos de muestreo empleados en el cálculo de concentraciones intracelulares y extracelulares de metabolitos. Los datos de la tabla representan el promedio de los tres puntos en el tiempo para cada cepa. X_{prom} = biomasa promedio; q_s = velocidad específica de consumo de Glc; q_p = productividad específica de SA; $Y_{SA/Glc}$ = rendimiento de SA a partir de Glc.

metabolitos intracelulares con una solución de acetonitrilo, metanol y agua (Materiales y métodos). Además, se agregó a cada muestra una solución de referencia de metabolitos (IDMS) que se extrae de la cepa silvestre de *E. coli* MG1655 crecida en un cultivo continuo con Glc marcada isotópicamente en todos los carbonos, para obtener una mezcla de intermediarios presentes durante el crecimiento balanceado [94]. De esa manera, cada

metabolito detectado en las muestras fue normalizado contra su concentración en el IDMS, permitiendo corregir variaciones producidas por el manejo de la muestra y hacer una comparación relativa entre las diferentes concentraciones de compuestos detectados. Desafortunadamente, el método de detección empleado no nos permitió determinar las concentraciones de oxaloacetato, piruvato, G3P ni acetil-CoA, por lo que deben buscarse alternativas para su cuantificación. Para contribuir al análisis metabolómico, se cuantificó también el rendimiento de glucógeno producido por unidad de peso seco.

En la Figura 14 se muestra la comparación del metaboloma para las tres cepas; las columnas, de izquierda a derecha, corresponden a las concentraciones en AR3, AR3e y AR36 y las barras de error indican la desviación estándar a lo largo del tiempo. Pueden observarse diferencias interesantes entre las tres cepas, que coinciden con los consumos de Glc calculados. En primer lugar, la cepa con plásmido vacío (AR3e), que consume glucosa más rápidamente que las demás, acumula una mayor concentración de productos de fermentación, intermediarios del TCA, glucosa 1-fosfato (G1P), que es precursor de glucógeno (GLG) y N-acetilglucosamina 1-fosfato (GlcNAc-1P), empleada en la síntesis de pared celular. Por el contrario, la cepa AR36 no acumula lactato, etanol o acetato, probablemente porque la sobreexpresión de los genes del operón impone una redirección del PEP hacia la vía del SA. Esto previene la formación de productos tóxicos como acetato y etanol, que se acumulan casi 100 veces más en AR3e que en AR36 (Fig. 14).

Como se esperaba para la cepa AR36, los niveles intracelulares de fructosa 1,6-bisfosfato (F1,6BP) son bajos debido a que la sobreexpresión de la transcetolasa (*tktA*) promueve la conversión de F6P en E4P. Sin embargo, los niveles de PEP no son menores que en las otras cepas a pesar de la sobreexpresión de la enzima DAHP sintasa (*aroG*). Los resultados indican que la cepa es capaz de compensar la baja en F1,6BP y de regenerar intermediarios glucolíticos a niveles normales a partir del nodo de 1,3 bisfosfoglicerato (1,3BPG). Este hallazgo es relevante porque implica que existen sistemas de control activos para evitar la falta de precursores esenciales como G6P, F6P y PEP que funcionan sin importar la tasa de consumo de Glc y que responden incluso a una perturbación fuerte como la expresión constitutiva en plásmido de una vía biosintética.

También fue inesperado encontrar una concentración 100 veces mayor de PEP que de E4P, lo que sugiere que cualquier modificación que logre aumentar la disponibilidad de E4P

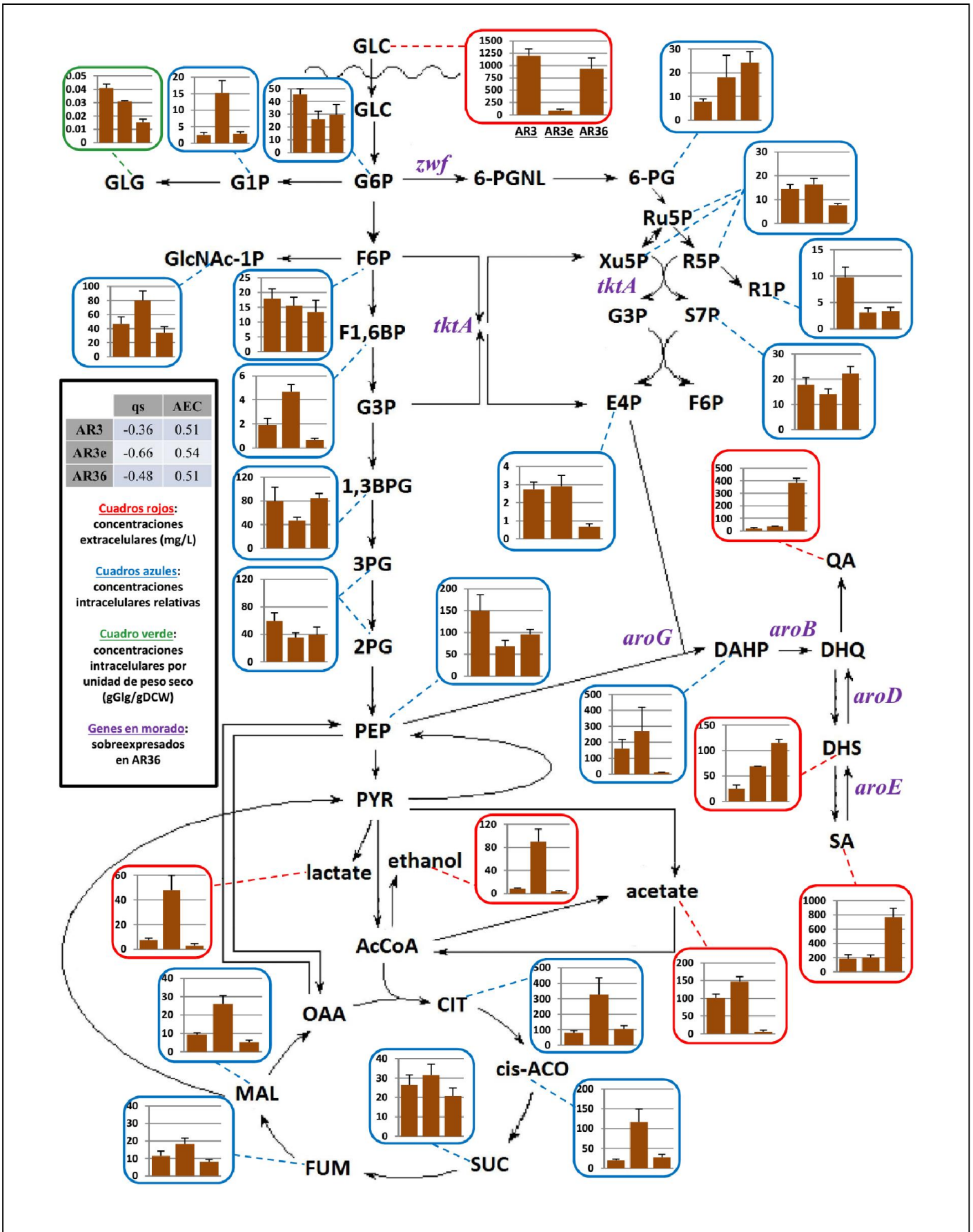


Figura 14. Metabolómica comparativa de las tres cepas estudiadas. La primera columna de izquierda a derecha corresponde a la cepa sin plásmido AR3, la columna central corresponde a la cepa con plásmido vacío AR3e y la tercera columna de izquierda a derecha corresponde a la cepa que expresa el operón sintético AR36. Los cuadros rojos indican metabolitos medidos en el sobrenadante; los cuadros azules indican la concentración relativa al IDMS de metabolitos intracelulares; el cuadro verde indica la concentración intracelular de glucógeno. Se indican además en morado los genes sobreexpresados en AR36, así como las velocidades específicas de consumo de Glc (qs) y los valores de la carga energética de adenilatos (AEC) calculados con las concentraciones intracelulares de ATP, ADP y AMP. Abreviaturas de intermediarios no definidos anteriormente: G1P: glucosa 1-fosfato; GLG: glucógeno; GlcNAc-1P: N-acetilglucosamina 1-fosfato; 6-PG: 6-fosfogluconato; 1,3BPG: 1,3-bisfosfoglicerato; 3PG: 3-fosfoglicerato; 2PG: 2-fosfoglicerato; R1P: ribosa 1-fosfato; cis-ACO: cis-aconitato; SUC: succinato; FUM: fumarato; MAL: malato.

podría tener una ventaja inmediata sobre el rendimiento y producción de SA. Este estado del sistema resulta muy útil para evaluar otras estrategias que incrementen el flujo de carbono hacia E4P, por ejemplo sobreexpresar a la enzima transaldolasa para aumentar la capacidad de generación de E4P a partir de sedoheptulosa 7-fosfato (S7P), la cual es 30 veces más abundante (Fig. 14). De manera similar, la sobreexpresión del gen *zwf* probablemente ocasionó una acumulación de 6-fosfogluconato (6-PG) que resultó ser 5 veces mayor que las concentraciones combinadas de tres pentosas fosfato (Ru5P, Xu5P y R5P), sugiriendo que la actividad específica de las enzimas que consumen 6-PG (Rpe y RpiA) no es suficiente para evitar su acumulación y puede ser otro blanco importante para su modificación artificial. Interesantemente, ninguna de estas aproximaciones ha sido evaluada para el mejoramiento de cepas productoras de SA.

Por otro lado, una de las razones por las cuales la cepa sin plásmido AR3 consume menos Glc y lo hace de manera más lenta, podría ser la acumulación observada de intermediarios metabólicos como G6P, F6P, R1P y PEP. Esto coincide con una teoría que postula que la presencia de un plásmido aumenta el consumo de glucosa porque induce un redireccionamiento de carbono hacia vías de síntesis de nucleótidos y aminoácidos y simultáneamente incrementa las rutas de generación de energía para poder replicarse y expresar las proteínas presentes en el mismo [42]. Sin embargo, parece que este incremento en el consumo (del 75% en nuestro caso) alcanza un estado en el cual las vías catabólicas son superiores a las anabólicas y se genera un sobreflujo de carbono hacia vías fermentativas. Es por ello que la expresión de una vía biosintética como la del SA que haga uso de este exceso de carbono, energía y poder reductor, servirá para generar SA con mayor

rendimiento y aminorar la acumulación de intermediarios que indican suficiencia energética como G1P, GLG y GlcNAc-1P. No obstante, puede esperarse que este cambio vaya acompañado de otros efectos adversos provocados por la expresión excesiva de proteínas codificadas en plásmido, usualmente difíciles de caracterizar por separado y nombrados “carga metabólica” por simplicidad [41]. De hecho, la expresión del operón disminuye la qs en 30%, aunque se mantiene por encima del valor de la cepa sin plásmido (Fig. 13). Se calculó también el valor de la carga energética de adenilatos (AEC) usando las concentraciones promedio de AMP, ADP y ATP de la siguiente manera: $AEC = ([ATP] + 1/2[ADP]) / ([ATP] + [ADP] + [AMP])$, y se observó que interesantemente la carga energética es muy similar en las tres cepas con un valor de 0.5 (Fig. 14), lo que corresponde a células metabólicamente activas en fase estacionaria, en comparación a un valor de 0.9 para células en fase exponencial [95]. Junto con los datos de transcriptoma, los datos de metaboloma resaltan nuevamente la plasticidad de la cepa ante diferentes condiciones y sugiere que la expresión constitutiva en plásmido impone una carga metabólica pero puede también darle un beneficio al aumentar la capacidad de consumo de Glc y producción de SA por periodos de tiempo más largos, evitando la acumulación de productos tóxicos.

Es importante resaltar que la adición de YE al medio de cultivo en las fermentaciones de producción seguramente aminorará algunos efectos negativos causados por la expresión constitutiva, pero también podría cambiar los perfiles energéticos y metabólicos al sumar la contribución de otros componentes. Por ejemplo, la abundancia de glutamato y aspartato en el YE podría generar energía y precursores por periodos de tiempo más largos, pero también podría imponer una mayor sobrecarga en las vías oxidativas como glucólisis y TCA. Los resultados de cultivos en lote con altas concentraciones de sustrato muestran que esta sobrecarga puede ser aliviada redirigiendo esqueletos de carbono provenientes de Glc y YE hacia SA, permitiendo que se consuman más de 100 g/L de Glc y 30 g/L de YE. Asimismo, las bajas concentraciones intracelulares de malato encontradas en AR36 en las células en reposo (análogo a dejar de suplementar YE o al agotamiento de precursores anapleróticos) soportan la hipótesis generada a partir de los datos de transcriptoma que sugiere que hay un flujo gluconeogénico importante a través de PckA que drena intermediarios de TCA en forma de oxaloacetato para generar PEP. Si se confirman estas aseveraciones con otro tipo de experimentos (actividades enzimáticas, flujos metabólicos

con mutantes en *PckA*, suplementación de aminoácidos gluconeogénicos, etc.), sería posible diseñar técnicas de alimentación de sustratos para prolongar la fase exponencial y aumentar el rendimiento y productividad de SA. Finalmente, aunque las respuestas metabólicas observadas aquí son similares a algunas reportadas anteriormente [91, 92, 96], existen otras que son particulares de las cepas evaluadas, como la presencia de un metabolismo glucolítico y gluconeogénico simultáneo, lo cual se ha demostrado en la parental PB12 creciendo con diversas fuentes de carbono [53, 57, 81, 97]. De este modo, las cepas construidas en este trabajo pueden convertirse en modelos de estudio interesantes, no solamente para la producción de SA o de compuestos derivados, sino también para indagar las interacciones de plásmidos con el resto de procesos celulares de *E. coli*. De cualquier manera, los resultados de la caracterización efectuada en este trabajo junto con los buenos parámetros de producción obtenidos justifica hacer más modificaciones racionales al sistema y condiciones de cultivo, buscando encontrar un balance apropiado entre el crecimiento y el rendimiento de SA para optimizar la productividad del proceso.

CONCLUSIONES

Hasta la fecha, *E. coli* es el microorganismo que ha dado los mejores resultados para la producción de SA, con cepas modificadas que pueden acumular hasta 84 g/L usando fermentadores de 10 L y cultivos en modo lote alimentado (Tabla 4) [10, 20]. En este trabajo se observó por primera vez que la inactivación simultánea de PTS y *pykF* es una estrategia viable para incrementar el título y el rendimiento de SA cuando se emplean cultivos en lote con altas concentraciones de Glc y YE. La cepa productora AR36 exhibió el mayor rendimiento reportado de SA (0.42 mol/mol) y una alta proporción molar de este compuesto a otros subproductos aromáticos (80%). Esto representa un avance muy importante en la identificación de las condiciones necesarias para redirigir eficientemente Glc hacia la vía de AAA, lo cual históricamente ha representado un reto para los ingenieros metabólicos a causa de la longitud y gran regulación de esta vía. Importantemente, la cepa AR36 puede tolerar y consumir altas concentraciones iniciales de Glc y se encontró una relación entre esta capacidad y la sobreexpresión de genes glucolíticos, de TCA y de asimilación de acetato. Además, se determinó que la sola presencia del plásmido aumenta

el consumo de Glc y que la expresión de las enzimas codificadas en el operón sintético podría funcionar como una vía de escape para el exceso de intermediarios glucolíticos que de otra forma se convertirían en acetato, lactato, etanol y otros metabolitos. En este trabajo se reporta el primer estudio metabolómico de una cepa productora de SA y uno de los resultados más interesantes es el descubrimiento de una limitación en la disponibilidad de E4P pero no de PEP. Lo anterior justifica la evaluación de otras modificaciones en la vía de las pentosas, una parte del metabolismo sorprendentemente poco estudiada con respecto a la producción de SA en *E. coli*. Una característica notable de la cepa AR36, a diferencia de otras cepas empleadas para la producción de compuestos aromáticos, es la posibilidad de co-utilizar Glc y sustratos gluconeogénicos como acetato y componentes del YE. Esta capacidad ha sido demostrada en su parental PB12 y el descubrimiento de la fuerte sobreexpresión de genes como *actP*, *acs* y *pckA*, indican que es probable que uno de los factores por los cuales el rendimiento de SA es alto en AR36 es que podría formarse PEP y otros precursores esenciales a partir de la asimilación varios sustratos de manera simultánea.

Por otro lado, aunque aquí se reporta que es posible utilizar cultivos en modo lote para producir altas concentraciones de SA, el hecho de que la producción no cesa sino hasta que se agota la Glc sugiere que se pueden alcanzar mayores títulos de SA con estrategias adecuadas de alimentación. Fermentaciones preliminares de AR36 en modo lote alimentado han resultado en la producción de 60 g/L de SA en 24 h, usando un fermentador de 2 L al que se alimentó durante 14 h con Glc y YE (Material adicional 5). Interesantemente, este proceso a escala laboratorio permitió la generación de una gran concentración celular (aprox. 170 OD₆₀₀) y resultó en la productividad volumétrica más alta obtenida para cualquier proceso de producción de SA, 2.46 gSA/L*h (Tabla 4). Incluso después de que la cepa AR36 consumiera alrededor de 180 g totales de Glc, el rendimiento de SA obtenido en este cultivo de alta densidad celular (0.42 mol/mol) fue similar a los observados en otras condiciones de cultivo evaluadas en este trabajo. En conjunto, los resultados obtenidos postulan a la cepa y a las estrategias de cultivo descritas en este trabajo como una alternativa competitiva para la sobreproducción de SA y otros compuestos derivados de interés industrial.

PERSPECTIVAS

La información obtenida en este estudio permite definir algunas perspectivas interesantes. La primera comprende evaluar la sobreexpresión constitutiva del gen codificante para la transaldolasa (*talB*) en la cepa AR36 y aumentar así la conversión de S7P a E4P, considerando que se detectó una diferencia de 30 veces en sus concentraciones intracelulares y que la limitación de E4P probablemente restringe el rendimiento alcanzado de SA. La secuencia codificante de este gen podría insertarse en el plásmido pTrcAro6 como parte del operón sintético, haciendo uso de los sitios únicos de restricción restantes y quedar controlado por el mismo promotor. La segunda propuesta es la sobreexpresión del operón de asimilación de acetato y del gen codificante para la enzima PEP carboxinasa (*actP*, *acs* y *pckA*) en etapas tardías del cultivo, haciendo un reemplazo cromosomal de sus promotores por otros constitutivos o de fase estacionaria. Esto le daría a la cepa resultante el potencial de consumir Glc en conjunto con sustratos gluconeogénicos como acetato o componentes del YE por periodos más largos y sería factible aprovechar esta característica en cultivos alimentados con varias fuentes de carbono. Además de minimizar la acumulación de acetato, esta estrategia debería promover una mayor disponibilidad de acetil-CoA y PEP, lo cual podría generar un efecto positivo sobre el crecimiento y la producción de SA. Finalmente, resulta interesante comprobar algunas de las hipótesis planteadas a partir de los resultados transcriptómicos y metabolómicos con la cuantificación de los flujos de carbono y determinación de actividades enzimáticas. La integración de estos diferentes niveles de información celular permitiría conocer otras características importantes para mejorar el proceso de producción.

MATERIALES Y MÉTODOS

Construcción de plásmidos y cepas de *Escherichia coli*

La cepa evolucionada en laboratorio PB12, derivada de PB11 (obtenida por la inactivación de PTS en la cepa JM101 [98]), fue la receptora de las modificaciones genéticas descritas en este trabajo [51, 52, 54]. Las cepas y plásmidos usados se enlistan en la Tabla 1. Las inactivaciones cromosomales de los genes *aroK*, *aroL* y *lacI* se efectuaron de manera

secuencial mediante recombinación homóloga de productos de PCR [99]. En todos los casos, se empleó al plásmido pKD3 como templado de PCR en conjunto con oligonucleótidos diseñados para incluir una homología de 45 bp con la secuencia objetivo en el cromosoma (Material adicional 4). El plásmido pKD46 permitió la expresión del sistema de recombinasa Red del bacteriófago lambda y el plásmido pCP20 facilitó la remoción del cassette de resistencia a cloranfenicol después de cada evento [100]. Se verificó cada paso por PCR, identificando las clonas que presentaron los tamaños de amplicón esperados al usar diferentes sets de oligonucleótidos y DNA cromosomal como templado (Material adicional 4). El gen inactivado *pykF* fue transducido en la cepa PB12 *aroK⁻ aroL⁻* usando un lisado del fago P1 obtenido a partir de la cepa PB28 (*pykF::gen*) [82]. Se seleccionaron transductantes en cajas con gentamicina (10 µg/mL) y la inactivación fue confirmada por PCR.

La construcción del operón Aro6 y del vector de expresión se llevó a cabo en varios pasos (Figura 3 y Material adicional 2). Primero, los genes *aroB*, *aroG^{fbr}*, *tktA* y *aroE* fueron amplificados por PCR utilizando a la DNA polimerasa Pfu y ligados secuencialmente en el polilinker del plásmido pBRINT-Ts Cm [69]. Se usó DNA cromosomal de la cepa JM101 como templado para la amplificación de los genes requeridos, con la excepción del gen *aroG^{fbr}*, el cual fue amplificado a partir del plásmido pJLBaroG^{fbr}*tktA* [22]. Diferentes pares de oligonucleótidos fueron empleados para la amplificación de cada gen (Material adicional 4), lo que generó también sitios de restricción en los costados del amplicón y sitios Shine-Dalgarno con secuencias consenso (AGGAGG) situadas 8 bp río arriba de cada secuencia codificante. Los productos de PCR fueron insertados en el polilinker en el orden siguiente: *aroB* en el sitio *SmaI*, *aroG^{fbr}* en el sitio *XhoI*, *tktA* en el sitio *EcoRV* y *aroE* en el sitio *ApaI*. De manera simultánea se construyó el plásmido pTrc327par*lacI⁺* (Material adicional 2) mediante la ligación de un fragmento amplificado por PCR que incluyó al gen *lacI*, el promotor Trc, el polilinker y los terminadores de la transcripción del plásmido pTrc99A [62] en los sitios *ScaI* y *PvuI* del plásmido pBR327par [63]. El operón de 4 genes presente en el plásmido pBRINT-Ts Cm fue amplificado por PCR con un par de oligonucleótidos (Material adicional 4) y ligado en el plásmido pTrc327par*lacI⁺* después de digerir ambos con *SacI* y *XbaI*. Posteriormente, se amplificó por PCR al gen *aroD* (flanqueado por sitios *NheI*) y ligado en el sitio compatible *XbaI* del pTrc327par*lacI⁺*. A

partir del interés de expresar el operón de manera constitutiva, se generó un plásmido derivado *lacI* (sin operón sintético) al que se nombró pTrc327par (Material adicional 2). El operón de 5 genes se transfirió a los sitios *SacI* y *NcoI* del pTrc327par, generando el plásmido pTrcAro5. Finalmente, se insertó el gen *zwf* en el sitio *XbaI*, creando un operón de 6 genes en pTrc327par. El plásmido resultante se denominó pTrcAro6 y fue utilizado para transformar a las cepas AR2 y AR3, generando a las derivadas AR26 y AR36, respectivamente (Tabla 1). Las cepas transformadas fueron seleccionadas en cajas de LB suplementadas con tetraciclina (30 µg/mL).

Cada paso en la clonación de genes y la construcción de los plásmidos fue verificado por digestión con endonucleasas y por PCR, visualizado por electroforesis en gel de agarosa teñido con bromuro de etidio y verificado por secuenciación de DNA (3730, Perkin-Elmer/Applied Biosystems, USA). Todas las enzimas y reactivos utilizados en los procedimientos de biología molecular fueron adquiridas de Fermentas (USA) y New England Biolabs (USA). Se emplearon kits para la purificación de DNA (Roche, Suiza) a partir de PCR, plásmido o gel de agarosa cuando fue requerido. Se usaron células TOP10 (Invitrogen, USA) como hospedero para el rastreo de ligaciones de DNA durante pasos intermedios en la construcción del vector.

Medio de cultivo y condiciones de crecimiento

Composición del medio de producción

El medio de producción de SA (ajustado a pH 7.0 con NaOH 10 N) contiene por litro: K₂HPO₄ (7.5 g), KH₂PO₄ (7.5 g), ácido cítrico monohidratado (2.1 g), citrato de amonio férrico (0.3 g), H₂SO₄ concentrado (1.2 mL), MgSO₄ (0.64 g), CaCl₂ (0.06 g), (NH₄)₆(Mo₇O₂₄) (0.0037 g), ZnSO₄ (0.0029 g), H₃BO₃ (0.0247 g), CuSO₄ (0.0025 g), MnCl₂ (0.0158 g), CoCl₂ (0.00129 g), hidrocloreuro de tiamina (0.001 g) y betaína (0.234 g) como osmoprotector. Se añadió también tetraciclina (30 µg/mL) a inóculos y cultivos para el mantenimiento de plásmidos. Se suplementó glucosa (esterilizada por filtración) y extracto de levadura (adicionado antes de esterilizar en autoclave) a las concentraciones requeridas para cada experimento. La glucosa fue comprada en JT Baker (USA) y el extracto de levadura autolisado en BD Difco (USA).

Cultivos en matraces agitados

La preparación de los inóculos empezó con la adición de alícuotas congeladas de 1 mL a matraces baffleados con 25 mL de medio de producción suplementado con glucosa (25 g/L) y extracto de levadura (15 g/L). Los inóculos fueron crecidos a 37°C y 300 rpm hasta la mitad de la fase exponencial, donde se transfirió aproximadamente 5 % del volumen final a los matraces donde se efectuaron los experimentos y fueron incubados a las mismas condiciones controladas con 15 g/L de glucosa y 5 g/L de extracto de levadura. El crecimiento celular se midió al monitorear la densidad óptica a 600 nm (OD_{600}) en un espectrofotómetro DU700 (Beckman, USA) y se colectaron y centrifugaron muestras de manera periódica. Los sobrenadantes de las muestras tomadas fueron almacenados a -20°C para su posterior análisis de metabolitos. Estos experimentos se efectuaron cuando menos por triplicado. Todos los cultivos empezaron aproximadamente en 0.3 OD_{600} .

Cultivos en fermentadores

Se llevaron a cabo cultivos en lote, cuando menos por duplicado, usando reactores de vidrio de 1 L (Applikon, Países Bajos) con 500 mL de volumen de trabajo. Los bioreactores fueron conectados a un controlador Applikon ADI 1010 para monitorear la temperatura, el pH, la velocidad de agitación y la tensión de oxígeno disuelto (DOT). El pH se mantuvo en 7.0 mediante la adición de H_3PO_3 (3.3%) y de NH_4OH (10%). La DOT en el medio de cultivo se mantuvo mediante la adición constante de aire filtrado (1 vvm) y controlando manualmente la velocidad del impulsor (desde 500 hasta 1000 rpm) para asegurarse que la DOT se mantuviera por encima de 20% en todo momento. La preparación del inóculo para los fermentadores empezó con la adición de una alícuota congelada a matraces baffleados de 500 mL con 50 mL de medio de producción suplementado con glucosa (25 g/L) y extracto de levadura (15 g/L). Los inóculos fueron crecidos a 37°C y 300 rpm hasta la mitad de la fase exponencial, donde se transfirió aproximadamente 5 % del volumen final a los reactores previamente preparados con medio de producción. Todas las fermentaciones se llevaron a cabo en presencia de tetraciclina (30 $\mu\text{g/mL}$). El crecimiento celular se midió al monitorear la densidad óptica a 600 nm (OD_{600}) en un espectrofotómetro DU700 (Beckman, USA) y se colectaron muestras de manera periódica. Los sobrenadantes de las

muestras tomadas fueron almacenados a -20°C para su posterior análisis de metabolitos. Todas las fermentaciones empezaron aproximadamente en 0.3 OD_{600} .

Cuantificación de metabolitos extracelulares

El sobrenadante de cada muestra se diluyó cuando fue requerido y se filtró con membranas de nylon de $0.45 \mu\text{M}$. Se determinaron las concentraciones de shikimato (SA), 3-deshidroshikimato (DHS), 3-deshidroquinato (DHQ), ácido quínico (QA), ácido gálico (GA), ácido acético y glucosa por HPLC mediante un sistema Waters (bomba cuaternaria 600E, inyector automático 717, y detectores de índice de refracción 2410 y arreglo de diodos 996; USA) equipado con una columna Aminex HPX-87H ($300 \times 7.8 \text{ mm}$; $9 \mu\text{m}$; Bio-Rad, USA). La fase móvil fue H_2SO_4 5mM con un flujo de 0.5 mL/min , mantenido a 50°C . Las concentraciones de 3-desoxi-D-arabinoheptulosoato 7-fosfato (DAHP) se determinaron colorimétricamente mediante el ensayo del ácido tiobarbitúrico [101]. Este método no distingue entre DAHP y su forma desfosforilada, DAH, por lo que las concentraciones reportadas en este trabajo representan la suma de ambas especies.

Análisis de datos y cálculos realizados

Las concentraciones medidas de metabolitos y biomasa fueron normalizadas al volumen inicial para considerar los cambios de volumen derivados del control en fermentadores. Los datos de experimentos independientes fueron promediados y presentados en las gráficas correspondientes, donde las barras de error indican la desviación estándar para cada punto. La concentración de biomasa (X) se determinó con una curva de calibración que relaciona el peso seco celular y la OD_{600} , resultando en la ecuación $X = 0.3587 * \text{OD}_{600}$. La velocidad específica de crecimiento (μ) se determinó al ajustar linealmente la concentración de biomasa con respecto al tiempo durante la fase exponencial con la siguiente ecuación: $\ln X = \ln X_0 + \mu * t$ (donde t es el tiempo y X_0 es la concentración inicial de biomasa), mostrando valores de R^2 superiores a 0.97. El rendimiento de SA a partir de Glc ($Y_{\text{SA/Glc}}$) se calculó con las concentraciones molares promedio de SA y Glc, producidas y consumidas, respectivamente, en el punto más alto de concentración de SA. El rendimiento de compuestos aromáticos totales a partir de Glc ($Y_{\text{TAC/Glc}}$) se calculó con los rendimientos molares combinados de DAHP, DHQ, DHS, SA, QA y GA en el punto más alto de

concentración de SA. El rendimiento máximo teórico de compuestos aromáticos ha sido previamente estimado como $0.86 \text{ mol}_{\text{TAC}}/\text{mol}_{\text{Glc}}$ para cepas PTS⁻ creciendo en Glc como única fuente de carbono [10]. La productividad volumétrica global de SA ($Q_{p_{\text{global}}}$) y la velocidad volumétrica global de consumo de Glc ($Q_{s_{\text{global}}}$) fueron calculadas tomando en cuenta el tiempo necesario para alcanzar la máxima concentración de SA. Además de los cálculos anteriores, se llevaron a cabo linearizaciones para obtener rendimientos aparentes de biomasa sobre sustrato ($Y_{X/S}$) y producto sobre biomasa ($Y_{P/X}$). A pesar de que estos rendimientos aparentes no toman en cuenta el consumo de extracto de levadura, los valores de correlación encontrados en todos los experimentos fueron superiores a 0.95, permitiendo hacer comparaciones entre ellos. Estos rendimientos se usaron para calcular la productividad específica y la velocidad específica de consumo de sustrato en la fase exponencial ($q_{p_{\text{exp}}}$ y $q_{s_{\text{exp}}}$, respectivamente) con las siguientes ecuaciones: $q_{p_{\text{exp}}} = Y_{P/X} * \mu$; $q_{s_{\text{exp}}} = \mu * Y_{X/S}$. La productividad volumétrica y la velocidad volumétrica de consumo de Glc en fase estacionaria fueron determinadas mediante la linearización de los primeros puntos obtenidos para esta etapa. Las velocidades volumétricas fueron empleadas para calcular las velocidades específicas de producción y consumo en fase estacionaria ($q_{p_{\text{sta}}}$ y $q_{s_{\text{sta}}}$, respectivamente) al dividir las entre la concentración de biomasa promedio.

Extracción de RNA, síntesis de cDNA y análisis por RT-qPCR

Se colectaron muestras de las fermentaciones en modo lote de la cepa AR36 crecida con 50 g/L de Glc y 15 g/L de YE para extracción de RNA en la fase exponencial temprana, EE (2 h ~ 1 OD₆₀₀), fase exponencial tardía, LE (8 h ~ 12 OD₆₀₀), y fase estacionaria, ST (24 h ~ 17 OD₆₀₀) para determinar los niveles de expresión. Para comparación de los datos, también se tomaron y procesaron muestras de cultivos en fase exponencial temprana de la cepa con plásmido vacío AR3e (3.5 h ~ 1 OD₆₀₀), crecida en las mismas condiciones. Se extrajo RNA utilizando fenol caliente equilibrado con agua y la síntesis de cDNA se llevó a cabo usando el kit RevertAid H First Strand cDNA Synthesis (Fermentas, USA) y una mezcla específica de oligonucleótidos, cuyas secuencias se han reportado previamente [52, 54]. Los experimentos de qPCR fueron efectuados con un equipo ABI Prism 7300 Real Time PCR System (Applied Biosystems, USA) usando el kit Maxima SYBRGreen PCR Master Mix (Fermentas, USA) y condiciones de reacción descritas anteriormente [52, 54]. La

técnica de cuantificación usada para comparar los datos fue el método $2^{-\Delta\Delta CT}$ [102] y los resultados se normalizaron usando el gen *ihfB* como control interno. Se ha detectado el mismo nivel de expresión reproducible para este gen en las cepas y condiciones analizadas [54]. Todos los experimentos de qPCR cumplieron los lineamientos MIQE para la publicación de experimentos de PCR cuantitativo en tiempo real [103]. Se extrajo RNA y se llevó a cabo la síntesis correspondiente de cDNA a partir de células obtenidas de dos fermentaciones separadas. Los valores de expresión génica fueron determinados por triplicado para cada muestra y se graficaron los valores promedio. Las barras de error representan la desviación estándar, la cual fue menor al 30% en todos los casos. Los ensayos de qPCR para verificar los genotipos de las cepas construidas (Fig. 5) se llevaron a cabo con 5 ng de DNA templado, plásmidico o cromosomal, extraídos por lisis alcalina y mediante un kit de purificación (UltraClean Microbial DNA Isolation Kit, MO BIO, USA), bajo las condiciones de amplificación reportadas anteriormente [54].

Determinación de metabolitos intracelulares por LC-MS

Medio mínimo P-003 para metabolómica

El medio de cultivo P-003 es un medio muy bajo en fosfatos y sulfatos, compatible con análisis de espectrometría de masas y contiene por litro: $\text{Na}_2\text{HPO}_4\cdot 12\text{H}_2\text{O}$ (3.48 g), KH_2PO_4 (0.61 g), NaCl (0.51 g), NH_4Cl (2.04 g), MgSO_4 (0.098 g), CaCl_2 (4.38 mg), $\text{Na}_2\text{EDTA}\cdot 2\text{H}_2\text{O}$ (15 mg), $\text{ZnSO}_4\cdot 7\text{H}_2\text{O}$ (4.5 mg), $\text{CoCl}_2\cdot 6\text{H}_2\text{O}$ (0.3 mg), $\text{MnCl}_2\cdot 4\text{H}_2\text{O}$ (1 mg), H_3BO_3 (1 mg), $\text{Na}_2\text{MoO}_4\cdot 2\text{H}_2\text{O}$ (0.4 mg), $\text{FeSO}_4\cdot 7\text{H}_2\text{O}$ (3 mg), $\text{CuSO}_4\cdot 5\text{H}_2\text{O}$ (0.3 mg) e hidrocloreuro de tiamina (0.001 g). Además se utilizó una concentración de 3 g/L de Glc como única fuente de carbono, filtrada por separado y añadida al medio después de esterilizar éste por autoclave.

Preparación de células en reposo

La preparación del sistema de células en reposo comprendió los pasos descritos a continuación: 1) se tomó una muestra de células AR3, AR3e y AR36 con un asa estéril a partir de gliceroles almacenados a -70°C y se crecieron pre-inóculos en 3 mL de medio LB con 2 g/L de Glc en presencia de los antibióticos respectivos a 37°C y 180 rpm por 8 horas; 2) a partir de estos tubos se inocularon matraces con fondo plano de 500 mL con 50 mL de

medio de producción de SA suplementado con 25 g/L de Glc, 15 g/L de YE y antibióticos, empleando 300 μ L del cultivo con AR3, 75 μ L del cultivo con AR3e y 100 μ L del cultivo con AR36 (para tomar en cuenta las diferencias en velocidades de crecimiento), a 37°C y 180 rpm por aproximadamente 14 horas; 3) se centrifugó el volumen de células necesario de cada inóculo para comenzar el experimento a 4 OD₆₀₀ (aproximadamente 30 mL) por 4 minutos a 3000 rpm y 22°C. Las células fueron lavadas con 10 mL de medio P-003 a temperatura ambiente y nuevamente centrifugadas para remover completamente el extracto de levadura; 4) para comenzar el experimento, estas células se resuspendieron en 2 mL de medio P-003 y añadidas a matraces baffleados de 250 mL con 25 mL de volumen del mismo medio suplementado con 3 g/L de Glc, donde se monitoreó el consumo de sustrato y la producción de metabolitos. La auxotrofia de las cepas a aminoácidos aromáticos facilitó la preparación de este sistema en medio mínimo sin necesidad de agregar cloranfenicol o limitar nitrógeno, como es requerido en otros casos para inhibir el crecimiento.

Procesamiento de muestras

Se tomaron muestras en los tiempos: 0 h, 0.5 h, 1.5 h y 2.5 h a partir del momento de inoculación, donde se midió la concentración celular para verificar que no hubiera variaciones importantes durante el tiempo del experimento y además se tomó muestra de los sobrenadantes para analizar metabolitos extracelulares. La concentración de Glc en los sobrenadantes fue determinada con un analizador bioquímico YSI 2700 Select (USA) para garantizar un consumo constante desde el inicio del experimento. La preparación de extractos celulares para el análisis de muestras por cromatografía líquida acoplada a espectrometría de masas (LC-MS) consistió en la adición de 120 μ L de medio completo con células a 1 mL de solución de apagado o *quenching* (metanol:acetonitrilo:agua, 40:40:20) a -20°C, reposado por al menos una hora para favorecer una lisis completa. A este tubo se añadió también un volumen de 120 μ L de una solución de metabolitos de referencia, IDMS (isotope dilution mass spectrometry), usada como estándar interno para el análisis de abundancia relativa. Este estándar interno se prepara a partir de un cultivo continuo de *E. coli* MG1655 con glucosa isotópicamente marcada en todos los carbonos (U-¹³C) y extraído en fase exponencial para obtener un perfil de metabolitos propio del crecimiento

balanceado de este organismo [94]. Se almacenaron y analizaron 3 muestras por cada punto en el tiempo de cada una de las dos réplicas biológicas de las 3 cepas en tubos eppendorf de 2 mL. Los tubos con los metabolitos y el IDMS fueron centrifugados brevemente para separar el sobrenadante de los restos celulares y éste se evaporó en un concentrador de vacío (SC110A SpeedVac Plus, ThermoFisher, USA) por aproximadamente 12 horas. Los metabolitos precipitados en cada muestra fueron resuspendidos en 120 μ L de agua ultra-pura y almacenados a -80°C hasta su utilización.

La suspensión de metabolitos fue entonces transferida a viales apropiados y analizada con un método descrito anteriormente [90, 104]. Brevemente, se utilizó cromatografía de intercambio aniónico de alta resolución (Dionex ICS 2000 System, USA) acoplado a un espectrómetro de masas de triple cuadrupolo QTrap 4000 (AB Sciex, USA) y analizado por monitoreo de reacciones múltiples (MRM), con un volumen de inyección de 15 μ L. La identificación de intermediarios metabólicos se hizo de manera manual usando el programa Maven (maven.princeton.edu) [105] con respecto a su relación masa/carga y a los tiempos de retención estandarizados previamente. Posteriormente, se obtuvieron los valores ajustados a la concentración relativa de compuestos presentes en el IDMS de cada muestra [94] y normalizados a la biomasa promedio obtenida en cada experimento. De esta manera, para cada metabolito y cepa se graficaron los valores obtenidos del promedio de todas las muestras tomadas y las barras de error representan la desviación estándar con respecto a los valores obtenidos a lo largo del experimento.

Ensayo para la cuantificación de glucógeno

Se extrajo glucógeno (GLG) de los cultivos con células en reposo empleando una adaptación de métodos previamente reportados [106, 107]. Las células obtenidas a partir de 1 mL de cultivo fueron resuspendidas en 250 μ L de acetato de sodio 50 mM y hervidas a 90°C por 10 minutos. Cada muestra fue sonicada con cuatro sets de pulsos de 30 segundos por 5 minutos y posteriormente centrifugada 20 minutos a velocidad máxima para separar los restos celulares. Se precipitó el GLG existente en el sobrenadante al añadir 2 volúmenes de etanol absoluto y centrifugando 20 minutos en velocidad máxima a 4°C . Se decantó el sobrenadante y se evaporaron los restos de etanol con un concentrador de vacío (SC110A SpeedVac Plus, ThermoFisher, USA). La cuantificación de GLG se llevó a cabo por un

método colorimétrico después de ponerlo a reaccionar con un reactivo de yodo (I_2 0.15 M + KI 1 M) en presencia de altas concentraciones de sal de calcio [108]. Para ello, a cada muestra evaporada se le agregaron 65 μ L de agua, 65 μ L de acetato de sodio 50 mM y 870 μ L de solución calcio-yodo (32.5 mL de $CaCl_2$ 6 M + 125 μ L de reactivo de yodo). Finalmente, se determinó la absorbancia a 460 nm y se calculó la concentración de GLG con base en la siguiente relación: $1 OD_{460nm} = 0.834 \text{ gGLG/L}$, obtenida al ajustar linealmente los valores de absorbancia obtenidos con concentraciones estándar en el intervalo de 0.1 a 1.0 gGLG/L, de manera similar a lo reportado anteriormente [108]. Las concentraciones fueron ajustadas con los valores de biomasa presentes en cada una las muestras.

REFERENCIAS

1. Nie LD, Ding W, Shi XX, Quan N, Lu X: **A novel and high-yielding asymmetric synthesis of oseltamivir phosphate (Tamiflu) starting from (-)-shikimic acid.** *Tetrahedron: Asymmetry* 2012, **23**:742–747.
2. Ward P, Small I, Smith J, Suter P, Dutkowski R: **Oseltamivir (Tamiflu®) and its potential for use in the event of an influenza pandemic.** *J Antimicrob Chemother* 2005, **55**.
3. Farina V, Brown JD: **Tamiflu: the supply problem.** *Angew Chem, Int Ed Engl* 2006, **45**:7330–7334.
4. Rawat G, Tripathi P, Saxena RK: **Expanding horizons of shikimic acid. Recent progresses in production and its endless frontiers in application and market trends.** *Appl Microbiol Biotechnol* 2013, **97**:4277–87.
5. Ghosh S, Chisti Y, Banerjee UC: **Production of shikimic acid.** *Biotechnol Adv* 2012, **30**:1425–31.
6. Escalante A, Calderón R, Valdivia A, de Anda R, Hernández G, Ramírez OT, Gosset G, Bolívar F: **Metabolic engineering for the production of shikimic acid in an evolved *Escherichia coli* strain lacking the phosphoenolpyruvate:carbohydrate phosphotransferase system.** *Microb Cell Fact* 2010, **9**:21.
7. Rodríguez A: **Incorporación de modificaciones genéticas a una cepa de *Escherichia coli* para modular la expresión de genes que permitan acumular shikimato.** Tesis de Maestría, UNAM; 2010.

8. Gutiérrez-Mendoza LM, Schwartz B, Méndez de Lira JDJ, Wirtz VJ: **Oseltamivir storage, distribution and dispensing following the 2009 H1N1 influenza outbreak in Mexico.** *Bull World Health Organ* 2012, **90**:782–7.
9. Rodriguez A, Martínez JA, Flores N, Escalante A, Gosset G, Bolivar F: **Engineering *Escherichia coli* to overproduce aromatic amino acids and derived compounds.** *Microb Cell Fact* 2014, **13**:126.
10. Chandran SS, Yi J, Draths KM, von Daeniken R, Weber W, Frost JW: **Phosphoenolpyruvate availability and the biosynthesis of shikimic acid.** *Biotechnol Prog* 2003, **19**:808–814.
11. Yi J, Li K, Draths KM, Frost JW: **Modulation of phosphoenolpyruvate synthase expression increases shikimate pathway product yields in *E. coli*.** *Biotechnol Prog* 2002, **18**:1141–8.
12. Krämer M, Bongaerts J, Bovenberg R, Kremer S, Müller U, Orf S, Wubbolts M, Raeven L: **Metabolic engineering for microbial production of shikimic acid.** *Metab Eng* 2003, **5**:277–283.
13. Flores N, Xiao J, Berry A, Bolivar F, Valle F: **Pathway engineering for the production of aromatic compounds in *Escherichia coli*.** *Nat Biotechnol* 1996, **14**:620–623.
14. Sprenger GA: **From scratch to value: engineering *Escherichia coli* wild type cells to the production of L-phenylalanine and other fine chemicals derived from chorismate.** *Appl Microbiol Biotechnol* 2007, **75**:739–49.
15. Patnaik R, Spitzer RG, Liao JC: **Pathway engineering for production of aromatics in *Escherichia coli*: confirmation of stoichiometric analysis by independent modulation of AroG, TktA, and Pps activities.** *Biotechnol Bioeng* 1995, **46**:361–370.
16. Meza E, Becker J, Bolivar F, Gosset G, Wittmann C: **Consequences of phosphoenolpyruvate:sugar phosphotransferase system and pyruvate kinase isozymes inactivation in central carbon metabolism flux distribution in *Escherichia coli*.** *Microb Cell Fact* 2012, **11**:127.
17. Postma PW, Lengeler JW, Jacobson GR: **Phosphoenolpyruvate: carbohydrate phosphotransferase systems.** In *Escherichia coli Salmonella Cell Mol Biol*. 2nd edition. Edited by Neidhart FC. ASM Press, USA; 1996:1149–1174.
18. Escalante A, Salinas-Cervantes A, Gosset G, Bolivar F: **Current knowledge of the *Escherichia coli* phosphoenolpyruvate-carbohydrate phosphotransferase system: peculiarities of regulation and impact on growth and product formation.** *Appl Microbiol Biotechnol* 2012.

19. Sprenger G, Siewe R, Martin K, Sonke T: **Microbial Preparation of Substances from Aromatic Metabolism/I. WO patent 98/18936.** 1998.
20. Frost JW., Knop DR: **Biocatalytic Synthesis of Shikimic Acid. US patent 6,472,169 B1.** 2002.
21. Yi J, Draths KM, Li K, Frost JW: **Altered glucose transport and shikimate pathway product yields in *E. coli*.** *Biotechnol Prog* 2003, **19**:1450–9.
22. Balderas-Hernández VE, Sabido-Ramos A, Silva P, Cabrera-Valladares N, Hernández-Chávez G, Baez-Viveros JL, Martínez A, Bolivar F, Gosset G: **Metabolic engineering for improving anthranilate synthesis from glucose in *Escherichia coli*.** *Microb Cell Fact* 2009, **8**:19.
23. Valle F, Mejia N, Berry A: **Application of Glucose Transport Mutants for Production of Aromatic Pathway Compounds. WO patent 96/34961.** 1996.
24. Patnaik R, Liao JC: **Engineering of *Escherichia coli* Central Metabolism for Aromatic Metabolite Production with Near Theoretical Yield.** *Appl Environ Microbiol* 1994, **60**:3903–3908.
25. Liao JC: **Microorganisms and Methods for Overproduction of DAHP by Cloned pps Gene. WO patent 96/08567.** 1996.
26. Liao JC, Hou SY, Chao YP: **Pathway analysis, engineering, and physiological considerations for redirecting central metabolism.** *Biotechnol Bioeng* 1996, **52**:129–40.
27. Gulevich AY, Biryukova IV, Zimenkov DV, Skorokhodova AY, Kivero AD, Belareva AV, Mashko SV: **Method for producing an L-amino acid using a bacterium having enhanced expression of the pckA gene. US Patent 2006/0035348 A1.** 2006.
28. Zhao G, Winkler ME: **An *Escherichia coli* K-12 tktA tktB mutant deficient in transketolase activity requires pyridoxine (vitamin B6) as well as the aromatic amino acids and vitamins for growth.** *J Bacteriol* 1994, **176**:6134–8.
29. Draths KM, Pompliano DL, Conley DL, Frost JW, Berry A, Disbrow GL, Stavarsky RJ, Lievens JC: **Biocatalytic Synthesis of Aromatics from D-Glucose: The Role.** *J Am Chem Soc* 1992, **114**:3956–3962.
30. Baez JL, Bolivar F, Gosset G: **Determination of 3-deoxy-D-arabino-heptulosonate 7-phosphate productivity and yield from glucose in *Escherichia coli* devoid of the glucose phosphotransferase transport system.** *Biotechnol Bioeng* 2001, **73**:530–535.
31. Lu J, Liao JC: **Metabolic engineering and control analysis for production of aromatics: Role of transaldolase.** *Biotechnol Bioeng* 1997, **53**:132–138.

32. Ahn J, Chung BKS, Lee D, Park M, Karimi IA, Jung J, Lee H: **NADPH-dependent pgi-gene knockout *Escherichia coli* metabolism producing shikimate on different carbon sources.** *FEMS Microbiol Lett* 2011, **324**:10–16.
33. Yakandawala N, Romeo T, Friesen AD, Madhyastha S: **Metabolic engineering of *Escherichia coli* to enhance phenylalanine production.** *Appl Microbiol Biotechnol* 2008, **78**:283–291.
34. Li K, Frost JW: **Microbial synthesis of 3-dehydroshikimic acid: a comparative analysis of D-xylose, L-arabinose, and D-glucose carbon sources.** *Biotechnol Prog* 1999, **15**:876–83.
35. Ahn JO, Lee HW, Saha R, Park MS, Jung JK, Lee DY: **Exploring the effects of carbon sources on the metabolic capacity for shikimic acid production in *Escherichia coli* using in silico metabolic predictions.** *J Microbiol Biotechnol* 2008, **18**:1773–1784.
36. Bongaerts J, Krämer M, Müller U, Raeven L, Wubbolts M: **Metabolic engineering for microbial production of aromatic amino acids and derived compounds.** *Metab Eng* 2001, **3**:289–300.
37. Jones KL, Kim SW, Keasling JD: **Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria.** *Metab Eng* 2000, **2**:328–338.
38. Hammer K, Mijakovic I, Jensen PR: **Synthetic promoter libraries-tuning of gene expression.** *Trends Biotechnol* 2006, **24**:53–55.
39. Keasling JD: **Gene-expression tools for the metabolic engineering of bacteria.** *Trends Biotechnol* 1999, **17**:452–460.
40. Silva F, Queiroz JA, Domingues FC: **Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*.** *Biotechnol Adv* 2012, **30**:691–708.
41. Glick BR: **Metabolic load and heterologous gene expression.** *Biotechnol Adv* 1995, **13**:247–261.
42. Diaz-Ricci JC, Hernandez ME: **Plasmid effects on *Escherichia coli* metabolism.** *Crit Rev Biotechnol* 2000, **20**:79–108.
43. Lütke-Eversloh T, Stephanopoulos G: **L-Tyrosine production by deregulated strains of *Escherichia coli*.** *Appl Microbiol Biotechnol* 2007, **75**:103–110.
44. Chen PT, Chiang C, Wang J, Lee M, Chao Y: **Genomic engineering of *Escherichia coli* for production of intermediate metabolites in the aromatic pathway.** *J Taiwan Inst Chem Eng* 2011, **42**:34–40.

45. Sabido A, Martínez LM, de Anda R, Martínez A, Bolívar F, Gosset G: **A novel plasmid vector designed for chromosomal gene integration and expression: Use for developing a genetically stable *Escherichia coli* melanin production strain.** *Plasmid* 2012, **69**:16–23.
46. Tyo KEJ, Ajikumar PK, Stephanopoulos G: **Stabilized gene duplication enables long-term selection-free heterologous pathway expression.** *Nat Biotech* 2009, **27**:760–765.
47. Cui Y-Y, Ling C, Zhang Y-Y, Huang J, Liu J-Z: **Production of shikimic acid from *Escherichia coli* through chemically inducible chromosomal evolution and cofactor metabolic engineering.** *Microb Cell Fact* 2014, **13**:21.
48. Alper H, Fischer C, Nevoigt E, Stephanopoulos G: **Tuning genetic control through promoter engineering.** *Proc Natl Acad Sci U S A* 2005, **102**:12678–83.
49. Gosset G: **Improvement of *Escherichia coli* production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system.** *Microb Cell Fact* 2005, **4**:14.
50. Sprenger GA: **Aromatic amino acids.** *Microbiol Monogr* 2006, **5**:93–127.
51. Flores N, Leal L, Sigala JC, de Anda R, Escalante A, Martínez A, Ramírez OT, Gosset G, Bolivar F: **Growth recovery on glucose under aerobic conditions of an *Escherichia coli* strain carrying a phosphoenolpyruvate: carbohydrate phosphotransferase system deletion by inactivating *arcA* and overexpressing the genes coding for glucokinase and galactose.** *J Mol Microbiol Biotechnol* 2007, **13**:105–116.
52. Flores N, Flores S, Escalante A, de Anda R, Leal L, Malpica R, Georgellis D, Gosset G, Bolivar F: **Adaptation for fast growth on glucose by differential expression of central carbon metabolism and gal regulon genes in an *Escherichia coli* strain lacking the phosphoenolpyruvate:carbohydrate phosphotransferase system.** *Metab Eng* 2005, **7**:70–87.
53. Martínez K, de Anda R, Hernández G, Escalante A, Gosset G, Ramírez OT, Bolivar F: **Couitilization of glucose and glycerol enhances the production of aromatic compounds in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.** *Microb Cell Fact* 2008, **7**:1.
54. Aguilar C, Escalante A, Flores N, de Anda R, Riveros-McKay F, Gosset G, Morett E, Bolivar F: **Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system.** *BMC Genomics* 2012, **13**:385.
55. Kedar P, Colah R, Shimizu K: **Proteomic investigation on the *pyk-F* gene knockout *Escherichia coli* for aromatic amino acid production.** *Enzym Microb Technol* 2007, **41**:455–465.

56. Gosset G, Yong-Xiao J, Berry A: **A direct comparison of approaches for increasing carbon flow to aromatic biosynthesis in *Escherichia coli***. *J Ind Microbiol* 1996, **17**:47–52.
57. Sabido A, Sigala JC, Hernández-Chávez G, Flores N, Gosset G, Bolívar F: **Physiological and transcriptional characterization of *Escherichia coli* strains lacking interconversion of phosphoenolpyruvate and pyruvate when glucose and acetate are coutilized**. *Biotechnol Bioeng* 2014, **111**:1150–1160.
58. Al Zaid Siddiquee K, Arauzo-Bravo MJ, Shimizu K: **Metabolic flux analysis of *pykF* gene knockout *Escherichia coli* based on 13 C-labeling experiments together with measurements of enzyme activities and intracellular metabolite concentrations**. *Appl Microbiol Biotechnol* 2004, **63**:407–417.
59. Blangy D, Buc H, Monod J: **Kinetics of the allosteric interactions of phosphofructokinase from *Escherichia coli***. *J Mol Biol* 1968, **31**:13–35.
60. Romeo T, Snoep JL: **Glycolysis and Flux Control**. *EcoSal Plus* 2005, **1**.
61. Johansson L, Lindskog A, Silfversparre G, Cimander C, Nielsen KF, Lidén G: **Shikimic acid production by a modified strain of *E. coli* (W3110.shik1) under phosphate-limited and carbon-limited conditions**. *Biotechnol Bioeng* 2005, **92**:541–552.
62. Amann E, Ochs B, Abel K: **Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli***. *Gene* 1988, **69**:301–315.
63. Zurita M, Bolivar F, Soberon X: **Construction and characterization of new cloning vehicles VII. Construction of plasmid pBR327par, a completely sequenced, stable derivative of pBR327 containing the par locus of PSC101**. *Gene* 1984, **28**:119–122.
64. Balbás P, Soberón X, Merino E, Zurita M, Lomelí H, Valle F, Flores N, Bolívar F: **Plasmid vector pBR322 and its special-purpose derivatives - a review**. *Gene* 1986, **50**:3–40.
65. Hoischen C, Bussiek M, Langowski J, Diekmann S: ***Escherichia coli* low-copy-number plasmid R1 centromere parC forms a U-shaped complex with its binding protein ParR**. *Nucleic Acids Res* 2008, **36**:607–615.
66. Okibe N, Suzuki N, Inui M, Yukawa H: **pCGR2 copy number depends on the par locus that forms a ParC-ParB-DNA partition complex in *Corynebacterium glutamicum***. *J Appl Microbiol* 2013, **115**:495–508.
67. Wilson CJ, Zhan H, Swint-Kruse L, Matthews KS: **The lactose repressor system: Paradigms for regulation, allosteric behavior and protein folding**. *Cell Mol Life Sci* 2007, **64**:3–16.

68. Boer H de, Comstock LJ, Vasser M: **The *tac* promoter: A functional hybrid derived from the *trp* and *lac* promoters.** *Proc Natl Acad Sci U S A* 1983, **80**(January):21–25.
69. Le Borgne S, Palmeros B, Valle F, Bolivar F, Gosset G: **pBRINT-Ts: a plasmid family with a temperature-sensitive replicon, designed for chromosomal integration into the *lacZ* gene of *Escherichia coli*.** *Gene* 1998, **223**:213–219.
70. Juminaga D, Baidoo EEK, Redding-Johanson AM, Batth TS, Burd H, Mukhopadhyay A, Petzold CJ, Keasling JD: **Modular engineering of L-tyrosine production in *Escherichia coli*.** *Appl Env Microbiol* 2012, **78**:89–98.
71. Knop DR, Draths KM, Chandran SS, Barker JL, von Daeniken R, Weber W, Frost JW: **Hydroaromatic equilibration during biosynthesis of shikimic acid.** *J Am Chem Soc* 2001, **123**:10173–10182.
72. Johansson L, Lidén G: **Transcriptome analysis of a shikimic acid producing strain of *Escherichia coli* W3110 grown under carbon- and phosphate-limited conditions.** *J Biotechnol* 2006, **126**:528–545.
73. Flores S, de Anda-Herrera R, Gosset G, Bolívar F.: **Growth-rate recovery of *Escherichia coli* cultures carrying a multicopy plasmid, by engineering of the pentose-phosphate pathway.** *Biotechnol Bioeng* 2004, **87**:485–494.
74. Lara AR, Caspeta L, Gosset G, Bolivar F, Ramírez OT: **Utility of an *Escherichia coli* strain engineered in the substrate uptake system for improved culture performance at high glucose and cell concentrations : An alternative to fed-batch cultures.** *Biotechnol Bioeng* 2008, **99**:893–901.
75. Wolfe AJ: **The acetate switch.** *Microbiol Mol Biol Rev* 2005, **69**:12–50.
76. Chen X, Li M, Zhou L, Shen W, Algasan G, Fan Y, Wang Z: **Metabolic engineering of *Escherichia coli* for improving shikimate synthesis from glucose.** *Bioresour Technol* 2014, **166**:64–71.
77. Michel G, Roszak AW, Sauvé V, Maclean J, Matte A, Coggins JR, Cygler M, Laphorn AJ: **Structures of shikimate dehydrogenase AroE and its paralog YdiB.** *J Biol Chem* 2003, **278**:19463–19472.
78. Lütke-Eversloh T, Stephanopoulos G: **Combinatorial pathway analysis for improved L-tyrosine production in *Escherichia coli*: identification of enzymatic bottlenecks by systematic gene overexpression.** *Metab Eng* 2008, **10**:69–77.
79. Chen K, Dou J, Tang S, Yang Y, Wang H, Fang H, Zhou C: **Deletion of the *aroK* gene is essential for high shikimic acid accumulation through the shikimate pathway in *E. coli*.** *Bioresour Technol* 2012, **119**:141–7.

80. Cortés-tolalpa L, Gutiérrez-ríos RM, Martínez LM, Anda R De, Gosset G, Bolívar F, Escalante A: **Global transcriptomic analysis of an engineered *Escherichia coli* strain lacking the phosphoenolpyruvate : carbohydrate phosphotransferase system during shikimic acid production in rich culture medium.** 2014:1–20.
81. Sigala JC, Flores S, Flores N, Aguilar C, Anda R de, Gosset G, Bolívar F: **Acetate metabolism in *Escherichia coli* strains lacking phosphoenolpyruvate: carbohydrate phosphotransferase system; evidence of carbon recycling strategies and futile cycles.** *J Mol Microbiol Biotechnol* 2009, **16**:224–235.
82. Ponce E, Flores N, Martinez A, Valle F, Bolívar F: **Cloning of the two pyruvate kinase isoenzyme structural genes from *Escherichia coli*: the relative roles of these enzymes in pyruvate biosynthesis.** *J Bacteriol* 1995, **177**:5719–22.
83. Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martinez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muniz-Rascado L, Ong Q, Paley S, Schroder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus RP, Paulsen I, Karp PD: **EcoCyc: fusing model organism databases with systems biology.** *Nucleic Acids Res* 2013, **41**:D605–D612.
84. Peng L, Arauzo-Bravo MJ, Shimizu K: **Metabolic flux analysis for a ppc mutant *Escherichia coli* based on ¹³C-labelling experiments together with enzyme activity assays and intracellular metabolite measurements.** *FEMS Microbiol Lett* 2004, **235**:17–23.
85. Yun-Peng Chao, Liao JC: **Alteration of growth yield by overexpression of phosphoenolpyruvate carboxylase and phosphoenolpyruvate carboxykinase in *Escherichia coli*.** *Appl Environ Microbiol* 1993, **59**:5.
86. Peng L, Shimizu K: **Global metabolic regulation analysis for *Escherichia coli* K12 based on protein expression by 2-dimensional electrophoresis and enzyme activity measurement.** *Appl Microbiol Biotechnol* 2003, **61**:163–178.
87. Sauer U, Eikmanns BJ: **The PEP – pyruvate – oxaloacetate node as the switch point for carbon flux distribution in bacteria.** *FEMS Microbiol Rev* 2005, **29**:765–794.
88. Ralston A: **Simultaneous gene transcription and translation in bacteria.** *Nat Educ* 2008, **1**.
89. Ellis DI, Goodacre R: **Metabolomics-assisted synthetic biology.** *Curr Opin Biotechnol* 2012, **23**:22–8.
90. Kiefer P, Nicolas C, Letisse F, Portais JC: **Determination of carbon labeling distribution of intracellular metabolites from single fragment ions by ion chromatography tandem mass spectrometry.** *Anal Biochem* 2007, **360**:182–8.

91. Diaz-Ricci JC, Bode J, Rhee JIL, Schugerl K: **Gene expression enhancement due to plasmid maintenance.** *J Bacteriol* 1995, **177**.
92. Wang Z, Xiang L, Shao J, Wegrzyn A, Wegrzyn G: **Effects of the presence of ColE1 plasmid DNA in *Escherichia coli* on the host cell metabolism.** *Microb Cell Fact* 2006, **5**:34.
93. Salgado H, Peralta-Gil M, Gama-Castro S, Santos-Zavaleta A, Muñiz-Rascado L, García-Sotelo JS, Weiss V, Solano-Lira H, Martínez-Flores I, Medina-Rivera A, Salgado-Osorio G, Alquicira-Hernández S, Alquicira-Hernández K, López-Fuentes A, Porrón-Sotelo L, Huerta AM, Bonavides-Martínez C, Balderas-Martínez YI, Pannier L, Olvera M, Labastida A, Jiménez-Jacinto V, Vega-Alvarado L, Del Moral-Chávez V, Hernández-Alvarez A, Morett E, Collado-Vides J: **RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more.** *Nucleic Acids Res* 2013, **41**(Database issue):D203–13.
94. Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, van Winden W a, van Gulik WM, Heijnen JJ: **Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ¹³C-labeled cell extracts as internal standards.** *Anal Biochem* 2005, **336**:164–71.
95. Chapman AG, Fall L, Atkinson DE: **Adenylate energy charge in *Escherichia coli* during growth and starvation.** *J Bacteriol* 1971, **108**:1072–1086.
96. Carneiro S, Ferreira EC, Rocha I: **Metabolic responses to recombinant bioprocesses in *Escherichia coli*.** *J Biotechnol* 2013, **164**:396–408.
97. Flores S, Gosset G, Flores N, De Graaf AA, Bolivar F: **Analysis of carbon metabolism in *Escherichia coli* strains with an inactive phosphotransferase system by (¹³C) labeling and NMR spectroscopy.** *Metab Eng* 2002, **4**:124–137.
98. Messing J: **Recombinant DNA technical bulletin.** *NIH Publ* 1979, **2**:43–48.
99. Datsenko K, Wanner B: **One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.** *Proc Natl Acad Sci USA* 2000, **97**:6640–6645.
100. Cherepanov PP, Wackernagel W: **Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of FIp-catalyzed excision of the antibiotic-resistance determinant.** *Gene* 1995, **158**:9–14.
101. Weissbach A, Hurwitz J: **The formation of 2-Keto-3-deoxyheptonic acid in extracts of *Escherichia coli* B.** *J Biol Chem* 1958, **234**:705–709.
102. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2(- $\Delta\Delta$ CT) method.** *Methods* 2001, **25**:402–408.

103. Bustin SA, Benes V, Garson JA, Hellemans J, Hugget J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: **The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments.** *Clin Chem* 2009, **55**:611–622.
104. Bolten CJ, Kiefer P, Letisse F, Portais JC, Wittmann C: **Sampling for metabolome analysis of microorganisms.** *Anal Chem* 2007, **79**:3843–3849.
105. Clasquin MF, Melamud E, Rabinowitz JD: **LC-MS data processing with MAVEN: a metabolomic analysis and visualization engine.** *Curr Protoc Bioinforma* 2012, **Chapter 14**.
106. Shim J-H, Park J-T, Hong J-S, Kim KW, Kim M-J, Auh J-H, Kim Y-W, Park C-S, Boos W, Kim J-W, Park K-H: **Role of maltogenic amylase and pullulanase in maltodextrin and glycogen metabolism of *Bacillus subtilis* 168.** *J Bacteriol* 2009, **191**:4835–44.
107. Park J-T, Shim J-H, Tran PL, Hong I-H, Yong H-U, Oktavina EF, Nguyen HD, Kim J-W, Lee TS, Park S-H, Boos W, Park K-H: **Role of maltose enzymes in glycogen synthesis by *Escherichia coli*.** *J Bacteriol* 2011, **193**:2517–26.
108. Krisman C: **A method for the colorimetric estimation of glycogen with iodine.** *Anal Biochem* 1962, **23**:17–23.

MATERIAL ADICIONAL 1

Secuencia nucleotídica del operón sintético construido en este trabajo y presente en el plásmido pTrcAro6. El gen *aroG^{fbr}* incluido en esta construcción fue un regalo de DuPont-Genencor® por lo que su secuencia no puede ser divulgada. Cada nucleótido de *aroG^{fbr}* está indicado con una “n” excepto por los que corresponden a sus codones de inicio y término.

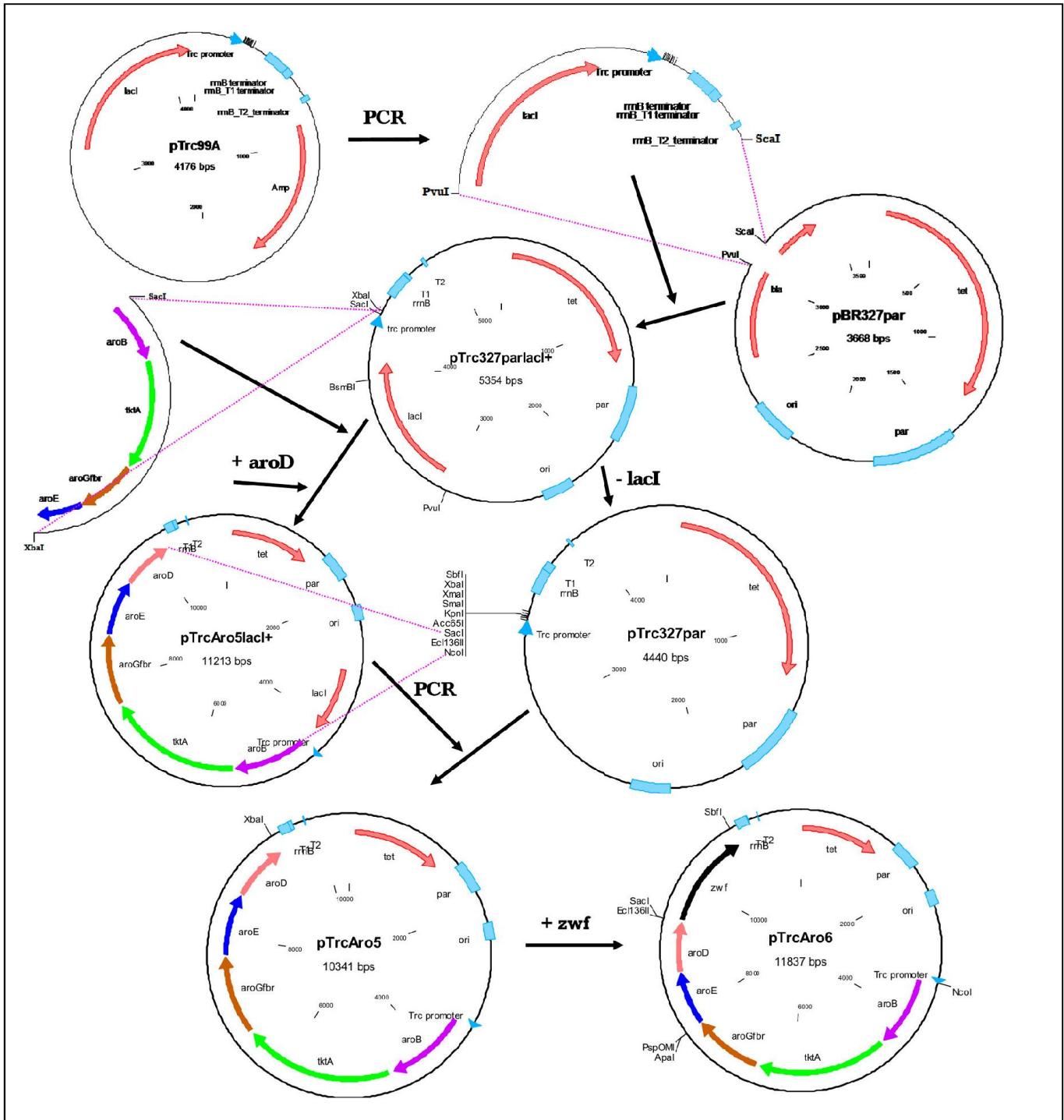
1	gccgacatca	taacggttct	ggcaaatatt	ctgaaatgag	ctggtgacaa	ttaatcatcc
61	ggctcgtata	atgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	acagaccatg
121	gaggaggtgc	tgctatggag	aggattgtcg	ttactctcgg	ggaacgtagt	tacccaatta
181	ccatcgcatc	tggtttgttt	aatgaaccag	cttcattctt	accgctgaaa	tcgggcgagc
241	aggatcatgtt	ggtcaccaac	gaaaccctgg	ctcctctgta	tctcgataag	gtccgcggcg
301	tacttgaaca	ggcgggtgtt	aacgtcgata	gcgttatcct	ccctgacggc	gagcagtata
361	aaagcctggc	tgtactcgat	accgtcttta	cggcgttggt	acaaaaaccg	catggctcgcg
421	atactacgct	ggtggcgctt	ggcggcgggc	tagtgggcca	tctgaccggc	ttcgcgggcg
481	cgagttatca	gcgcggtgtc	cgtttcattc	aagtcccgac	gacgttactg	tcgcaggtcg
541	attcctccgt	tggcggcaaa	actgcggcoa	accatcccct	cggtaaaaac	atgattggcg
601	cgttctacca	acctgcttca	gtgggtggtg	atctcgactg	tctgaaaacg	cttccccgcg
661	gtgagttagc	gtcggggctg	gcagaagtea	tcaaatacgg	cattatctct	gacgggtcgt
721	tttttaactg	gctggaagag	aatctggatg	cgttggtgcg	tctggacggg	ccggcaatgg
781	cgtactgtat	tcgccgttgg	tgtgaactga	aggcagaagt	tgtcgcggcc	gacgagcggc
841	aaaccgggtt	acgtgcttta	ctgaactctg	gacacacctt	tggtcatgcc	attggaagctg
901	aaatggggta	tggcaattgg	ttacatggtg	aagcggctcg	tgcgggtatg	gtgatggcgg
961	cgcggacgtc	ggaacgtctc	gggcagttta	gttctgcca	aacgcagcgt	attataacc
1021	tgctcaagcg	ggctgggtta	ccggtcaatg	ggcgcgcgca	aatgtccggc	cagggcgtatt
1081	taccgcatat	gctgcgtgac	aagaaagtcc	ttgcgggaga	gatgcgctta	attcttccgt
1141	tggcaattgg	taagagtga	gttcgcagcg	gcgtttcgca	cgagcttggt	cttaaccgca
1201	ttgccgattg	tcaatcagcg	taagtcaagt	cagcgtcagg	aggggctgca	ggaattcgat
1261	gctacgatct	agctaggagg	gagtcaaat	gtcctcacgt	aaagagcttg	ccaatgctat
1321	tcgtgcgctg	agcatggacg	cagtacagaa	agccaaatcc	ggtcaccccg	gtgccctat
1381	gggtatggct	gacattgccg	aagtccctgtg	cgctgatttc	ctgaaacaca	accgcgagaa
1441	tcctgctcgg	gctgaccgtg	accgcttcgt	gctgtccaac	ggccaacggc	ccatgctgat
1501	ctacagcctg	ctgcacctca	ccggttacaa	tctgccgatg	gaagaactga	aaaacttccg
1561	tcagctgcac	tctaaaactc	cgggtcacc	ggaagtgggt	tacaccgctg	gtgtggaaac
1621	caccaccggt	ccgtcgggtc	agggtattgc	caacgcagtc	ggatggcgca	ttgcagaaaa
1681	aacgctggcg	gcgcagttta	accgtccggg	ccacggcatt	gtcgaccact	acacctacgc
1741	cttcatgggc	gacggctgca	tgatggaagg	catctcccac	gaagtttget	ctctggcggg
1801	tacgctgaag	ctgggtaaac	tgattgcatt	ctacgatgac	aacggtatct	ctatcgatgg
1861	tcacgttgaa	ggctggttca	ccgacgacac	cgcaatgcgt	ttcgaagctt	acggctggca
1921	cgttattcgc	gacatcgacg	gtcatgacgc	ggcatctatc	aaacgcgcag	tagaagaagc
1981	gcgcgcagtg	actgacaaa	cttccctgct	gatgtgcaaa	accatcatcg	gtttcggttc
2041	cccgaacaaa	gccggtacc	acgactccca	cggtgcccg	ctgggocgac	ctgaaattgc
2101	cctgaccgcg	gaacaaactg	gctggaata	tgcgcggtt	gaaatcccg	ctgaaatcta
2161	tgctcagtg	gatgcgaaag	aagcaggcca	ggcgaagaa	tccgcatgga	acgagaaatt
2221	cgctgcttac	gcgaaagctt	atccgcagga	agccgctgaa	tttaccgcc	gtatgaaagg
2281	cgaaatgccc	tctgacttcc	acgctaaagc	gaaagagttc	atcgctaaac	tgcaggctaa
2341	tccggcgaaa	atcgccagcc	gtaaaagcgc	tcagaatgct	atcgaagcgt	tcggctccgct
2401	gttgccggaa	ttcctcggcg	gttctgctga	cctggcgcgg	tctaactga	ccctgtggtc
2461	tggttctaaa	gcaatcaacg	aagatgctgc	gggtaactac	atcactacg	gtgttcgca
2521	gttcggatag	accgcgattg	ctaacggtat	ctccctgcac	gggtggttcc	tgcggtacac
2581	ctccaccttc	ctgatgttcc	tggaaatcgc	acgtaaccgc	gtacgtatgg	ctgcgctgat
2641	gaaacagcgt	caggtgacgg	tttacaccca	cgactccatc	ggctcggggc	aagacggccc
2701	gactcaccag	ccggttgagc	aggctccttc	tctgcgcgta	accccgaaca	tgtctacatg
2761	gcgtccgtgt	gaccaggttg	aatccgcgg	cgcgtggaaa	tacgggtgtg	agcgtcagga
2821	cggcccagcc	gactgatcc	tctcccgtca	gaacctggcg	cagcaggaac	gaactgaaga
2881	gcaactggca	aacatcgcgc	gcgggtggtta	tgtgctgaaa	gactgcgcgg	gtcagccgga
2941	actgattttc	atcgctaccg	gttcagaagt	tgaactggct	ggtgctgctc	acgaaaaact
3001	gactgccgaa	ggcgtgaaag	cgcgcgtgg	gtccatgccc	tctaccgacg	catttgacaa
3061	gcaggatgct	gcttaccgtg	aatccgtact	gccgaaagcg	gttactgcac	gcgttgctgt
3121	agaagcgggt	attgctgact	actggtacaa	gtatggtggc	ctgaacggtg	ctatcgctcg
3181	tatgaccacc	ttcgggtgaat	ctgctccggc	agagctgctg	tttgaagagt	tcggcttcac
3241	tgttgataac	gttggtgcca	aagcaaaaga	actgctgtaa	catactgagt	tcgtaggagg
3301	atcaagctta	tcgataccgt	cgacctcgac	aggagggaac	agacatgann	nnnnnnnnnn
3361	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3421	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn

3481	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3541	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3601	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3661	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3721	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3781	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3841	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3901	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3961	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4021	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4081	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4141	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4201	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4261	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4321	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4381	nnnnnnnnnn	nnnntaagtc	gagggggggc	ccaggaggca	gataatggaa	acctatgctg
4441	tttttggtaa	tccgatagcc	cacagcaaat	cgccattcat	tcatcagcaa	tttgctcagc
4501	aactgactat	tgaacatccc	tatgggcgcg	tgttggcacc	catcaatgat	ttcatcaaca
4561	cactaaacgc	tttctttagt	gctggtggtg	aagggtcgaa	tgtgacggtg	ccttttaaag
4621	aagaggcttt	tgccagagcg	gatgagctta	ctgaacgggc	agcgttggct	ggtgctgtta
4681	ataccctcat	gcggttagaa	gatggacgcc	tgctgggtga	caataccgat	ggtgtaggct
4741	tgttaagcga	tctggaacgt	ctgtctttta	tccgccctgy	tttactgtatt	ctgcatatcg
4801	gcgctggtgg	agcatctcgc	ggcgtaacta	tgccactcct	ttccctggac	tgtgcggtga
4861	caataactaa	tccgacggta	tcccgcgcgg	aagagttgcy	taaattgttt	gcgcacactg
4921	gcagtatcca	ggcgttgagt	atggacgaac	tggaaaggtca	tgagtttgat	ctcattatta
4981	atgcaacatc	cagtggcatc	agtggtgata	ttccggcgat	cccgtcatcg	ctcattcatc
5041	caggcattta	ttgctatgac	atgttctatc	agaaaggaaa	aactcctttt	ctggcatggt
5101	gtgagcagcg	aggctcaaag	cgtaatgctg	atggttttagg	aatgctggtg	gcacaggcgg
5161	ctcatgcctt	tcttctctgg	cacggtgttc	tgctgacgt	agaaccagtt	ataaagcaat
5221	tgcaggagga	attgtccgcy	tgaatcagta	tctagcagga	ggcagataat	gaaaaccgta
5281	actgtaaaa	atctcgatc	tggtaacggc	gcaacctaaa	tcatcgctc	gctgatggcg
5341	aaagatatcg	ccagcgtgaa	atccgaagct	ctcgcctatc	gtgaagcggg	ccttgatatt
5401	ctggaatggc	gtgtggacca	ctatgccgac	ctctccaatg	tggagtctgt	catggcgcca
5461	gcaaaaatc	tcogtgagac	catgccagaa	aaaccgctgc	tgtttacct	ccgcagtgcc
5521	aaagaaggcg	gcgagcaggc	gatctccacc	gaggcttata	ttgcactcaa	tcgtgcagcc
5581	atcgacagcg	gcctgggtga	tatgatcgat	ctggagttat	ttaccgggta	tgatcaggtt
5641	aaagaaaccg	tcgctacgc	ccacgcgc	gatgtgaaag	tagtcatgtc	caaccatgac
5701	ttccataaaa	cgccggaagc	cgaagaaatc	attgccctgc	tgcgcataat	gcaatccttc
5761	gacgcccata	ttcctaagat	tgcgctgatg	ccgcaaagta	ccagcgatgt	gctgacgttg
5821	cttgccgcga	cctggagat	gcaggagcag	tatgccgatc	gtccaattat	cacgatgtcg
5881	atggcaaaaa	ctggcgtaat	ttctcgtctg	gctgggtgaag	tatttggctc	ggcggcaact
5941	tttgggtcgg	taaaaaaagc	gtctgcgcca	gggcaaatct	cggtaaatga	tttgcgcacg
6001	gtattaacta	ttttacacca	ggcataagag	ctcggtaacc	ggggatcctc	tagaaggagg
6061	ctcgcactat	ggcggtaacg	caaacagccc	aggcctgtga	cctggtcatt	ttcggcgcca
6121	aaggcgacct	tgcgctcgt	aaattgctgc	cttccctgta	tcaactggaa	aaagccggtc
6181	agctcaaccc	ggacacccgg	attatcggcg	tagggcgtgc	tgactgggat	aaagcggcat
6241	ataccaaaagt	tgtccgcgag	gcgctcgaaa	ctttcatgaa	agaaaccatt	gatgaaggtt
6301	tatgggacac	cctgagtgca	cgtctggatt	tttgtaatct	cgatgtcaat	gacactgctg
6361	cattcagccg	tctcggcgcg	atgetggatc	aaaaaaatcg	tatcaccatt	aactactttg
6421	ccatgcccg	cagcactttt	ggcgcaattt	gcaaagggtc	tggcgaggca	aaactgaatg
6481	ctaaaccggc	acgcgtagtc	atggagaaac	cgctggggac	gtcgtggcg	acctcgacgg
6541	aaatcaatga	tcaggttggc	gaatacttcg	aggagtgcc	ggtttaccgt	atcgaccact
6601	atcttggtaa	agaaacggtg	ctgaacctgt	tggcgctcgy	ttttgctaac	tccctgtttg
6661	tgaataactg	ggacaatcgc	accattgatc	atggttgagat	taccgtggca	gaagaagtgg
6721	ggatcgaagg	gcgctggggc	tattttgata	aagccggctca	gatgcgcgac	atgatccaga
6781	accacctgct	gcaaatctt	tgcgatattg	cgatgtctc	gccgtctgac	ctgagcgacg
6841	acagcatccg	cgatgaaaaa	gtgaaagtac	tgaagtctct	gcgcgcgac	gaccgctcca
6901	acgtacgcga	aaaaaccgta	cgcgggcaat	atactgcggg	cttcgcccag	ggcaaaaaag

6961	tgccgggata	tctggaagaa	gagggcgcga	acaagagcag	caatacagaa	actttcgtgg
7021	cgatccgcgt	cgacattgat	aactggcgct	gggccgggtgt	gccattctac	ctgcgtactg
7081	gtaaactgct	gccgaccaa	tgttctgaag	tcgtgggtcta	tttcaaaaca	cctgaactga
7141	atctgtttaa	agaatcgtgg	caggatctgc	cgcagaataa	actgactatc	cgtctgcaac
7201	ctgatgaagg	cgtggatata	caggtactga	ataaagttcc	tggccttgac	cacaaacata
7261	acctgcaaat	caccaagctg	gatctgagct	atcagaaac	ctttaatcag	acgcactctg
7321	cggatgecta	tgaacgtttg	ctgctggaaa	ccatgcgtgg	tattcaggca	ctgtttgtac
7381	gtcgcgacga	agtggaagaa	gcctggaaat	gggtagactc	cattactgag	gcgtgggcga
7441	tggacaatga	tgcgccgaaa	ccgtatcagg	ccggaacctg	gggacctgtt	gcctcgggtg
7501	cgatgattac	ccgtgatggt	cgttcctgga	atgagtttga	gtaatctaga	gtcgacctgc
7561	aggcatgcaa	gcttggctgt	tttggcggat	gagagaagat	tttcagcctg	atacagatta
7621	aatcagaacg	cagaagcggg	ctgataaaa	agaatttgcc	tggcggcagt	agcgcgggtg
7681	tcccacctga	ccccatgccg	aactcagaag	tgaaacgccg	tagcggccgat	ggtagtgtgg
7741	ggtctcccca	tgcgagagta	gggaactgcc	aggcatcaaa	taaaacgaaa	ggctcagtcg
7801	aaagactggg	cctttcgttt	tatctgttgt	ttgtcgggtga	acgctctcct	gagtaggaca
7861	aatccgccgg	gagcggattt	gaacgttgcg	aagcaacggc	ccggaggggtg	gcgggcagga
7921	cgcccgccat	aaactgccag	gcatacaatt	aagcagaagg	ccatcctgac	ggatggactc
7981	aaccaagtca	ttctgagaat	agtgtatgcg	gcgaccgagt	tg	

MATERIAL ADICIONAL 2

Esquema detallado de la construcción del plásmido pTrcAro6. La secuencia de reacciones se describe en Materiales y métodos.



MATERIAL ADICIONAL 3

Valores de transcripción relativa obtenidos por qPCR, correspondientes al promedio y desviación estándar de tres mediciones hechas para cada una de las dos réplicas biológicas.

gene	1 OD (EE)	12 OD (LE)	15 OD (ST)
aceB	0.10 + 0.01	1.06 + 0.01	1.04 + 0.14
aceE	3.01 + 0.77	2.19 + 0.15	0.39 + 0.05
aceF	2.76 + 0.80	1.73 + 0.16	0.23 + 0.02
ackA	0.82 + 0.07	0.78 + 0.16	0.66 + 0.11
acnA	1.48 + 0.53	2.66 + 0.49	1.00 + 0.19
acs	2.97 + 0.39	10.92 + 2.54	0.34 + 0.03
actP	3.91 + 0.28	16.86 + 4.01	1.70 + 0.18
aroB	361 + 28.20	209. + 46.3	209.99 + 18.11
aroD	1707.74 + 294.48	1242.03 + 228.13	1516.53 + 220.75
aroE	214.94 + 58.06	174.27 + 9.40	135.33 + 6.87
aroG	109.67 + 17.99	62.59 + 7.13	31.62 + 2.05
aroK	0	0	0
eda	1.29 + 0.38	0.84 + 0.18	0.17 + 0.01
edd	1.81 + 0.43	0.79 + 0.12	0.32 + 0.07
eno	2.67 + 0.46	2.29 + 0.60	
fbaA	1.95 + 0.42	1.70 + 0.07	0.17 + 0.01
fumC	0.73 + 0.02	0.48 + 0.02	0.10 + 0.02
galP	2.75 + 0.10	2.43 + 0.15	0.41 + 0.01
gapA	4.04 + 0.83	2.11 + 0.43	0.16 + 0.04
glk	3.24 + 0.07	2.91 + 0.03	0.57 + 0.01
gltA	0.71 + 0.08	0.86 + 0.12	0.15 + 0.03
gnd	2.18 + 0.73	0.96 + 0.20	0.34 + 0.01
icdA	2.81 + 1.41	2.11 + 0.23	0.21 + 0.02
lacl	0	0	0
lpdA	N/D	2.04 + 0.02	0.44 + 0.06
maeB	1.23 + 0.19	0.90 + 0.07	0.10 + 0.02
mdh	1.42 + 0.02	2.05 + 0.04	0.17 + 0.01
pckA	6.22 + 0.79	3.74 + 0.33	0.24 + 0.05
pfkA	1.92 + 0.48	1.87 + 0.25	1.26 + 0.34
pgi	3.34 + 0.70	3.18 + 0.85	0.25 + 0.03
pgl	0.85 + 0.11	0.98 + 0.06	0.28 + 0.02
poxB	0.77 + 0.00	1.05 + 0.24	0.22 + 0.01
ppc	0.36 + 0.05	0.31 + 0.03	0.04 + 0.00
ppsA	0.66 + 0.21	0.84 + 0.11	0.19 + 0.01
pta	0.99 + 0.03	0.60 + 0.06	0.26 + 0.01
ptsH	0	0	0
pykA	0.93 + 0.02	0.99 + 0.15	0.31 + 0.02
pykF	0	0	0
sdhA	2.00 + 0.43	2.34 + 0.13	0.07 + 0.01
sdhC	1.53 + 0.45	1.75 + 0.06	0.07 + 0.00
sucA	4.80 + 1.32	4.28 + 0.36	0.12 + 0.00
sucC	4.21 + 1.30	4.76 + 0.58	0.11 + 0.02
talA	1.24 + 0.26	1.95 + 0.32	0.64 + 0.07
tktA	232.58 + 21.75	106.56 + 18.71	62.07 + 4.26
zwf	24.09 + 1.00	11.24 + 1.37	14.72 + 1.65

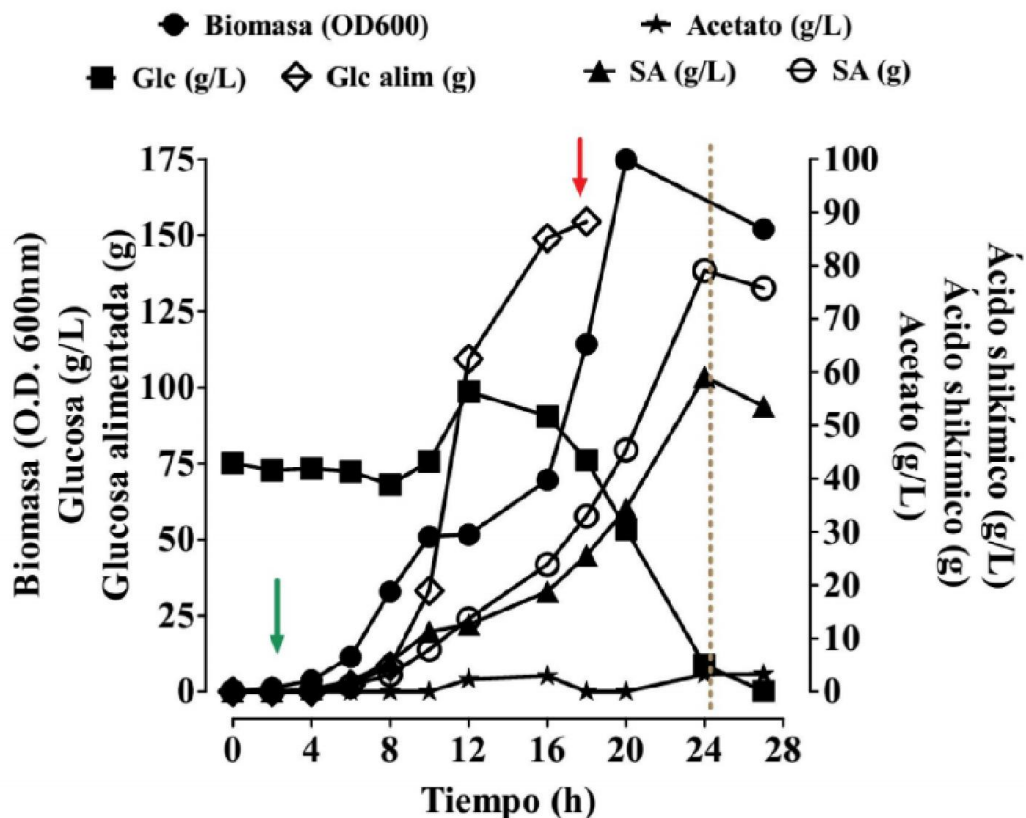
MATERIAL ADICIONAL 4

Oligonucleótidos utilizados en este trabajo. Las secuencias de oligonucleótidos empleados para los estudios de qPCR han sido previamente reportadas [52, 54].

Nombre (bp)	Secuencia (5'- 3')	Función
aroKcatFw (67)	AATAGTCTTAGTAGTACCGAAAAAATGGCAGAGAAACGCAATA TCTGTGTAGGCTGGAGCTGCTTCG	Amplificación del gen cat de pDK3 con región homóloga para aroK
aroKcatRv (65)	AGACGAGTGTATATAAAGCCAGAAATTAGTTGCTTTCCAGCATGT GATGGGAATTAGCCATGGTCC	Amplificación del gen cat de pDK3 con región homóloga para aroK
aroLcatFw (66)	GCGACCTATTGGGGAAAACCCACGATGACACAACCTCTTTTCT GTGTGTAGGCTGGAGCTGCTTC	Amplificación del gen cat de pDK3 con región homóloga para aroL
aroLcatRv (65)	ACGTAAAGTATAGGCGCTCGAAAATCAACAATTGATCGTCTGTG CATGGGAATTAGCCATGGTCC	Amplificación del gen cat de pDK3 con región homóloga para aroL
laclcatFw (65)	TGTCTTCGGTATCGTCTGATCCCACTACCGAGATATCCGCACCAA GTGTAGGCTGGAGCTGCTTC	Amplificación del gen cat de pDK3 con región homóloga para lacl
laclcatRv (65)	GCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTC ATGGGAATTAGCCATGGTCC	Amplificación del gen cat de pDK3 con región homóloga para lacl
-500aroK5' (20)	CATTCCCTGGTTCGGGCAAT	Verificación de la inactivación de aroK
+500aroK3' (20)	GGGATGGTTGACCGCAGTTT	Verificación de la inactivación de aroK
-500aroL5' (20)	TTGTCCGGCAGTGTGAAGCG	Verificación de la inactivación de aroL
+500aroL3' (20)	TTACCGGGCAAGTGTGAAGC	Verificación de la inactivación de aroL
lacl5'verif (21)	GCAACAGCTGATTGCCCTTCA	Verificación de la inactivación de lacl
lacl3'verif (23)	AAGCGGCGATGGCGGAGCTGAAT	Verificación de la inactivación de lacl
-700pykF5' (20)	CAGCTCACTGCGCTGAATGC	Verificación de la inactivación de pykF
+800pykF3' (20)	CACTATCGGCAGAAGAAGCAG	Verificación de la inactivación de pykF
aroB5' (41)	CATACTGAGTTTCGTAGGAGGGTCCGCTTATGGAGAGGATTG	Amplificación de aroB para clonar
aroB3' (40)	CCTCCTGACGCTGACTTGACTTACGCTGATTGACAATCGG	Amplificación de aroB para clonar
aroG5' (36)	GCGCGCGTCGACAGGAGGGAACAGACATGAATTATC	Amplificación de aroG ^{ibr} para clonar
aroG3' (32)	GGTACGGTTCGACTTACCCGCGACGCGCTTTTA	Amplificación de aroG ^{ibr} para clonar
tktA5' (39)	GCTACGATCTAGCTAGGAGGGAGTCAAAATGTCCTCACG	Amplificación de tktA para clonar
tktA3' (40)	CCTCCTACGAACTCAGTATGTTACAGCAGTTCTTTTGCTT	Amplificación de tktA para clonar
aroE5' (47)	GCGTCGAGGGCCAGGAGGCAGATAATGAAAACCTATGCTGTTT TTG	Amplificación de aroE para clonar
aroE3' (40)	ATCTAGCGGGCCCTACTGATTCACGCGGACAATTCCTCCT	Amplificación de aroE para clonar
aroD5' (47)	GCGCCGCGCTAGCAGGAGGCAGATAATGAAAACCGTAACTGTA AAAG	Amplificación de aroD para clonar
aroD3' (41)	CGCTGGTCTCGCGGCTAGCTTATGCCTGGTGTAATAAGTT	Amplificación de aroD para clonar
zwf5' (47)	GCTATGCTCTAGAAGGAGGCTCGCACTATGGCGGTAACGCAAAC AGC	Amplificación de zwf para clonar
zwf3' (38)	CGCGACGCTCTAGATTACTCAAACCTATTCCAGGAACG	Amplificación de zwf para clonar
5'trc (46)	GTCATCACCGATCGATATCTGCAGGTGGTGAACAGGCCAGCCA CG	Clonación del fragmento de pTrc99A en pBR327par
3'trc (27)	CCATCCGTCAGGATGGCCTTCTGCTTA	Clonación del fragmento de pTrc99A en pBR327par
5'Aro4 (34)	CATGCAGAGCTCATGGAGAGGATTGTCGTTACTC	Clonación del operón de 4 genes en pTrc327parlacl+
3'Aro4 (34)	GCCGTGTCTAGATACTGATTACGCGGACAATTC	Clonación del operón de 4 genes en pTrc327parlacl+
SHK5'A (50)	GTAGTACGCCATGGAGGAGGTGCTGCTATGGAGAGGATTGTGCG TTACTC	Transferencia del operón de 5 genes a pTrc327par
SHK3'A (36)	CGCTGCGGAGCTCTTATGCTGGTGTAAAATAGTT	Transferencia del operón de 5 genes a pTrc327par
SHK5'B (22)	GCCGACATCATAACGGTTCTGG	Amplificación de todo el operón Aro6 con promotor y terminadores
SHK3'B (23)	CAACTCGGTCGCCGCATACACTA	Amplificación de todo el operón Aro6 con promotor y terminadores

MATERIAL ADICIONAL 5

Perfil de producción de SA obtenido al cultivar a la cepa AR36 en un fermentador de 2 L en modo lote alimentado, suministrando simultáneamente Glc y YE durante 14 horas. El volumen inicial de la fermentación fue de 500 mL y el final de 1400 mL. Las concentraciones iniciales de sustratos fueron 80 g/L de Glc y 40 g/L de YE. Las flechas verde y roja indican el inicio y término de la alimentación, respectivamente.

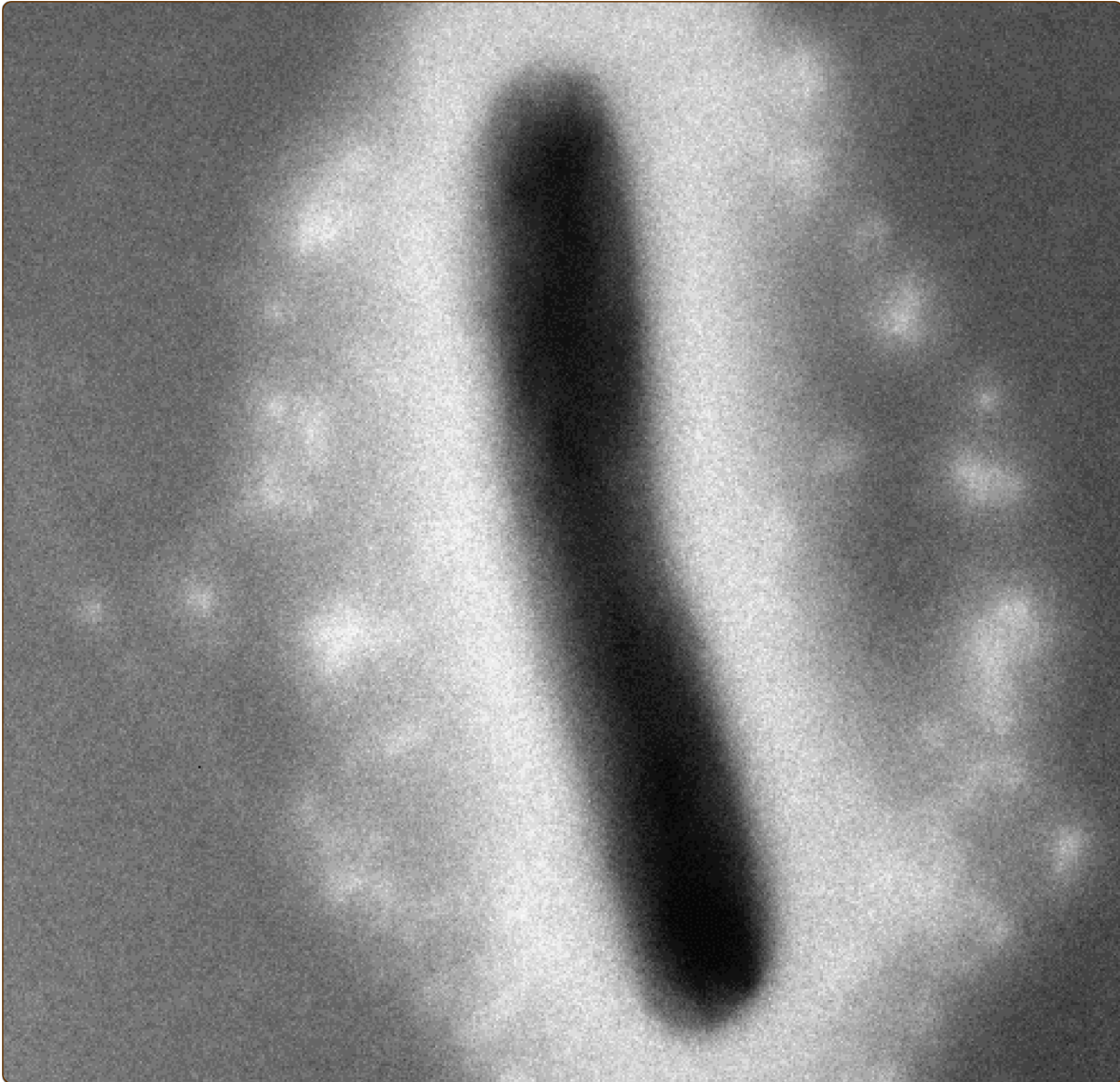


Fermentador 2 L (500 ml vol. inicial – 1400 ml vol. final); **80 g/L Glc + 40 g/L YE**

• $Q_{p_{global}} = 2.46 \text{ gSA/L} \cdot \text{h}$ (vs $1.44 \text{ gSA/L} \cdot \text{h}$ en modo lote; incremento de 70%)

• **Título real de SA = 59 g/L** (vs 33 g/L en modo lote; incremento de 80%)

• $Y_{SA/Glc} = 0.42 \text{ mol/mol}$ (vs 0.42 mol/mol en modo lote)



Constitutive expression of selected genes from the pentose phosphate and aromatic pathways increases the shikimic acid yield in high-glucose batch cultures of an *Escherichia coli* strain lacking PTS and *pykF*

Rodriguez *et al.*

RESEARCH

Open Access

Constitutive expression of selected genes from the pentose phosphate and aromatic pathways increases the shikimic acid yield in high-glucose batch cultures of an *Escherichia coli* strain lacking PTS and *pykF*

Alberto Rodriguez¹, Juan A Martínez¹, José L Báez-Viveros², Noemí Flores¹, Georgina Hernández-Chávez¹, Octavio T Ramírez³, Guillermo Gosset^{1*} and Francisco Bolívar¹

Abstract

Background: During the last two decades many efforts have been directed towards obtaining efficient microbial processes for the production of shikimic acid (SA); however, feeding high amounts of substrate to increase the titer of this compound has invariably rendered low conversion yields, leaving room for improvement of the producing strains. In this work we report an alternative platform to overproduce SA in a laboratory-evolved *Escherichia coli* strain, based on plasmid-driven constitutive expression of six genes selected from the pentose phosphate and aromatic amino acid pathways, artificially arranged as an operon. Production strains also carried inactivated genes coding for phosphotransferase system components (*ptsHlcr*), shikimate kinases I and II (*aroK* and *aroL*), pyruvate kinase I (*pykF*) and the lactose operon repressor (*lacI*).

Results: The strong and constitutive expression of the constructed operon permitted SA production from the beginning of the cultures, as evidenced in 1 L batch-mode fermentors starting with high concentrations of glucose and yeast extract. Inactivation of the *pykF* gene improved SA production under the evaluated conditions by increasing the titer, yield and productivity of this metabolite compared to the isogenic *pykF*⁺ strain. The best producing strain accumulated up to 43 g/L of SA in 30 h and relatively low concentrations of acetate and aromatic byproducts were detected, with SA accounting for 80% of the produced aromatic compounds. These results were consistent with high expression levels of the glycolytic pathway and synthetic operon genes from the beginning of fermentations, as revealed by transcriptomic analysis. Despite the consumption of 100 g/L of glucose, the yields on glucose of SA and of total aromatic compounds were about 50% and 60% of the theoretical maximum, respectively. The obtained yields and specific production and consumption rates proved to be constant with three different substrate concentrations.

Conclusions: The developed production system allowed continuous SA accumulation until glucose exhaustion and eliminated the requirement for culture inducers. The obtained SA titers and yields represent the highest reported values for a high-substrate batch process, postulating the strategy described in this report as an interesting alternative to the traditionally employed fed-batch processes for SA production.

Keywords: Shikimic acid, Synthetic operon, *Escherichia coli*, *pykF*, Aromatic compounds

* Correspondence: gosset@ibt.unam.mx

¹Departamento de Ingeniería Celular y Biotecnología, Instituto de Biotecnología, Universidad Nacional Autónoma de México (UNAM), Apdo. Postal 510-3, Cuernavaca, Morelos 62250, Mexico

Full list of author information is available at the end of the article

Background

Shikimic acid (SA) is an intermediate compound in the aromatic amino acid (AAA) biosynthetic pathway in plants and bacteria (Figure 1). This metabolite is utilized as starting material in the chemical synthesis of oseltamivir phosphate (Tamiflu), used for influenza treatment [1-3]. Several genetic strategies have been reported for improving SA productivity and yield in *Escherichia coli*. These strategies aim to increase the availability of the direct precursors of the AAA pathway, erythrose 4-phosphate (E4P) and phosphoenolpyruvate (PEP), by genetic alterations that promote a convenient redistribution of carbon fluxes in the central metabolism [4,5]. Complementary approaches include the interruption of the AAA pathway after SA formation by inactivation of the genes coding for

shikimate kinases (*aroK* and *aroL*), as well as enhancements in carbon channeling towards SA by overexpression of feedback-resistant DAHP synthases, shikimate dehydrogenase, transketolase, and DHQ synthase enzymes (coded by *aroFGH^{abr}*, *aroE*, *tktA*, and *aroB*, respectively) (Figure 1) [6-9]. In an attempt to further increase the intracellular availability of PEP, strains overexpressing PEP synthase (coded by *ppsA*), or lacking the PEP:carbohydrate phosphotransferase system (PTS) and the pyruvate kinase isozymes (coded by *pykF* and *pykA*), have also been evaluated (Figure 1) [10-13].

Although the implementation of these modifications along with bioengineering strategies has led to diverse *E. coli* strains capable of accumulating SA, the yields obtained to date are still far from the theoretical maximum

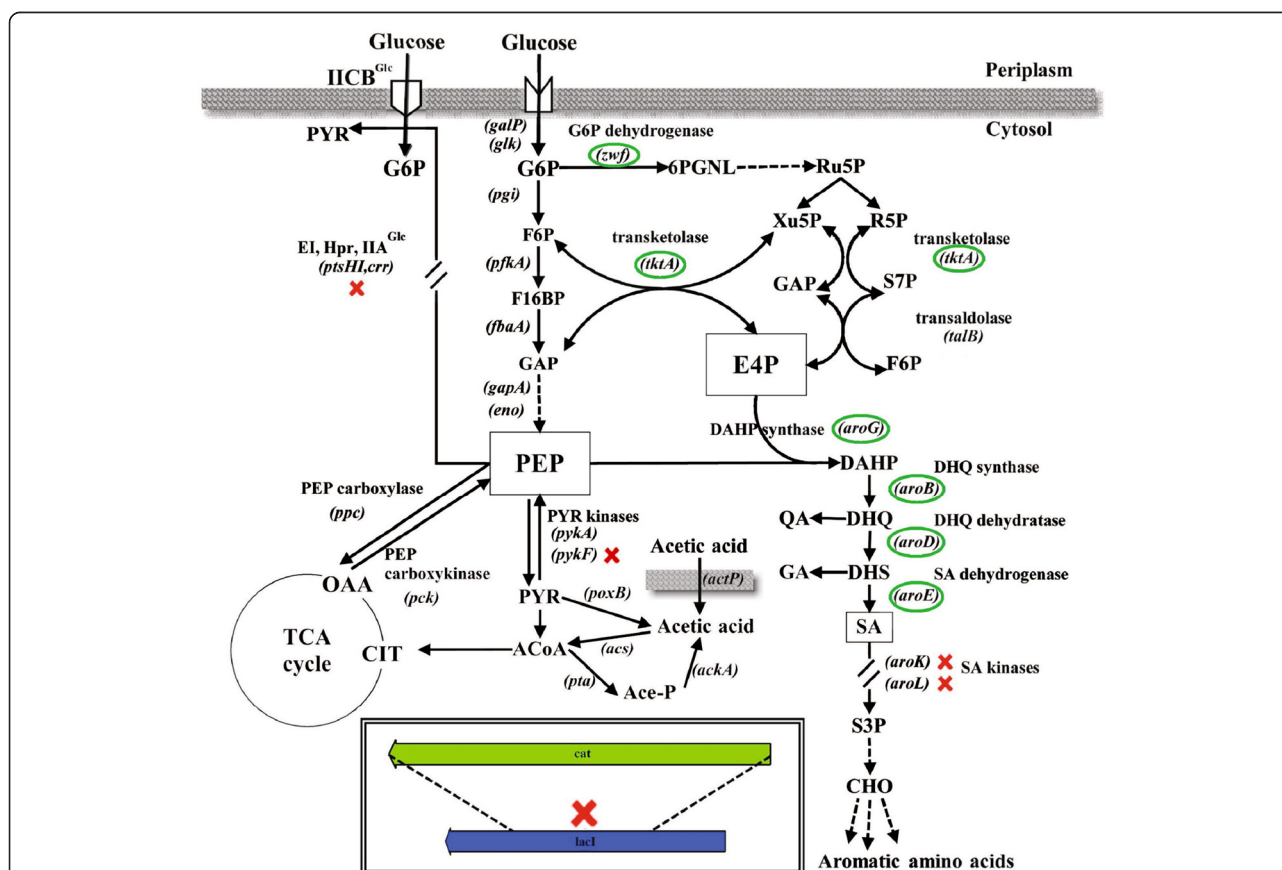


Figure 1 Genetic modifications applied in this report to enhance the production of shikimic acid from glucose in the laboratory-evolved *E. coli* strain PB12. Inactivated genes are indicated with a cross and plasmid-expressed genes are circled (see Methods for details). Dashed arrows indicate more than one catalytic step. G6P = glucose 6-phosphate; F6P = fructose 6-phosphate; GAP = glyceraldehyde 3-phosphate; 6PGNL = 6-phosphogluconolactone; Ru5P = ribulose 5-phosphate; R5P = ribose 5-phosphate; Xu5P = xylulose 5-phosphate; S7P = sedoheptulose 7-phosphate; E4P = erythrose 4-phosphate; PEP = phosphoenolpyruvate; PYR = pyruvate; ACoA = acetyl-coenzyme A; Ace-P = acetyl phosphate; CIT = citrate; OAA = oxaloacetate; DAHP = 3-deoxy-D-arabinoheptulosonate 7-phosphate; DHQ = 3-dehydroquinic acid; DHS = 3-dehydroshikimic acid; QA = quinic acid; GA = gallic acid; SA = shikimic acid; S3P = shikimate 3-phosphate; CHO = chorismate; IICB^{Glc} = membrane component of glucose-specific PTS permease; E1 = PTS enzyme 1; Hpr = PTS histidine protein; IIA^{Glc} = cytosolic component of glucose-specific PTS permease. Genes coding for enzymes not named in the figure: *galP*, galactose permease; *glk*, glucokinase; *pgi*, phosphoglucose isomerase; *pfkA*, 6-phosphofructokinase I; *fbaA*, fructose biphosphate aldolase class II; *gapA*, glyceraldehyde 3-phosphate dehydrogenase; *eno*, enolase; *actP*, acetate permease; *acs*, acetyl-coenzyme A synthetase; *pta*, phosphate acetyltransferase; *ackA*, acetate kinase; *poxB*, pyruvate oxidase.

[10,11,14,15]. This can be partially attributed to the fact that most expression systems used involve genes controlled by a mixture of inducible and native promoters of variable strengths, contained in more than one type of plasmid. These imbalances often cause a metabolic burden and heterogeneities on the intensity and temporality of gene expression, which may translate into suboptimal production capabilities of the recombinant strains, resulting in low productivity and yield of SA [16-19]. Consequently, optimized DNA expression systems and genetic backgrounds are needed for promoting a more efficient carbon channeling towards SA formation.

With the goal of producing aromatic compounds, our group has constructed and characterized strains lacking PTS, the major glucose transport system [20]. One of such strains is PB11, which grows poorly on glucose due to the inactivation of PTS [21,22]. Strain PB12, a derivative of PB11 with a 400% increased growth rate, was isolated in a short laboratory adaptive evolution process to foster derivatives growing in glucose [21,22]. This strain can simultaneously utilize glucose and other carbon sources (acetate, glycerol and various carbohydrates) in minimal medium due to the lack of catabolite repression exerted by PTS [21,23]. Whole genome analysis allowed the identification of the genetic changes that occurred in PB12, suggesting that the deletion of 12 genes, including *rppH*, *galR* and *mutH*, is the main reason for its rapid growth on glucose [24].

It was reported that PB12, which assimilates glucose by the non-PTS symporter GalP [25], can be engineered to accumulate SA in culture media containing glucose (Glc) and yeast extract (YE). For instance, when PB12 was transformed with two plasmids encoding four biosynthetic genes, the variant with both functional pyruvate kinases accumulated the highest SA concentration (up to 7 g/L), but the highest yield of aromatic compounds was achieved by a derivative with an inactivated *pykF* gene [11]. This result may be related to other reported effects caused by the inactivation of *pykF*, such as an increase in plasmid copy number per cell [26], low acetate production due to less glycolytic overflux [13,27,28], or higher concentrations of the AAA pathway enzymes [29]. Interestingly, in spite of the aforesaid features that can be beneficial for SA production, the metabolic engineering efforts to overproduce this compound have been mainly applied to strains with a *pykF*⁺ background, probably because of their typically higher glucose consumption rates compared to the *pykF*⁻ counterparts [11,30].

Here, we propose that a PTS⁻ *pykF*⁻ strain has the potential to increase the yield and titer of SA when compared to an isogenic *pykF*⁺ strain, provided that the gene expression system permits an appropriate temporal coordination in the synthesis of the enzymes required to channel the carbon towards SA, while reducing

the accumulation of acetate and intermediate compounds in the AAA pathway. In order to accomplish this goal a synthetic operon was constructed containing the coding sequences of six genes selected from the pentose phosphate (PPP) and AAA pathways (Figure 1), controlled by a single constitutive Trc promoter [31] and inserted it into a high-copy plasmid containing a region that confers segregational stability (Figure 2) [32]. The resulting plasmid was transformed into a modified PB12 strain with inactive *aroK*, *aroL*, *pykF*, and *lacI* genes, and was cultured in fermentors using mineral media containing Glc and supplemented with YE. Overall, the strategy proposed in this report allowed the overproduction of SA from the beginning of the culture, resulting in a high titer and yield of SA with relatively low accumulation of acetate and aromatic byproducts. It was also found that, under the high-substrate conditions tested, the SA titer was independent of the YE concentration and the maximum biomass produced depended exclusively on the initial YE concentration but not on the amount of glucose.

Results and discussion

Construction of strains derived from PB12 *aroK* *aroL*⁻ containing a plasmid designed for the constitutive expression of a synthetic operon used in the production of shikimic acid

Unpublished evidence from our laboratory indicates that the production of aromatic compounds in the laboratory-evolved strain PB12 can attain higher levels when the transcriptional induction of the genes involved in canalizing carbon flux into the AAA pathway occurs at the beginning of fermentations. Taking into account this observation, a new strategy was developed for optimizing the production of SA in PB12 carrying inactive *aroK* and *aroL* genes (Figure 1). This strategy included the design and construction of a plasmid for the strong and stable expression of six key genes arranged in the form of a synthetic operon, controlled exclusively by a single Trc promoter. In order to reduce metabolic burden, a single plasmid derived from pBR327 carrying the *par* locus for increased plasmid stability was utilized as the vector [32], after incorporating a fragment containing the promoter, polylinker, and transcriptional terminators from pTrc99A (Figure 2).

The initial part of the operon was constructed by sequential amplification and ligation of the first 4 coding sequences (*aroB*, *tktA*, *aroG*^{fbt}, and *aroE*) into the polylinker of plasmid pBRINT-Ts Cm, used as a cloning scaffold (see Methods). Later, the 4-gene construction was transferred to the hybrid plasmid pTrc327par in conjunction with 2 more genes (*aroD* and *zwf*), leading to an 8Kb operon contained in a 12Kb plasmid (Figure 2). The resulting plasmid, termed pTrcAro6, was transformed into the PB12 *aroK*⁻ *aroL*⁻ strain devoid of the *lacI*^q gene,

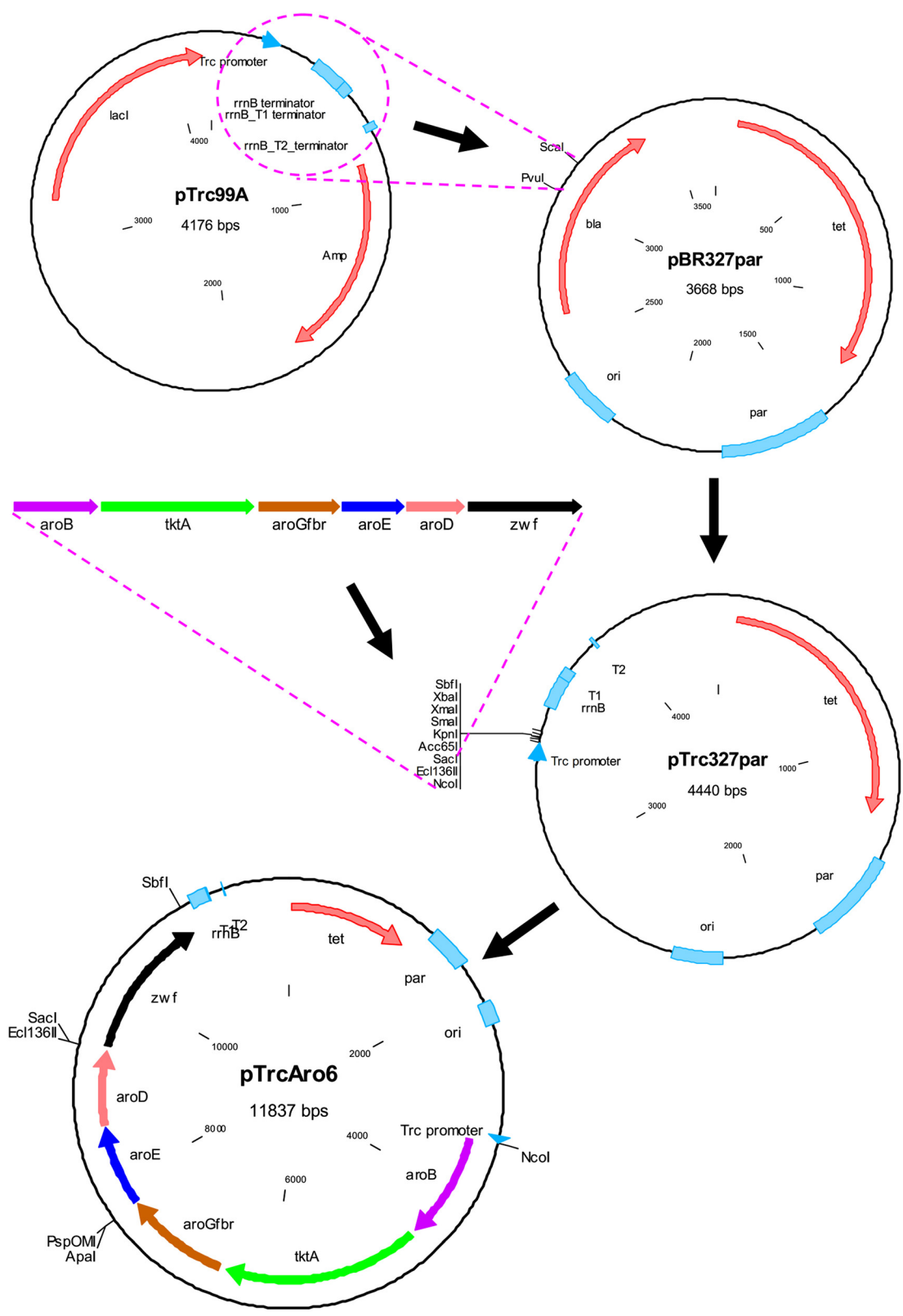


Figure 2 (See legend on next page.)

(See figure on previous page.)

Figure 2 Simplified scheme of the steps required in the construction of plasmid pTrcAro6, carrying 6 synthetic genes under the control of the Trc promoter. First, a segment of pTrc99A was amplified and ligated into pBR327par, creating plasmid pTrc327par. A synthetic operon comprising the required genes (*aroB*, *tktA*, *aroG^{fbr}*, *aroE*, *aroD*, and *zwf*) was assembled separately and transferred to pTrc327par, generating pTrcAro6. The dotted lines indicate the site and orientation of some of the performed ligation reactions. Only the relevant restriction sites are displayed. A more detailed scheme of the constructions is presented in Additional file 3.

allowing constitutive expression of the genes of interest (Table 1). For simplicity, the generated PB12 *aroK⁻ aroL⁻ lacI⁻* strain was termed AR2. After the *pykF* gene was inactivated in AR2, the resulting strain was named AR3. Strains derived from AR2 and AR3 carrying plasmid pTrcAro6 were named AR26 and AR36, respectively (Table 1).

The spatial arrangement of the coding sequences that constitute the synthetic operon in pTrcAro6, flanked by the Trc promoter and transcriptional terminators, is shown in Figure 2. *aroB* is the first gene in the operon since several evidences indicate that its low expression is one of the limiting steps in the production of aromatic compounds [33-35]. Plasmid pTrcAro6 also carries the *tktA* and *aroG^{fbr}* genes, whose products are involved in E4P synthesis and its condensation with PEP to form DAHP, the first aromatic compound (Figure 1). *aroD* and *aroE* genes were also included to promote an efficient conversion of DHQ to SA. Additionally, this plasmid carries

the *zwf* gene, coding for the first enzyme of the PPP (Figure 1). The decision to include this gene was based on the following observations: 1) the overexpression of *zwf* substantially recovered the growth rate loss due to plasmid metabolic load in strain JM101 growing on glucose as only carbon source [36]; 2) it has been reported that strain PB12 displays a particularly low carbon flux partition at the glucose 6-phosphate (G6P) node towards the PPP (5% of the consumed G6P compared to 22% in the parental strain JM101) [25]. Therefore, an overexpression of this gene should increase NADPH availability, required in catalytic amounts by the enzyme shikimate dehydrogenase (AroE), and may alleviate potential growth affectations by redirecting more G6P towards nucleotide and amino acid biosynthesis in strains derived from PB12 [37]. However, the experiments presented in this report did not aim to dissect the specific effect of any utilized gene but instead sought to characterize the consequences of expressing all of them as an operon.

Table 1 *Escherichia coli* strains and plasmids utilized in this report

Strains		
Name	Characteristics	Reference
JM101	F' <i>traD36 proA⁺ proB⁺ lacIq lacZΔM15/supE thi Δ(lac-proAB) rpoS(33 am)</i>	Messing [41]
PB11	JM101 Δ(<i>ptsH, ptsI, crr</i>): <i>kan</i>	Flores et al. [21]; Flores et al. [22]
PB12	PB11, PTS ⁻ Glc ⁺ ; laboratory-evolved strain	Flores et al. [21]; Flores et al. [22]
AR2	PB12 <i>lacI⁻ aroK⁻ aroL⁻</i>	This work
AR3	PB12 <i>lacI⁻ aroK⁻ aroL⁻ pykF⁻</i>	This work
AR26	AR2 + pTrcAro6 (Trc/ <i>aroB⁺tktA⁺aroGfbr⁺aroE⁺aroD⁺zwf⁺</i>)	This work
AR36	AR3 + pTrcAro6 (Trc/ <i>aroB⁺tktA⁺aroGfbr⁺aroE⁺aroD⁺zwf⁺</i>)	This work
AR2e	AR2 + pTrc327par (plasmid vector without synthetic operon)	This work
AR3e	AR3 + pTrc327par (plasmid vector without synthetic operon)	This work
Plasmids		
Name	Characteristics	Reference
pKD3	PCR template for amplification of chloramphenicol resistance gene flanked by homologous sequences	Datsenko and Wanner [42]
pKD46	Plasmid expressing λ-Red recombinase system with thermosensitive origin of replication	Datsenko and Wanner [42]
pCP20	FLP recombinase expression plasmid	Cherepanov and Wackernagel [43]
pBR327par	Derivative of pBR322 exhibiting increased copy number and segregational stability	Zurita et al. [32]
pTrc99A	Multipurpose expression plasmid bearing <i>lacI</i> gene and a polylinker in front of Trc promoter	Amann et al. [31]
pTrc327par	Contains the promoter, polylinker, and terminators of pTrc99A, and <i>par</i> and <i>ori</i> regions of pBR327par	This work (Figure 2)
pTrcAro6	pTrc327par containing a 6-gene synthetic operon to enhance the production of shikimate	This work (Figure 2)

In order to promote an efficient translation of every gene, each coding sequence was amplified using designated primers that introduced a consensus Shine-Dalgarno sequence located 8 bp upstream of the translation start site. The nucleotide sequence of the constructed operon is presented in Additional file 1.

Assessment of the effects caused by *pykF* inactivation in strains expressing the Aro6 operon

To evaluate the effects caused by *pykF* inactivation on the production of SA, the performance of production strains AR26 (*pykF*⁺) and AR36 (*pykF*⁻) was compared using shake flasks containing 15 g/L of Glc and 5 g/L of YE. As a control, the same strains containing an empty pTrc327par plasmid (without the Aro6 operon), AR2e and AR3e, were also included.

Even though SA accumulated in all cases, as expected for mutants in *aroK* and *aroL*, the strains containing pTrcAro6 reached higher SA concentrations than the ones with an empty plasmid (Figure 3b). Moreover, the SA titer was almost two times higher in AR36 than in AR26 (6.1 g/L vs. 3.3 g/L). A decrease in Glc consumption was observed in strain AR26 after approximately 18 h of culture, correlating with high acetate concentration and an arrest in the production of SA. In contrast, strain AR36 exhibited constant Glc consumption and negligible amounts of acetate were produced (Figure 3c, 3d). These results demonstrate that the genes present in the artificial operon are functional and promote the production of SA since the beginning of the culture. Their constitutive expression diminished the specific growth rate (μ) by 25% in the *pykF*⁺ background, and marginally increased it in the *pykF*⁻ variant, but did not cause significant changes to the maximum biomass produced (X_{\max}) compared to strains with an empty plasmid (Figure 3a). Remarkably, in the operon-expressing strains under these growth conditions, the inactivation of the *pykF* gene increased the production of SA, eliminated the accumulation of acetate, and allowed steady Glc consumption.

To determine if the higher acetate production and lower SA production in AR26 compared to AR36 is a consequence of the inherently low oxygen availability and acidification of the medium in shake flask cultures, both strains were cultured in 1 L batch fermentors under controlled conditions of pH and dissolved oxygen tension (DOT). As an approach to increase the SA titer, the initial concentration of Glc in these experiments was raised to 100 g/L, and the YE concentration was concomitantly increased to 15 g/L to allow higher biomass generation.

Under these conditions strain AR36 produced 42 g/L of SA in 60 h, consuming all the Glc, and accumulating 12 g/L of acetate. In contrast, after 47 h strain AR26 produced a maximum of 13 g/L of SA, did not exhaust the

Glc, and accumulated 29 g/L of acetate (Figure 3e and Table 2). Regardless of the controlled conditions in the fermentors, where the pH was kept at 7 and the DOT was higher than 20% at all times, the production profiles of both strains resembled the behavior observed in shake flasks, with AR26 producing more acetate and less SA. Even when the global volumetric Glc consumption rate ($Q_{S_{\text{global}}}$), μ and X_{\max} attained by both strains were similar, the productivity, yield, and titer were more than twofold higher in AR36 than in AR26 (Figure 3e and Table 2).

It is remarkable that such large differences in acetate and SA production were observed by disrupting only one gene, which demonstrates the advantages of the combined inactivation of PTS and *pykF* when using a constitutive expression system in an evolved *E. coli* strain. To account for the observed improvements in SA production, we suggest that the early and constant expression of enzymes encoded in the operon could maintain a steady consumption of glycolytic intermediates throughout the cultures, preventing high fluctuations in their intracellular concentrations. We hypothesize that the combination of this steady metabolic state with a reduced flux from PEP to pyruvate caused by the inactivation of the *pykF* gene may increase the availability of PEP and other glycolytic precursors for SA production without decreasing the Glc consumption rate. However, we acknowledge that in the absence of measured intracellular metabolite concentrations, these remarks are speculative.

Fermentation profiles of AR36 in batch cultures

Taking into account the previous results, AR36 was selected for further characterization of its kinetic and stoichiometric performance in 1 L fermentors. To accomplish such purpose, the production of SA was tested with three different high-substrate culture conditions. Growth, Glc and byproducts were measured for each case, which in turn allowed a comparison of the productivities and yields.

First, 50 g/L of Glc and 15 g/L of YE were utilized (Figure 4a). Growth occurred during the first 10 h, generating 6.3 g/L of dry cell weight with a μ of 0.53 h⁻¹. Under this condition, 24 g/L of SA were produced in 32 h. Glc consumption and SA production occurred since the beginning of the fermentation and lasted until Glc exhaustion, although the specific Glc consumption rate (q_s) and specific SA productivity (q_p) were higher in exponential phase (Table 3). The resulting yield of SA on Glc ($Y_{SA/Glc}$) was 0.47 mol/mol and the global volumetric SA productivity ($Q_{p_{\text{global}}}$) was 0.74 gSA/L*h (Table 3). With respect to the accumulation of byproducts in the SA pathway, concentrations of 2.4 g/L of DAHP, 2.1 g/L of DHS, 1.4 g/L of QA, 0.4 g/L of GA, and 0.3 g/L of DHQ, were present in the supernatant at the end of the fermentation (Figure 5a). Under these conditions, virtually no acetate was produced during

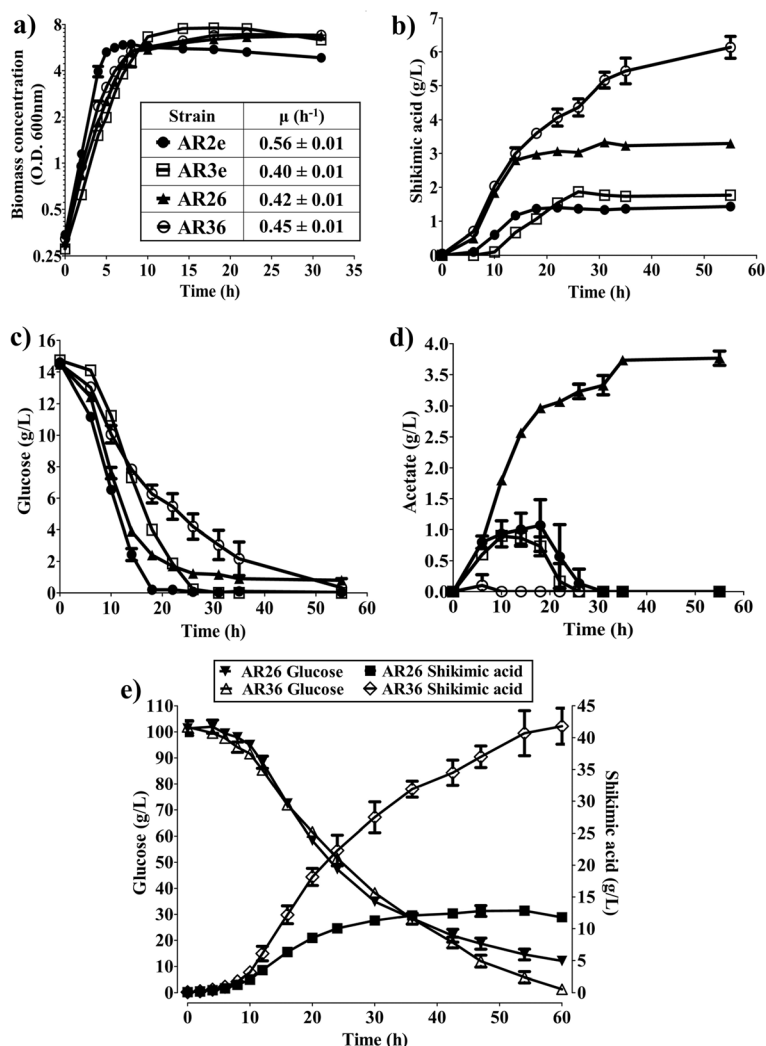


Figure 3 Behavior of strains AR26, AR36, and their empty-plasmid derivatives, AR2e (*pykF*⁺) and AR3e (*pykF*⁻), using shake flasks containing 15 g/L of Glc and 5 g/L of YE (a,b,c,d), and 1 L fermentors containing 100 g/L of Glc and 15 g/L of YE (e). a) Growth; b) SA production; c) Glc consumption; d) acetate production; e) Glc consumption and SA production of AR26 and AR36 in fermentors. Error bars represent standard deviation.

the course of the fermentation, reaching a maximum concentration of 1.5 g/L after 32 h (Figure 4a).

Considering that 50 g/L of Glc were consumed completely, a second batch experiment was initiated with 100 g/L of Glc and 15 g/L of YE. As stated in the comparison with AR26 in the previous section, AR36 grown under these conditions produced approximately 42 g/L of SA in 60 h (Figure 4b). In this case, after consuming about 100 g/L of glucose and attaining the maximum concentration of SA, the strain produced 12 g/L of acetate. The values obtained for $Y_{SA/Glc}$, $Q_{P_{global}}$, $Q_{S_{global}}$, X_{max} , and μ , were similar to those obtained with 50 g/L of Glc and 15 g/L of YE (Table 3). These experiments show that when using the same YE concentration, twice the amount of Glc is consumed in almost twice the time, indicating that

the average glucose consumption rate is maintained between both culture conditions. Concentrations of 4.8 g/L of DAHP, 2.8 g/L of DHS, 3.4 g/L of QA, 0.7 g/L of GA, and 0.9 g/L of DHQ, were present in the supernatant after 60 h (Figure 5b). Interestingly, when doubling the Glc concentration the intermediate products of the AAA pathway increased in a fairly proportional manner with the SA, indicating that the consumption of 100 g/L of Glc did not apparently generate new carbon flux bottlenecks. As a result, the amount of SA formed with respect to the total aromatic compounds produced was close to 80% in both experiments (Figure 6).

The effect of increasing the YE on SA productivity was investigated with a third set of experiments, using 100 g/L of Glc and 30 g/L of YE. Although the biomass was

Table 2 Comparative data from 1 L batch fermentations of strains AR26 and AR36, using 100 g/L of Glc and 15 g/L of YE as substrates

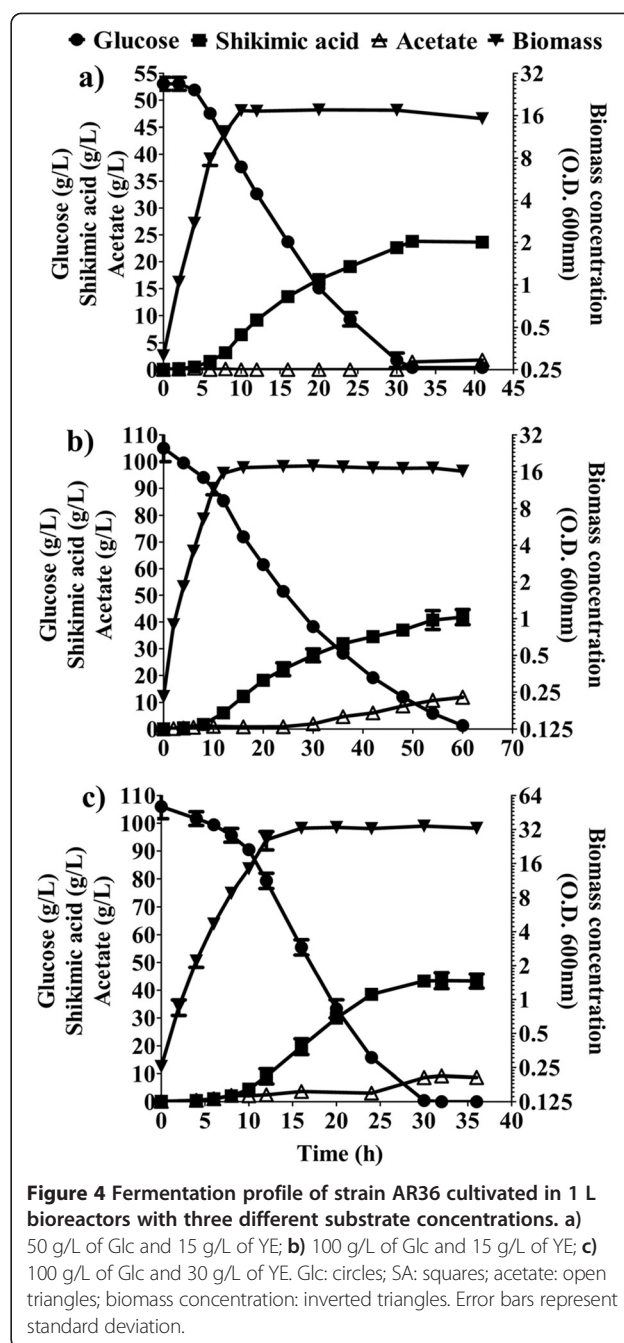
Strain	AR26	AR36
SA titer (g/L)	12.95 ± 0.64	41.80 ± 2.83
Glc consumed (g/L)	82.65 ± 4.88	103.70 ± 6.79
Duration of culture (h)	47	60
$Y_{SA/Glc}$ (mol/mol)	0.16 ± 0.02	0.42 ± 0.00
Acetate titer (g/L)	29.35 ± 0.21	11.90 ± 0.14
X_{max} (g/L)	6.18 ± 0.10	6.54 ± 0.09
μ (h^{-1})	0.45 ± 0.01	0.45 ± 0.02
$Q_{p_{global}}$ (gSA/L*h)	0.27 ± 0.02	0.75 ± 0.07
$Q_{s_{global}}$ (gGlc/L*h)	-1.76 ± 0.10	-1.73 ± 0.11

$Y_{SA/Glc}$ = SA yield from Glc; X_{max} = maximum attained biomass expressed as dry cell weight; $Q_{p_{global}}$ = volumetric SA productivity for the entire fermentation; $Q_{s_{global}}$ = volumetric Glc consumption rate for the entire fermentation; μ = specific growth rate.

doubled when using twice the concentration of YE, the SA titer, μ and $Y_{SA/Glc}$ were very similar to those obtained in the culture with 100 g/L of Glc and 15 g/L of YE (Figure 4b and Figure 4c). In conjunction with data obtained from the other two conditions, these findings suggest that the amount of YE primarily determines the maximum biomass that can be achieved. Additionally, an increment in the initial YE concentration did not alter the SA titer, and supports the observation that SA is mainly being produced from glucose. The direct relation between the initial YE concentration and the maximum biomass generated, regardless of the initial Glc concentration tested in these growth conditions, suggests that one or more limiting nutrients are being supplied by the YE. It would also appear that such nutrients cannot be synthesized from Glc, hence their depletion from YE limits growth long before Glc is exhausted. It is expected that the aromatic amino acids and vitamins present in the YE that are needed to counteract AR36 auxotrophy will become limiting; however, other compounds in this complex media may also play a role in growth limitation over time.

For a starting YE concentration of 30 g/L, a total of 106 g/L of Glc and 43 g/L of SA were consumed and produced, respectively, in approximately half the time than the fermentation with 15 g/L of YE. With 30 g/L of YE, the $Q_{s_{global}}$ and $Q_{p_{global}}$ increased twofold, in comparison with the fermentations with 15 g/L of YE, even though the SA titer remained unchanged (Table 3). Since the biomass also increased twofold, the calculated qp and qs were similar between the three experiments, both in exponential and stationary phases, exhibiting the metabolic robustness of the engineered strain under the tested conditions.

In addition, the results showed that an increase in YE concentration did not increase considerably the



concentration of SA pathway intermediates (Figure 5c). In this respect, it has been acknowledged that the presence of high quantities of pathway intermediates can negatively impact the recovery of SA from the fermentation broth [7,38]. This concern has directed some efforts into the subject, leading to the testing of culture conditions, genetic backgrounds, and the use of non-metabolizable glucose analogs, as attempts to minimize byproduct generation [39].

In these experiments, a high proportion of SA relative to byproducts was detected without applying any further

Table 3 Comparison of measured metabolites and calculated kinetic and stoichiometric parameters between three fermentations of strain AR36 with different substrate concentrations

Strain	AR36	AR36	AR36
Culture conditions	Batch 50 g/L Glc + 15 g/L YE	Batch 100 g/L Glc + 15 g/L YE	Batch 100 g/L Glc + 30 g/L YE
SA titer (g/L)	23.80 ± 0.00	41.80 ± 2.83	43.30 ± 0.57
Glc consumed (g/L)	52.65 ± 1.20	103.70 ± 6.79	105.55 ± 4.45
Duration of culture (h)	32	60	30
Y_{SA/Glc} (mol/mol)	0.47 ± 0.01	0.42 ± 0.00	0.42 ± 0.01
Acetate titer (g/L)	1.45 ± 0.00	11.90 ± 0.14	8.65 ± 0.92
X_{max} (g/L)	6.30 ± 0.09	6.54 ± 0.09	12.54 ± 0.06
μ (h⁻¹)	0.53 ± 0.03	0.45 ± 0.02	0.45 ± 0.00
Q_pglobal (gSA/L*h)	0.74 ± 0.00	0.75 ± 0.07	1.44 ± 0.02
Q_sglobal (gGlc/L*h)	-1.65 ± 0.04	-1.74 ± 0.11	-3.52 ± 0.15
q_pexp (gSA/gDCW*h)	0.38 ± 0.00	0.34 ± 0.03	0.46 ± 0.06
q_sexp (gGlc/gDCW*h)	-1.14 ± 0.15	-1.11 ± 0.00	-1.18 ± 0.03
q_psta (gSA/gDCW*h)	0.16 ± 0.01	0.19 ± 0.02	0.20 ± 0.01
q_ssta (gGlc/gDCW*h)	-0.36 ± 0.01	-0.42 ± 0.01	-0.42 ± 0.03

Y_{SA/Glc} = SA yield from Glc; X_{max} = maximum attained biomass expressed as dry cell weight; Q_pglobal = volumetric SA productivity for the entire fermentation; Q_sglobal = volumetric Glc consumption rate for the entire fermentation; q_pexp = specific SA productivity for exponential phase; q_sexp = specific Glc consumption rate for exponential phase; q_psta = specific SA productivity for stationary phase; q_ssta = specific Glc consumption rate for stationary phase; μ = specific growth rate.

modification to the strain or process. The concentration of each pathway intermediate was compared against the sum of all aromatic intermediates, and their percentages were used to calculate the molar ratio of SA to each byproduct at the end of the fermentations (Figure 6). The ratio of SA turned out to be higher than 10 for DHS, QA, or DAHP, and higher than 40 for GA or DHQ for all the substrate concentrations tested. Remarkably, in all the conditions the obtained SA yields were close to 50% of the theoretical maximum and the yields of total aromatic compounds (TAC) were above 60% of the theoretical maximum, estimated as 0.86 mol_{TAC}/mol_{Glc} (see Methods and Figure 6). This reflects the efficient redirection of glucose towards the AAA pathway in strain AR36, even when using high-glucose batch cultures. The ratio of SA to byproducts, as well as the obtained SA and TAC yields are fairly constant for all the conditions evaluated, and represent to our knowledge the highest reported values for a SA production fermentation process. These improvements can be justified by taking into account that the platform present in the engineered strain allows a more homogeneous expression of the necessary enzymes on an efficient genetic background. This, in contrast with other expression systems where the required genes are expressed from separate plasmids, under different promoters, or in strains not optimized for efficient use of high levels of Glc. In addition to the advantages concerning the dynamics of gene expression, the fact that IPTG is not needed to induce the Aro6 operon represents an

important economical benefit for the production process, since the high price of IPTG restricts its use in large-scale fermentations.

Insights on the glycolytic and acetate metabolisms of strain AR36 by RT-qPCR

To gain a deeper insight of the metabolic changes induced by the constitutive expression of the Aro6 synthetic operon in strain AR36, transcript levels of several genes were measured at three different growth stages in cultures with 50 g/L of Glc and 15 g/L of YE. As detailed in Methods, data obtained from early exponential phase (EE), late exponential phase (LE), and stationary phase (ST) were normalized against the values measured from strain AR3e at EE, grown under the same culture conditions.

The results indicate that the presence and expression of the operon in strain AR36 increases the transcriptional levels of several genes coding for glycolytic enzymes during the EE and LE phases (Figure 1 and Figure 7a). The rise in expression of genes *galP* and *glk* is particularly interesting because it has been reported that their products control the import and phosphorylation of glucose in PB12, the parental strain of AR36 [21,25]. Furthermore, there is a significant increase in the transcriptional levels of *pgi* and *eno*, but not *pykA*. These changes may translate into higher availability of PEP and fructose 6-P (which can be directly converted into E4P by plasmid-encoded transketolase), increasing the yield of aromatic compounds. We theorize that the observed

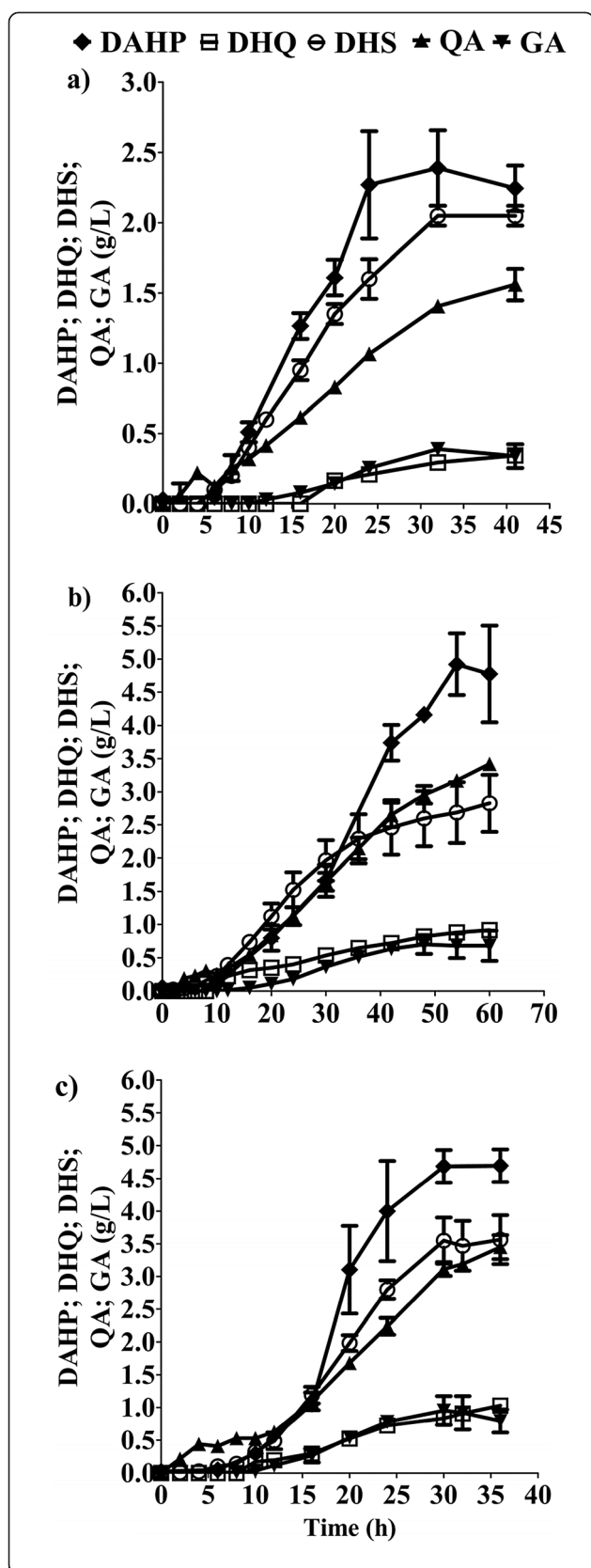
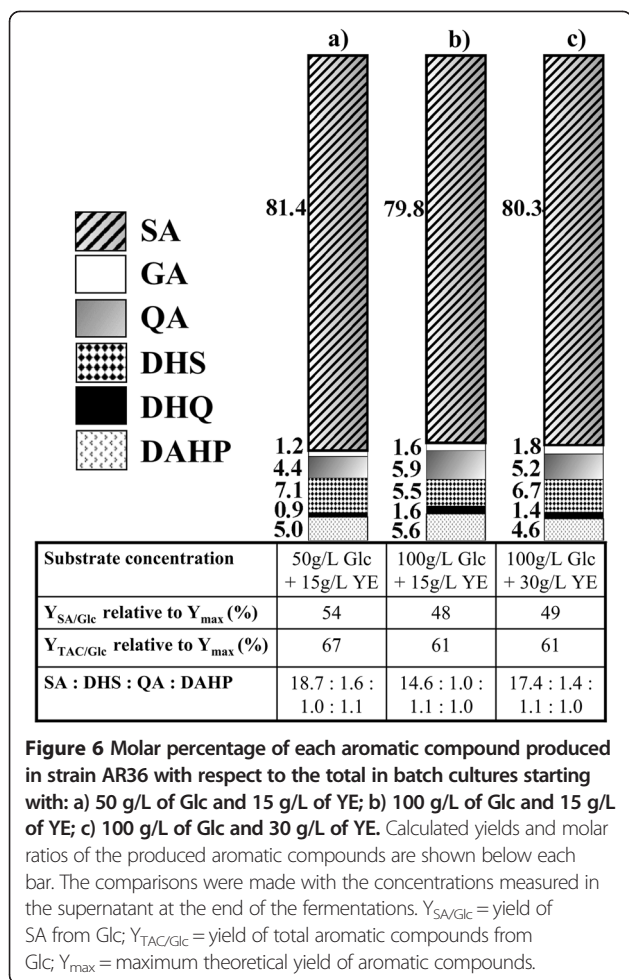


Figure 5 Aromatic byproducts of the SA pathway detected in 1 L fermentor cultures of strain AR36 using three different substrate concentrations. **a)** 50 g/L of Glc and 15 g/L of YE; **b)** 100 g/L of Glc and 15 g/L of YE; **c)** 100 g/L of Glc and 30 g/L of YE. Diamonds: DAHP (3-deoxy-D-arabinoheptulose 7-phosphate); squares: DHQ (3-dehydroquinic acid); circles: DHS (3-dehydroshikimic acid); triangles: QA (quinic acid); inverted triangles: GA (gallic acid). Error bars represent standard deviation.

upregulation of glycolytic genes in strain AR36 could be one of the consequences to low levels of some glycolytic intermediates (glucose 6-phosphate, fructose 6-phosphate and PEP), caused by the strong and constitutive expression of the operon-encoded enzymes that consume these metabolites.

On the other side, the transcriptional levels of genes coding for enzymes involved in acetate biosynthesis (*poxB*, *ackA* and *pta*) were not modified by the presence of the synthetic operon, while *actP* and *acs*, coding for enzymes involved in acetate assimilation, were strongly upregulated in the EE and LE phases (Figure 7b). Upregulation of *actP* and *acs* genes has also been detected in the exponential growth phase in the parental strain PB12 that is capable of co-utilizing Glc and acetate in minimal medium [21]. These findings correlate with the low levels of acetate in the assayed growth condition (Figure 4a). Importantly, the transcriptional values of these genes involved in acetate assimilation were low in ST phase (Figure 7b). If this response is representative of the other growth conditions used, it could partially explain the acetate accumulation observed in fermentations with 100 g/L of Glc, which consume higher amounts of Glc during stationary phase (Figure 4b and Figure 4c). These results highlight *actP* and *acs* as potential gene targets to artificially increase their expression in late culture stages, taking advantage of the expected capabilities of strain AR36 to utilize simultaneously Glc and acetate, present in its parental strain PB12 [21,40].

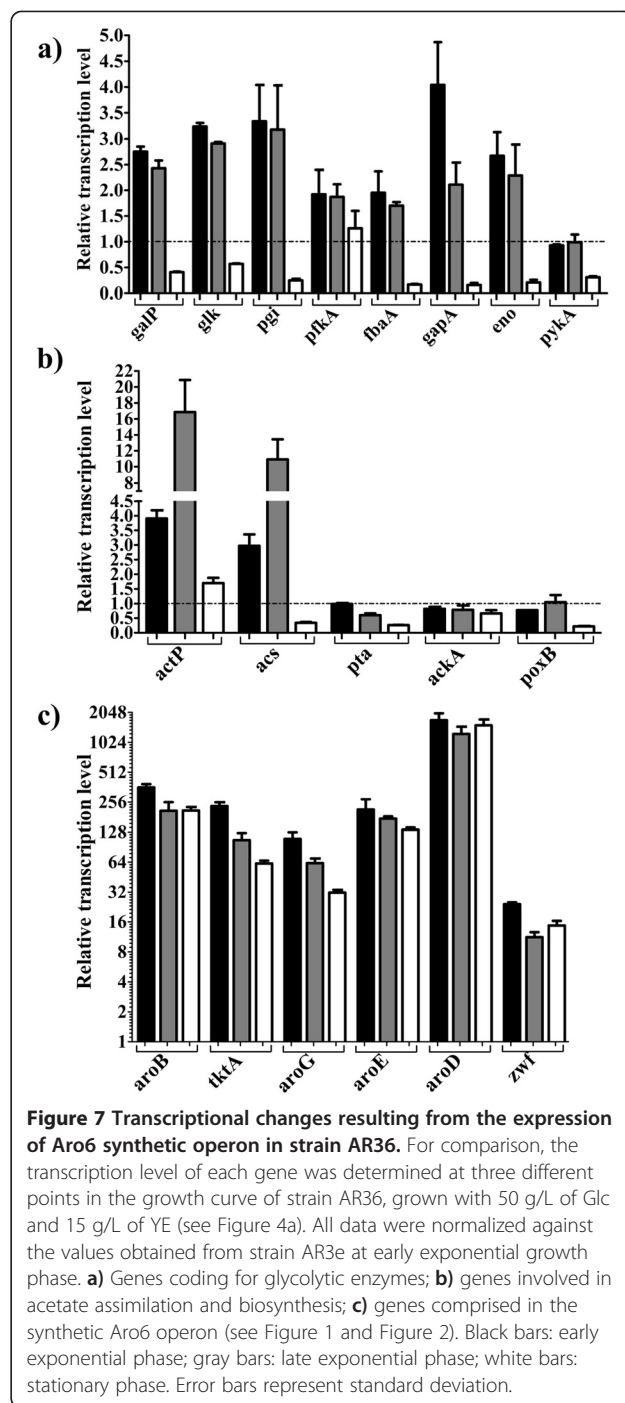
The genes present in the synthetic operon showed very strong expression levels (even in stationary phase), reflecting the constitutive nature of the promoter and high copy number of the plasmid (Figure 7c). These results correlate with the uninterrupted Glc consumption and SA production observed during the entire fermentation (Figure 4a), suggesting that the enzymes coded by the genes in the operon are present throughout the cultivation time. It can be seen in Figure 7c that the transcript levels of *aroD* and *zwf* are comparatively higher and lower, respectively, than the other four genes in the operon. This observation should be taken with caution because the six genes in the operon are being compared to the ones present in the chromosome of reference strain AR3e. Since the values obtained for the six genes are not normalized



between them, variations amongst their chromosomal expression in strain AR3e can alter the relative comparisons with strain AR36. Nevertheless, the transcriptomic data is consistent with the high ratio of SA to aromatic intermediates obtained in the tested conditions, which is to be expected if all the genes in the operon were adequately expressed. Together with kinetic and stoichiometric data, these results highlight the benefits of employing a constitutively-expressed synthetic operon as an alternate strategy to increase the yield of SA from Glc in an evolved strain that lacks PTS and *pykF*.

Conclusions

E. coli is the microorganism that has given the best results for SA production, with engineered strains that can accumulate up to 85 g/L using 10 L fermentors in fed-batch processes [10]. In this report, we showed that the constitutive and synchronous expression of a six-gene synthetic operon, in a laboratory-evolved strain bearing simultaneous PTS and *pykF* inactivations, resulted in a competitive process for the production of SA. The expression of Aro6 operon in the PTS⁻ *pykF*⁻ derivative



resulted in higher $Q_{p,global}$ and similar $Q_{s,global}$ than its PTS⁻ *pykF*⁺ counterpart. In addition, the Glc consumption and SA production profiles of strain AR36 are consistent with the observed increase in transcription levels of glycolytic genes as a response to the constitutive expression of the operon in this strain. These features translated into significant improvements in growth and production parameters in strain AR36 (producing 0.74 gSA/L*h with 54% yield, using 50 g/L of Glc + 15 g/L of YE), compared

to the PTS⁻ *pykF* strain reported by Escalante et al. in 2010 (producing 0.11 gSA/L*h with 26% yield, using 25 g/L of Glc + 15 g/L YE).

Albeit fed-batch fermentations in the past have given the best results with respect to SA production, here we report that under the appropriate conditions, batch cultures of strain AR36 with an initially high concentration of Glc can also be efficiently used to produce SA. However, the production profiles obtained suggest that fed-batch fermentations could also yield good results with this strain. The fact that the SA production ceases only when the glucose is exhausted suggests that higher titers could be achieved by adequate Glc feeding strategies and improvements in the acetate uptake capabilities of this strain, considering that it lacks PTS and could co-utilize Glc and acetate in these growing conditions. The fermentations reported here yielded an elevated ratio of SA to other byproducts of the pathway (10 times more SA than the main byproducts generated). Besides increasing the SA yield, this behavior is relevant for purifying SA from the culture broth. In fact, preliminary experiments concerning SA purification from the broth obtained from these cultures, resulted in an almost quantitative purification process (unpublished results). Furthermore, the highest yield of total aromatic compounds obtained represents 67% of the theoretical maximum, demonstrating the efficient redirection of carbon to the AAA pathway by strain AR36. Nevertheless, the relatively low cellular concentration present in the cultures, even when administering high concentrations of YE, represents a significant problem to this production system because it restricts the productivity of the process. Other strategies need to be utilized in order to increase the biomass concentration without increasing the supplemented YE, which will constitute an important improvement for scaling-up the process. Minimizing the metabolic load imposed by a high-copy plasmid while maintaining a sufficient gene dosage of the operon, should improve the distribution of resources that are directed towards biomass generation and SA production.

Methods

Construction of *Escherichia coli* derivatives and plasmids

The laboratory-evolved strain PB12, a derivative of PB11 (obtained by the inactivation of PTS in strain JM101 [41]), was the receptor of the genetic modifications described in this work [21,22,24]. The strains and plasmids used in this report are listed in Table 1. The chromosomal inactivations of *aroK*, *aroL*, and *lacI* genes were performed sequentially by homologous recombination of PCR products [42]. In all cases, plasmid pKD3 was used as a PCR template in conjunction with tailored oligonucleotides containing 45 bp homology with the target chromosomal sequence (Additional file 2). Plasmid pKD46 expressed the

Red recombinase system of bacteriophage lambda, and plasmid pCP20 allowed removal of the chloramphenicol resistance cassette after each event [43]. Every step was verified by PCR, identifying the clones that presented the expected amplicon sizes when using different sets of oligonucleotides and chromosomal DNA as a template (Additional file 2). The *pykF* inactivated gene was transduced to PB12 *aroK aroL*⁻ using a P1 phage lysate obtained from strain PB28 (*pykF*::gen) [44]. Transductants were selected on gentamycin plates (10 µg/ml), and the inactivation was confirmed by PCR.

The construction of the Aro6 operon and expression vector was accomplished in several steps (Figure 2 and Additional file 3). First, *aroB*, *aroG*^{fbr}, *tktA* and *aroE* genes were amplified by PCR using Pfu DNA polymerase and ligated sequentially into the polylinker of plasmid pBRINT-Ts Cm [45]. Chromosomal DNA from strain JM101 was used as a template for amplification of the required genes, with the exception of *aroG*^{fbr}, which was amplified from plasmid pJLBaroG^{fbr}*tktA* [46]. Different sets of oligonucleotides were employed for the amplification of each gene (Additional file 2), which also generated flanking restriction sites and consensus Shine-Dalgarno sequences (AGGAGG) situated 8 bp upstream of the start of each coding sequence. The PCR products were inserted into the polylinker in the following order: *aroB* in the *SmaI* site, *aroG*^{fbr} in the *XhoI* site, *tktA* in the *EcoRV* site, and *aroE* in the *ApaI* site. Simultaneously, plasmid pTrc327*parlacI*⁺ (Additional file 3) was built by ligating a PCR-amplified fragment containing the *lacI* gene, Trc promoter, polylinker, and transcriptional terminators of pTrc99A [31], into the *ScaI* and *PvuI* sites of pBR327*par* [32]. The 4-gene operon present in plasmid pBRINT-Ts Cm was amplified by PCR with a unique set of oligonucleotides (Additional file 1) and ligated into pTrc327*parlacI*⁺ after digesting both with *SacI* and *XbaI*. Later, *aroD* was amplified by PCR (flanked by *NheI* sites) and ligated into compatible *XbaI* site of pTrc327*parlacI*⁺. Because of our interest in expressing the operon in a constitutive manner, a *lacI* derivative of the initial pTrc327*parlacI*⁺ plasmid (without synthetic operon) was generated, and called pTrc327*par* (Additional file 3). The 5-gene operon was then transferred into *SacI* and *NcoI* sites of pTrc327*par*, giving rise to pTrcAro5. Finally, the *zwf* gene was inserted into the *XbaI* site, creating a 6-gene operon in pTrc327*par*. The resulting plasmid was named pTrcAro6 and transformed into AR2 and AR3 strains, generating AR26 and AR36, respectively (Table 1). The transformed strains were selected in LB plates supplemented with tetracycline (30 µg/ml).

Each step in the gene cloning and plasmid construction schemes was screened by endonuclease digestion and PCR, visualized with gel electrophoresis, and verified by DNA sequencing (3730, Perkin-Elmer/Applied Biosystems, USA). All the enzymes and reagents used in the molecular biology

procedures were purchased from Fermentas (USA) and New England Biolabs (USA). When required, kits for the purification of PCR, plasmid, and agarose-embedded DNA were utilized (Roche, Switzerland). TOP10 cells (Invitrogen, USA) were used as a host for screening of DNA ligations during intermediate steps in vector construction.

Cultivation media and growth conditions

Composition of production medium

SA production medium (adjusted to pH 7.0 with 10 N NaOH) contained per liter: K_2HPO_4 (7.5 g), KH_2PO_4 (7.5 g), citric acid monohydrate (2.1 g), ammonium iron (III) citrate (0.3 g), concentrated H_2SO_4 (1.2 ml), $MgSO_4$ (0.64 g), $CaCl_2$ (0.06 g), $(NH_4)_6(Mo_7O_{24})$ (0.0037 g), $ZnSO_4$ (0.0029 g), H_3BO_3 (0.0247 g), $CuSO_4$ (0.0025 g), $MnCl_2$ (0.0158 g), $CoCl_2$ (0.00129 g), thiamine (0.001 g), and betaine (0.234 g) as an osmoprotectant. Tetracycline (30 μ g/ml) was added to inocula and cultures whenever needed for plasmid maintenance. Glucose (filter-sterilized) and yeast extract (added before autoclaving) were supplied at the concentrations indicated for each experiment. The glucose was purchased from JT Baker (USA) and the autolysed yeast extract from BD Difco (USA).

Shake flask cultures

The inoculum preparation for the shake flask cultures started by the addition of 1 ml frozen aliquots to 250 ml shake flasks containing 25 ml of production medium supplemented with glucose (25 g/L) and yeast extract (15 g/L). The inoculum was grown at 37°C and 300 rpm until mid-exponential phase, and approximately 5% of the final volume was transferred to the test shake flasks and incubated under the same controlled conditions with media containing 15 g/L of glucose and 5 g/L of yeast extract. Cell growth was measured by monitoring the optical density at 600 nm (OD_{600}) in a DU700 spectrophotometer (Beckman, USA), and samples were taken periodically, centrifuged, and the supernatant was stored at -20°C for metabolite analysis. These experiments were performed at least in triplicate. All cultures started at approximately 0.3 OD_{600} .

Fermentor cultures

Batch cultures were performed at least in duplicate using 1 L autoclavable glass bioreactors (Applikon, The Netherlands) with 500 ml of working volume. Bioreactors were connected to an Applikon ADI 1010 BioController and ADI 1025 controllers to monitor temperature, pH, impeller speed, and dissolved oxygen tension (DOT). The pH was kept at 7.0 by the addition of H_3PO_3 (3.3%) and NH_4OH (10%). DOT in the culture medium was maintained by a continuous supply of filtered air (1 vvm), and by manually controlling the impeller speed (ranging from 500 to 1000 rpm) to ensure that DOT was

kept above 20% at all times. The inoculum preparation for the fermentors started by the addition of 1 ml frozen aliquots to 500 ml shake flasks containing 50 ml of production medium supplemented with glucose (25 g/L) and yeast extract (15 g/L). The strains were grown at 37°C and 300 rpm until mid-exponential phase and approximately 5% of the final volume was transferred from each inoculum to previously prepared bioreactors containing the production medium. All fermentations were performed in presence of tetracycline (30 μ g/ml). Cell growth was measured by monitoring optical density at 600 nm (OD_{600}) in a spectrophotometer (DU700, Beckman, USA), and samples were taken periodically, centrifuged, and the supernatant was stored at -20°C for metabolite analysis. All the fermentations started at approximately 0.3 OD_{600} .

Metabolite quantification

The supernatant from each sample was properly diluted and filtered through 0.45 μ m nylon membranes. Shikimic acid (SA), 3-dehydroshikimic acid (DHS), 3-dehydroquinic acid (DHQ), quinic acid (QA), gallic acid (GA), acetic acid, and glucose (Glc) concentrations were determined by HPLC using a Waters system (600E quaternary pump, 717 automatic injector, 2410 refraction index, and 996 photodiode array detectors; USA) equipped with an Aminex HPX-87H column (300 \times 7.8 mm; 9 μ m; Bio-Rad, USA). The mobile phase was 5 mM H_2SO_4 , with a flow rate of 0.5 ml/min, maintained at 50°C. 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) concentrations were determined colorimetrically by the thiobarbituric acid assay [47]. This method does not distinguish between DAHP and its unphosphorylated form, DAH, therefore in this work DAHP levels correspond to the sum of both compounds.

Data analysis and calculations

The measured concentrations of metabolites and biomass were normalized to the starting volume conditions to account for changes derived from pH control in fermentors. Data from independent experiments were averaged and presented in the corresponding graphs, where the error bars indicate the standard deviation for each point. Biomass concentration (X) was determined with a calibration curve between dry cellular weight and OD_{600} , resulting in the equation $X = 0.3587 \cdot OD_{600}$. Specific growth rate (μ) was determined by linearly fitting the biomass concentration to time during exponential phase with the following equation: $\ln X = \ln X_0 + \mu \cdot t$ (where t is time, and X_0 is the biomass concentration at initial time), displaying R^2 values >0.97 . The yield of SA from Glc ($Y_{SA/Glc}$) was calculated with the average molar concentrations of SA and Glc, produced and consumed, respectively, at the point of highest SA concentration. The yield of total aromatic compounds from glucose ($Y_{TAC/Glc}$) was calculated with the combined molar

yields of DAHP, DHQ, DHS, SA, QA, and GA at the point of highest SA concentration. The maximum theoretical yield of aromatic compounds was previously estimated as $0.86 \text{ mol}_{\text{TAC}}/\text{mol}_{\text{Glc}}$ for a PTS^- strain growing on glucose as only carbon source [10]. The global volumetric SA productivity ($Q_{p_{\text{global}}}$) and the global volumetric Glc consumption rate ($Q_{s_{\text{global}}}$) were calculated taking into account the time needed to reach the maximum SA concentration. Besides the previous calculations, linearizations were made to obtain apparent biomass on substrate ($Y_{X/S}$) and product on biomass ($Y_{P/X}$) yields. Although these apparent yields do not take into account the yeast extract consumption, correlation values for linearizations in all experiments were found to be >0.95 , allowing comparisons between them. These yields were used to calculate the specific productivity and specific consumption rate on the exponential phase ($q_{p_{\text{exp}}}$ and $q_{s_{\text{exp}}}$, respectively) with the following equations: $q_{p_{\text{exp}}} = Y_{P/X} \mu$; $q_{s_{\text{exp}}} = \mu/Y_{X/S}$.

The volumetric productivity and volumetric Glc consumption rate in stationary phase were determined by linearization of the first concentration data points at this stage versus time. The volumetric rates were utilized for calculation of specific production and consumption rates at stationary phase ($q_{p_{\text{sta}}}$ and $q_{s_{\text{sta}}}$, respectively) by dividing them by the average biomass concentration.

RNA extraction, cDNA synthesis and RT-qPCR analysis

Samples from batch fermentations of strain AR36 with 50 g/L of Glc and 15 g/L of YE were collected for RNA extraction at early exponential phase, EE (2 h ~ 1 OD_{600}), late exponential phase, LE (8 h ~ 12 OD_{600}), and stationary phase, ST (24 h ~ 17 OD_{600}), to determine gene expression levels. For comparison of the data, samples from early exponential phase (3.5 h ~ 1 OD_{600}) of strain AR3e (bearing an empty pTrc327par plasmid carrying only the tetracycline-resistance gene, Figure 2) cultured under the same conditions were also collected and processed. RNA was extracted using hot phenol equilibrated with water, and cDNA synthesis was performed using RevertAid H First Strand cDNA Synthesis kit (Fermentas, USA) and a mixture of specific DNA primers, as reported previously [21,24]. qPCR experiments were performed with the ABI Prism 7300 Real Time PCR System (Applied Biosystems, USA) using Maxima SYBRGreen PCR Master Mix (Fermentas, USA) and reaction conditions previously described [21,24]. The quantification technique used to compare data was the $2^{-\Delta\Delta\text{CT}}$ method [48] and the results were normalized using the *ihfB* gene as an internal control. The same reproducible expression level for this gene was detected in all the strains and conditions analyzed [24]. All qPCR experiments complied with the MIQE guidelines for publication of quantitative real-time PCR experiments [49]. Using cells from two separate

fermentations, RNA extraction and cDNA synthesis reactions were performed for each biological replicate at the indicated times and the gene expression values were measured by triplicate for each sample. Average values were graphed, with error bars representing standard deviation. Standard deviation was less than 30% in all cases.

Additional files

Additional file 1: Nucleotide sequence of the synthetic operon constructed in this work and present in plasmid pTrcAro6. The *aroG^{fb}* gene included in this construction was a gift from DuPont™-Genencor®, therefore its coding sequence cannot be disclosed. Each nucleotide of *aroG^{fb}* is indicated with an “n” except for the ones corresponding to its start and stop codons.

Additional file 2: Oligonucleotides utilized in this work.

Additional file 3: Detailed scheme of the construction of plasmid pTrcAro6.

Competing interests

The authors have declared no competing interests.

Authors' contributions

AR, JLB and FB conceived the study and designed the experiments. AR and JAM carried out all fermentation experiments and data analysis. JAM performed the calculation of fermentation parameters. NF carried out RNA extraction and RT-qPCR experiments. GH was involved in metabolite quantification and analysis. AR performed the molecular biology and strain construction procedures and wrote the manuscript. JAM, OTR, GG and FB critically revised the results and the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

The authors thank Ramón de Anda for technical support, José González for assistance with strain construction, Enrique Morett for helpful discussions, Paul Gaytán, Jorge Yáñez and Eugenio López for the synthesis of oligonucleotides and sequencing of DNA (Instituto de Biotecnología-UNAM, México), as well as Mercedes Enzaldo and Aurelia González for technical assistance. A. Rodríguez was supported by grants 215044 and 290604 from Consejo Nacional de Ciencia y Tecnología (CONACyT, México). This work was supported by grants 105782 and 177568 (CONACyT), 44126 and 126793 (FONSEC/SSA/ISSSTE/CONACyT), and IN224709 and IN205811 (DGAPA-PAPIIT-UNAM). The funding sources had no involvement in study design, collection and analysis of data, writing of the report or decision to publish.

Author details

¹Departamento de Ingeniería Celular y Biotatálisis, Instituto de Biotecnología, Universidad Nacional Autónoma de México (UNAM), Apdo. Postal 510-3, Cuernavaca, Morelos 62250, Mexico. ²Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos (UAEM), Av. Universidad 2000, Cuernavaca, Morelos 62250, Mexico. ³Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología, Universidad Nacional Autónoma de México (UNAM), Apdo. Postal 510-3, Cuernavaca, Morelos 62250, Mexico.

Received: 21 June 2013 Accepted: 28 August 2013

Published: 30 September 2013

References

1. Farina V, Brown JD: Tamiflu: the supply problem. *Angew Chem Int Ed Engl* 2006, **45**:7330–7334.
2. Kim HK, Park KJJ: A new efficient synthesis of oseltamivir phosphate (Tamiflu) from (-)-shikimic acid. *Tetrahedron Lett* 2012, **53**:1561–1563.
3. Nie LD, Ding W, Shi XX, Quan N, Lu X: A novel and high-yielding asymmetric synthesis of oseltamivir phosphate (Tamiflu) starting from (-)-shikimic acid. *Tetrahedron: Asymmetry* 2012, **23**:742–747.

4. Patnaik R, Spitzer RG, Liao JC: **Pathway engineering for production of aromatics in *Escherichia coli*: confirmation of stoichiometric analysis by independent modulation of AroG, TktA, and Pps activities.** *Biotechnol Bioeng* 1995, **46**:361–370.
5. Gosset G, Yong-Xiao J, Berry A: **A direct comparison of approaches for increasing carbon flow to aromatic biosynthesis in *Escherichia coli*.** *J Ind Microbiol* 1996, **17**:47–52.
6. Draths KM, Knop DR, Frost JW: **Shikimic acid and quinic acid: replacing isolation from plant sources with recombinant microbial biocatalysis.** *J Am Chem Soc* 1999, **121**:1603–1604.
7. Johansson L, Lindskog A, Silfversparre G, Cimander C, Nielsen KF, Lidén G: **Shikimic acid production by a modified strain of *E. coli* (W3110.shik1) under phosphate-limited and carbon-limited conditions.** *Biotechnol Bioeng* 2005, **92**:541–552.
8. Chen K, Dou J, Tang S, Yang Y, Wang H, Fang H, Zhou C: **Deletion of the *aroK* gene is essential for high shikimic acid accumulation through the shikimate pathway in *E. coli*.** *Bioresour Technol* 2012, **119**:141–147.
9. Flores N, Xiao J, Berry A, Bolivar F, Valle F: **Pathway engineering for the production of aromatic compounds in *Escherichia coli*.** *Nat Biotechnol* 1996, **14**:620–623.
10. Chandran SS, Yi J, Draths KM, Von Daeniken R, Weber W, Frost JW: **Phosphoenolpyruvate availability and the biosynthesis of shikimic acid.** *Biotechnol Prog* 2003, **19**:808–814.
11. Escalante A, Calderón R, Valdívila A, De Anda R, Hernández G, Ramírez OT, Gosset G, Bolivar F: **Metabolic engineering for the production of shikimic acid in an evolved *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.** *Microb Cell Fact* 2010, **9**:21.
12. Juminaga D, Baidoo EEK, Redding-Johanson AM, Bath TS, Burd H, Mukhopadhyay A, Petzold CJ, Keasling JD: **Modular engineering of L-tyrosine production in *Escherichia coli*.** *Appl Environ Microbiol* 2012, **78**:89–98.
13. Meza E, Becker J, Bolivar F, Gosset G, Wittmann C: **Consequences of phosphoenolpyruvate:sugar phosphotransferase system and pyruvate kinase isozymes inactivation in central carbon metabolism flux distribution in *Escherichia coli*.** *Microb Cell Fact* 2012, **11**:127.
14. Ahn JO, Lee HW, Saha R, Park MS, Jung JK, Lee DY: **Exploring the effects of carbon sources on the metabolic capacity for shikimic acid production in *Escherichia coli* using in silico metabolic predictions.** *J Microbiol Biotechnol* 2008, **18**:1773–1784.
15. Ghosh S, Chisti Y, Banerjee UC: **Production of shikimic acid.** *Biotechnol Adv* 2012, **30**:1425–1431.
16. Glick BR: **Metabolic load and heterologous gene expression.** *Biotechnol Adv* 1995, **13**:247–261.
17. Keasling JD: **Gene-expression tools for the metabolic engineering of bacteria.** *Trends Biotechnol* 1999, **17**:452–460.
18. Silva F, Queiroz JA, Domingues FC: **Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*.** *Biotechnol Adv* 2012, **30**:691–708.
19. Carneiro S, Ferreira EC, Rocha I: **Metabolic responses to recombinant bioprocesses in *Escherichia coli*.** *J Biotechnol* 2013, **164**:396–408.
20. Postma PW, Lengeler JW, Jacobson GR: **Phosphoenolpyruvate: carbohydrate phosphotransferase systems.** In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. 2nd edition. Edited by Neidhart FC. USA: ASM Press; 1996:1149–1174.
21. Flores N, Flores S, Escalante A, De Anda R, Leal L, Malpica R, Georgellis D, Gosset G, Bolivar F: **Adaptation for fast growth on glucose by differential expression of central carbon metabolism and *gal* regulon genes in an *Escherichia coli* strain lacking the phosphoenolpyruvate:carbohydrate phosphotransferase system.** *Metab Eng* 2005, **7**:70–87.
22. Flores N, Leal L, Sigala JC, De Anda R, Escalante A, Martínez A, Ramírez OT, Gosset G, Bolivar F: **Growth recovery on glucose under aerobic conditions of an *Escherichia coli* strain carrying a phosphoenolpyruvate: carbohydrate phosphotransferase system deletion by inactivating *arcA* and overexpressing the genes coding for glucokinase and galactose permease.** *J Mol Microbiol Biotechnol* 2007, **13**:105–116.
23. Martínez G, De Anda R, Hernández G, Escalante A, Gosset G, Ramírez OT, Bolivar F: **Couitilization of glucose and glycerol enhances the production of aromatic compounds in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system.** *Microb Cell Fact* 2008, **7**:1.
24. Aguilar C, Escalante A, Flores N, De Anda R, Riveros-McKay F, Gosset G, Morett E, Bolivar F: **Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system.** *BMC Genomics* 2012, **13**:385.
25. Flores S, Gosset G, Flores N, De Graaf AA, Bolivar F: **Analysis of carbon metabolism in *Escherichia coli* strains with an inactive phosphotransferase system by (¹³C) labeling and NMR spectroscopy.** *Metab Eng* 2002, **4**:124–137.
26. Cunningham DS, Liu Z, Domagalski N, Koepsel RR, Ataai MM, Domach MM: **Pyruvate kinase-deficient *Escherichia coli* exhibits increased plasmid copy number and cyclic AMP levels.** *J Bacteriol* 2009, **191**:3041–3049.
27. Ponce E: **Effect of growth rate reduction and genetic modifications on acetate accumulation and biomass yields in *Escherichia coli*.** *J Biosci Bioeng* 1999, **87**:775–780.
28. Zhu T, Phalakornkule C, Koepsel RR, Domach MM, Ataai MM: **Cell growth and by-product formation in a pyruvate kinase mutant of *E. coli*.** *Biotechnol Prog* 2001, **17**:624–628.
29. Kedar P, Colah R, Shimizu K: **Proteomic investigation on the *pyk-F* gene knockout *Escherichia coli* for aromatic amino acid production.** *Enzyme Microb Technol* 2007, **41**:455–465.
30. Al Zaid Siddiquee K, Arauzo-Bravo MJ, Shimizu K: **Metabolic flux analysis of *pykF* gene knockout *Escherichia coli* based on 13 C-labeling experiments together with measurements of enzyme activities and intracellular metabolite concentrations.** *Appl Microbiol Biotechnol* 2004, **63**:407–417.
31. Amann E, Ochs B, Abel K: **Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*.** *Gene* 1988, **69**:301–315.
32. Zurita M, Bolivar F, Soberon X: **Construction and characterization of new cloning vehicles VII. Construction of plasmid pBR327_{par}, a completely sequenced, stable derivative of pBR327 containing the *par* locus of pSC101.** *Gene* 1984, **28**:119–122.
33. Dell KA, Frost JW: **Identification and removal of impediments to biocatalytic synthesis of aromatics from D-Glucose: rate-limiting enzymes in the common pathway of aromatic amino acid biosynthesis.** *J Am Chem Soc* 1993, **115**:11581–11589.
34. Oldiges M, Kunze M, Degenring D, Sprenger GA, Takors R: **Stimulation, monitoring, and analysis of pathway dynamics by metabolic profiling in the aromatic amino acid pathway.** *Biotechnol Prog* 2004, **20**:1623–1633.
35. Lütke-Eversloh T, Stephanopoulos G: **Combinatorial pathway analysis for improved L-tyrosine production in *Escherichia coli*: identification of enzymatic bottlenecks by systematic gene overexpression.** *Metab Eng* 2008, **10**:69–77.
36. Flores S, De Anda-Herrera R, Gosset G, Bolivar F: **Growth-rate recovery of *Escherichia coli* cultures carrying a multicopy plasmid, by engineering of the pentose-phosphate pathway.** *Biotechnol Bioeng* 2004, **87**:485–494.
37. Sprenger GA: **Genetics of pentose-phosphate pathway enzymes of *Escherichia coli* K-12.** *Arch Microbiol* 1995, **164**:324–330.
38. Knop DR, Draths KM, Chandran SS, Barker JL, Von Daeniken R, Weber W, Frost JW: **Hydroaromatic equilibration during biosynthesis of shikimic acid.** *J Am Chem Soc* 2001, **123**:10173–10182.
39. Krämer M, Bongaerts J, Bovenberg R, Kremer S, Müller U, Orf S, Wubbolts M, Raeven L: **Metabolic engineering for microbial production of shikimic acid.** *Metab Eng* 2003, **5**:277–283.
40. Gimenez R, Nuñez MF, Badia J, Aguilar J, Baldoma L: **The gene *yjcG*, cotranscribed with the gene *acs*, encodes an acetate permease in *Escherichia coli*.** *J Bacteriol* 2003, **185**:6448–6455.
41. Messing J: **Recombinant DNA technical bulletin.** *NIH Publ* 1979, **2**:43–48.
42. Datsenko K, Wanner B: **One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products.** *Proc Natl Acad Sci U S A* 2000, **97**:6640–6645.
43. Cherepanov PP, Wackernagel W: **Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant.** *Gene* 1995, **158**:9–14.
44. Ponce E, Flores N, Martínez A, Valle F, Bolivar F: **Cloning of the two pyruvate kinase isoenzyme structural genes from *Escherichia coli*: the relative roles of these enzymes in pyruvate biosynthesis.** *J Bacteriol* 1995, **177**:5719–5722.
45. Le Borgne S, Palmeros B, Valle F, Bolivar F, Gosset G: **pBRINT-Ts: a plasmid family with a temperature-sensitive replicon, designed for chromosomal integration into the *lacZ* gene of *Escherichia coli*.** *Gene* 1998, **223**:213–219.

46. Balderas-Hernández VE, Sabido-Ramos A, Silva P, Cabrera-Valladares N, Hernández-Chávez G, Baez-Viveros JL, Martínez A, Bolívar F, Gosset G: **Metabolic engineering for improving anthranilate synthesis from glucose in *Escherichia coli*.** *Microb Cell Fact* 2009, **8**:19.
47. Weissbach A, Hurwitz J: **The Formation of 2-Keto-3-deoxyheptonic Acid in Extracts of *Escherichia coli* B.** *J Biol Chem* 1958, **234**:705–709.
48. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method.** *Methods* 2001, **25**:402–408.
49. Bustin SA, Benes V, Garson JA, Hellems J, Hugget J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: **The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments.** *Clin Chem* 2009, **55**:611–622.

doi:10.1186/1475-2859-12-86

Cite this article as: Rodriguez et al.: Constitutive expression of selected genes from the pentose phosphate and aromatic pathways increases the shikimic acid yield in high-glucose batch cultures of an *Escherichia coli* strain lacking PTS and *pykF*. *Microbial Cell Factories* 2013 **12**:86.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



REVIEW

Open Access

Engineering *Escherichia coli* to overproduce aromatic amino acids and derived compounds

Alberto Rodriguez, Juan A Martínez, Noemí Flores, Adelfo Escalante, Guillermo Gosset and Francisco Bolivar*

Abstract

The production of aromatic amino acids using fermentation processes with recombinant microorganisms can be an advantageous approach to reach their global demands. In addition, a large array of compounds with alimentary and pharmaceutical applications can potentially be synthesized from intermediates of this metabolic pathway. However, contrary to other amino acids and primary metabolites, the artificial channelling of building blocks from central metabolism towards the aromatic amino acid pathway is complicated to achieve in an efficient manner. The length and complex regulation of this pathway have progressively called for the employment of more integral approaches, promoting the merge of complementary tools and techniques in order to surpass metabolic and regulatory bottlenecks. As a result, relevant insights on the subject have been obtained during the last years, especially with genetically modified strains of *Escherichia coli*. By combining metabolic engineering strategies with developments in synthetic biology, systems biology and bioprocess engineering, notable advances were achieved regarding the generation, characterization and optimization of *E. coli* strains for the overproduction of aromatic amino acids, some of their precursors and related compounds. In this paper we review and compare recent successful reports dealing with the modification of metabolic traits to attain these objectives.

Keywords: Aromatic compounds, *Escherichia coli*, Metabolic engineering, Systems biotechnology, Synthetic biology, Shikimate pathway, Phenylalanine, Tyrosine, Tryptophan

Introduction

The aromatic amino acids (AAA), L-tryptophan (L-TRP), L-phenylalanine (L-PHE) and L-tyrosine (L-TYR), are the final products of the aromatic biosynthetic pathway comprising the shikimate (SHK) pathway, which connects central carbon metabolism (CCM) with the biosynthesis of chorismate (CHA), the last common precursor in the terminal branches for AAA biosynthesis (Figure 1) [1,2]. These pathways are present in bacteria and in several eukaryotic organisms such as ascomycetes fungi, apicomplexans, and plants [3,4]. The AAA are essential components in the diet of higher animals and humans, hence they are used as dietary supplements (e.g. diet of swine and poultry consisting of grains of corn and soybean is low in L-TRP) and key precursors of industrial and pharmaceutical compounds (e.g. L-PHE is the key ingredient in the synthesis of the artificial sweetener

aspartame, whereas L-TYR is an essential dietary component for phenylketonuria patients as the starter material for L-DOPA or melanin production) [5]. The annual worldwide production of amino acids is estimated to be above 4.5 million tons/year, with a market growth for most amino acids of ~10% and higher [6,7]. Among the aromatic amino acids, L-TRP has a market size of more than 14,000 tons/year [8] and the production of L-PHE exceeds 30,000 tons/year [9].

It is well established that the production of high-valued commodities can be performed cost-efficiently by the rational design, modification and cultivation of a recombinant microorganism. In particular, the development of efficient microbial processes for accumulation of compounds derived from the AAA biosynthetic pathway has not been an easy task for metabolic and bioprocess engineers. For more than 20 years, considerable efforts have been directed towards characterizing and purposely overriding the naturally tight metabolic regulation of this pathway. These continued efforts have relied on knowledge obtained from pioneer works on the biosynthesis of aromatic compounds by the

* Correspondence: bolivar@ibt.unam.mx
Departamento de Ingeniería Celular y Biotecnología, Instituto de Biotecnología, Universidad Nacional Autónoma de México (UNAM), Avenida Universidad 2001, Col. Chamilpa, Cuernavaca, Morelos 62210, México

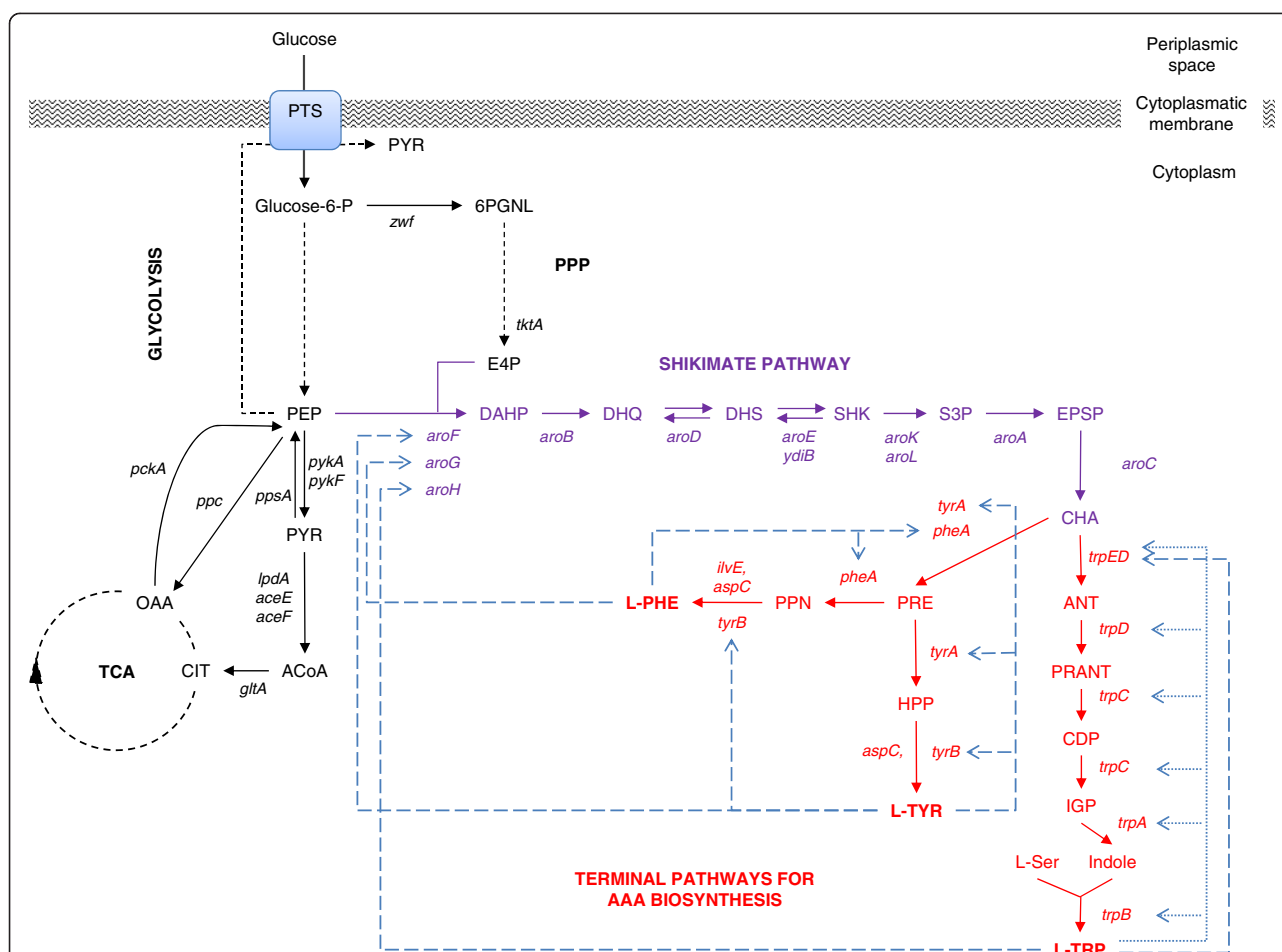


Figure 1 Schematic representation of the AAA pathway in *Escherichia coli* including its transcriptional and allosteric regulatory control circuits. Central carbon metabolism intermediates and genes shown: PPP (pentose phosphate pathway); TCA (tricarboxylic acid cycle); E4P (erythrose-4-P); PGNL (6-phospho D-glucono-1,5-lactone); PEP (phosphoenolpyruvate); PYR (pyruvate); ACoA (acetyl-CoA); CIT (citrate); OAA (oxaloacetate); *zwf* (glucose 6-phosphate-1-dehydrogenase); *tktA* (transketolase I); *pykA*, *pykF* (pyruvate kinase II and pyruvate kinase I, respectively); *lpdA*, *aceE*, and *aceF* (coding for PYR dehydrogenase subunits); *gltA* (citrate synthase); *pckA* (PEP carboxykinase); *ppc* (PEP carboxylase); *ppsA* (PEP synthetase). Shikimate pathway intermediates and genes shown: DAHP (3-deoxy-D-arabino-heptulosonate-7-phosphate); DHQ (3-dehydroquininate); DHS (3-dehydroshikimate); SHK (shikimate); S3P (SHK-3-phosphate); EPSP (5-enolpyruvyl-shikimate 3-phosphate); CHA (chorismate); *aroF*, *aroG*, *aroH* (DAHP synthase AroF, AroG and AroH, respectively); *aroB* (DHQ synthase); *aroD* (DHQ dehydratase); *aroE*, *ydiB* (SHK dehydrogenase and SHK dehydrogenase / quininate dehydrogenase, respectively); *aroA* (3-phosphoshikimate-1-carboxyvinyltransferase); *aroC* (CHA synthase). Terminal AAA biosynthetic pathways intermediates and genes shown: ANT (anthranilate); PRANT (*N*-5-phosphoribosyl-anthranilate); CDP (1-(*o*-carboxyphenylamino)-1'-deoxyribulose 5'-phosphate); IGP ((1*S*,2*R*)-1-C-(indol-3-yl)glycerol 3-phosphate); *trpE*, *trpD* (ANT synthase component I and II, respectively); *trpC* (indole-3-glycerol phosphate synthase / phosphoribosylanthranilate isomerase); *trpA* (indoleglycerol phosphate aldolase); *trpB* (tryptophan synthase); PRE (prephenate); PPN (phenylpyruvate); HPP (4-hydroxyphenylpyruvate); *tyrA*, *pheA* (TyrA and PheA subunits of the CHA mutase, respectively); *ivlE* (subunit of the branched-chain amino acid aminotransferase); *aspC* (subunit of aspartate aminotransferase); *tyrB* (tyrosine aminotransferase). Continuous arrows show single enzymatic reactions, black dashed arrows show several enzymatic reactions, long-dashed blue arrows indicate allosteric regulation and dotted blue arrows indicate transcriptional repression. Adapted from EcoCyc database [1].

groups of B.D. Davis, F. Gibson, C. Yanofsky, A.J. Pittard, K.M. Herrmann and J.W. Frost, among others, whose contributions have been comprehensively reviewed in the past [2,10-12].

Recently, the availability of omics-scale data has allowed significant advances in metabolic reconstruction and modeling, resulting in better strain development [13]. Likewise, the increased use of combinatorial and evolutionary approaches, fueled by a rapid expansion of synthetic

molecular tools, opened the possibility for testing novel and large combinations of gene expression systems and genetic backgrounds [14,15]. Additionally, efforts concerning the optimization of fermentation conditions have succeeded in scaling-up many AAA production processes, while simultaneously providing important feedback on the physiological behavior of engineered strains [16,17]. However, the availability of operational tools and techniques, as well as the amount of physiological and molecular information,

are unevenly distributed among the microorganisms currently used for the production of AAA. These circumstances have contributed to positioning *E. coli* as the organism with most reported success cases and has resulted in a wide array of well-characterized production strains [18,19].

In this paper we review some notable advances in the generation, characterization and optimization of *E. coli* strains for the overproduction of AAA, some of their important precursors and related compounds. Although these studies were classified in accordance to the main schemes employed for each case, the constant expansion and complementarity of such approaches has encouraged scientists to apply a systems-based perspective [20,21]. Therefore, recent and representative works on the subject using different strategies were selected and discussed.

Engineering of the CCM: glucose transport, glycolytic, gluconeogenic, and pentose phosphate pathways

Successful metabolic engineering efforts for the generation of *E. coli* strains that can overproduce AAA include: (i) increasing the availability of the direct precursors phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P); (ii) enhancement of the first enzymatic reaction in the SHK pathway to yield 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP); (iii) improving the carbon flow through the biosynthetic pathway of interest by removal of transcriptional and allosteric regulation; (iv) identifying and relieving rate-limiting enzymatic reactions; (v) preventing loss of carbon flow towards competing pathways; (vi) enhancement of product export; and (vii) prevention of product degradation or re-internalization.

Regarding PEP metabolism, *E. coli* uses the phosphotransferase system (PTS) as the main system for the translocation and phosphorylation of glucose from the periplasmic space to the cytoplasmic environment, consuming one PEP molecule which is converted to pyruvate (PYR) [22,23]. This reaction yields one molecule of glucose-6-phosphate which is catabolized by the glycolytic pathway, resulting in two PEP molecules (Figure 1). PEP is a precursor feeding several biosynthetic pathways and also participates in ATP generation, either by substrate-level phosphorylation of ADP or indirectly as an acetyl coenzyme-A (ACoA) precursor. When *E. coli* grows in mineral broth containing glucose as the sole carbon source the PTS consumes 50% of the available PEP, whereas the reactions catalyzed by other enzymes such as PEP carboxylase, PYR kinases, UDP-N-acetylglucosamine enolpyruvyl transferase, and DAHP synthases (DAHPS), consume approximately 16%, 15%, 16%, and 3% of remaining PEP, respectively [23,24]. Therefore, PEP can be converted to PYR by PTS and PYR kinases I and II (coded by *pykF* and *pykA* respectively), and PYR is converted to ACoA by the PYR dehydrogenase

multienzyme complex (coded by *aceE*, *aceF* and *lpd*), a reaction connecting the glycolytic pathway with the tricarboxylic acid cycle (TCA) [1]. Moreover, PEP and PYR are key intermediates of the CCM as they are substrate of at least six enzymes which determine the metabolic fate of these intermediates (biosynthetic/catabolic pathways and glycolytic/gluconeogenic capabilities of the cell): DAHPS isoenzymes (AroF, AroG, and AroH coded by *aroF*, *aroG* and *aroH*, respectively) [3,11]; PYR kinases I and II; PEP synthetase (PpsA coded by *ppsA*); PEP carboxylase (Ppc, coded by *ppc*); and PEP carboxykinase (PckA coded by *pckA*) [25] (Figure 1).

Detailed knowledge of these nodes permitted the development of strategies that allowed higher PEP availability for the biosynthesis of aromatic compounds, including the replacement of glucose transport and phosphorylation capabilities of the PTS by alternative enzymes such as the glucose facilitator and glucokinase from *Zymomonas mobilis* (coded by *glf* and *glk*, respectively) [26-28], the galactose permease and glucokinase from *E. coli* (coded by *galP* and *glk*, respectively) [29,30], or the use of an adaptive evolution process to select PTS⁻ derivatives growing at high specific growth rates (μ) on glucose [31,32]. Additionally, high PEP availability has been achieved by modulation of the carbon flux from PEP to the TCA caused by the inactivation of one or both of the PYR kinases [33,34], as well as improving the recycling of PYR to PEP by a plasmid-encoded copy of PEP synthetase [35-37]. The overexpression of *pckA*, in combination with an enhanced carbon flow through the glyoxylate shunt, has also been proposed as a strategy to increase the yield of aromatic compounds [38,39]. An alternative approach to increase PEP is the attenuation of CsrA, a regulatory protein of carbohydrate metabolism, either by direct gene knockout or by increasing the expression of its negative regulatory RNA, coded by *csrB* [40,41].

On the other hand, E4P is a metabolite that participates in reversible reactions present in the non-oxidative branch of the pentose phosphate pathway (PPP), as well as a substrate in irreversible reactions that lead to the production of aromatic amino acids or vitamin B₆ [42]. E4P can also be directly produced from sedoheptulose-1,7-bisphosphate in a reaction that is probably favored when the intracellular levels of sedoheptulose-7-phosphate are high [43]. Metabolic engineering reports have shown that a considerable increase in availability of E4P (inferred by the increased production of aromatic compounds and pathway intermediates, such as DAHP) can be achieved by overexpression of genes coding for a transketolase (*tktA*) [35,44-46] or a transaldolase (*talB*) [26,47]. Additional attempts to increase the carbon flow towards the PPP for enhanced production of aromatic compounds include the use of mutants lacking the enzyme phosphoglucose isomerase [48,49], the overexpression of enzyme

glucose-6-phosphate dehydrogenase [41,50], or the use of multiple carbon sources, mainly hexoses, pentoses and glycerol [51-54]. After an adequate supply of precursors has been established, it is essential to commit this carbon towards the SHK pathway and to remove control points and limiting steps to increase the production of target compounds.

Deregulation of the AAA pathway: identifying and relieving rate-limiting steps

In *E. coli*, the DAHPS isoenzymes AroG, AroF and AroH contribute to the total DAHPS activity and are subjected to allosteric control by L-PHE, L-TYR and L-TRP, respectively (Figure 1). AroG contributes about 80% of the overall DAHPS activity, AroF about 15%, and the remaining activity corresponds to AroH DAHPS [3,11]. Both the AroG and AroF isoenzymes are completely inhibited by about 0.1 mM of the corresponding amino acids, but AroH is only partially inhibited by L-TRP. Apparent inability of L-TRP to totally inhibit this isoenzyme is proposed to be a mechanism to ensure a sufficient supply of CHA for the biosynthesis of other aromatic compounds when AAA are present in excess in the growth medium [3]. Specific amino acid residues involved in the allosteric sites have been identified by structural analysis of feedback-insensitive mutant enzymes, resulting in the targeted generation of the feedback resistant (*fbr*) variants AroG^{fbr} and AroF^{fbr} [28,31,55]. Additionally to allosteric control of DAHPS isoenzymes, their transcriptional expression can be controlled by the *tyr*- and *trp*- repressors complexed with the AAA [3,11].

Consequently, amplification and deregulation of DAHPS activity is an essential strategy to overproduce aromatic compounds and its precursor SHK. Introduction of plasmid-encoded copies of *aroF*^{fbr} and *aroG*^{fbr} combined with additional plasmid-cloned gene *tktA*, or their chromosomal integrations in gene clusters, have resulted in increased carbon flow from the CCM to the SHK pathway for the production of L-PHE [11,55,56], L-TYR [5,57,58] and L-TRP [59-61]. Positive results were also obtained with the insertion of an *aroG*^{fbr} gene into the chromosome of an L-PHE producing strain while being controlled by a promoter that is active during late cultivation stages, in order to counteract the fall of DAHPS activity in stationary phase [62].

Further increases in carbon flux through the SHK pathway have been attained by the removal of transcriptional and allosteric control points and by relieving limiting enzymatic reactions [2,11,19,23]. The reactions catalyzed by DHQ synthase (encoded by *aroB*) and SHK kinase isoenzymes I and II (encoded by *aroK* and *aroL*, respectively) are considered as rate-limiting [63-65]. In addition, the reaction catalyzed by the enzyme quinate/shikimate dehydrogenase (coded by *ydiB*) was also reported as

limiting in the development of L-TYR production strains [58]. Either the overexpression of some of these genes by plasmid-cloned copies [28,66], their co-expression in a modular operon under control of diverse promoters [50,58,67], or their expression by chromosomal integration of additional gene copies and promoter engineering by chromosomal evolution [68], have relieved to a great extent these rate-limiting steps typically encountered during the development of SHK and AAA overproducing strains (Table 1). To date, genetically modified *E. coli* strains can overproduce SHK from glucose with yields in the range of 0.08 to 0.42 mol SHK / mol glucose under diverse culture conditions [28,50,68-70]. SHK is a key intermediate of the common biosynthetic aromatic pathway (Figure 1) gaining relevance as the substrate for the chemical synthesis of the drug oseltamivir phosphate, known commercially as Tamiflu®, an efficient inhibitor of the surface protein neuraminidase of seasonal influenza, avian influenza H5N1, and human influenza H1N1 viruses [71-74].

In addition to modifications in the SHK pathway, metabolic engineering approaches to overproduce L-TYR typically include alterations in TyrR and/or *trp* regulons. The TyrR regulon comprises diverse essential genes implicated in AAA biosynthesis and transport [1,75]. TyrR acts as a dual transcriptional activator and repressor; however, the repression mechanism requires the ATP-dependent binding of AAA to the central protein domain. L-TYR is the major effector of TyrR-mediated repression, although some repression occurs with L-PHE as co-repressor for *aroE*, *aroL*, *tyrP* (coding for a L-TYR specific permease), *aroP* (coding for an aromatic amino acid permease) and *aroG* genes, whereas activation does not apparently involve an ATP-dependent binding of aromatic amino acids [1,2,5,11]. Inactivation of TyrR-mediated regulation by deletion of *tyrR* and overexpression of *aroG*^{fbr} and *tyrA*^{fbr}, combined with the overexpression of CCM genes (e.g. *ppsA* and *tktA*) and genes of the L-TYR biosynthetic pathway (e.g. *tyrB*, *aroC*, *aroA*) have improved the production of L-TYR in diverse *E. coli* strains [5,11,58].

Similar results were obtained for L-PHE in resting cells by overexpression of a feedback-resistant or an evolved (*ev*) CHA mutase/prephenate dehydratase enzymes (coded by *pheA*^{fbr} and *pheA*^{ev}, respectively) [55,76]. The bifunctional enzyme chorismate mutase/prephenate dehydrogenase TyrA, catalyzes the shared first step in L-PHE and L-TYR final biosynthetic pathways (the conversion of CHA to prephenate), as well as the second step in L-TYR biosynthesis (the subsequent NAD⁺-dependent oxidative decarboxylation of prephenate to 4-hydroxyphenylpyruvate) (Figure 1). TyrA catalyzes both reactions in separate domains of the protein and the CHA mutase/prephenate dehydrogenase is feedback-inhibited by L-TYR (up to 95% inhibition of the prephenate dehydrogenase and 45% of the CHA mutase activity) [1,2]. The bifunctional enzyme CHA

Table 1 Relevant *E. coli* strains engineered for the overproduction of compounds derived from the aromatic biosynthetic pathway

Strain	Relevant characteristics	Main compound produced (titer ^a , and/or yield ^b , ^c). Relevant culture conditions	References
SP1.1pts/pSC6.090B (RB791 derivative)	$\Delta ptsHlcr \Delta aroK \Delta aroL serA::aroB$ / (plasmid) $aroF^{fbr} tktA$, $P_{tac} aroE serA$, $P_{tac} glf glk^c$	SHK (84, 0.33 ^b). 10 L fed-batch reactors with glucose, AAA and 15 g/L of yeast extract	[28]
AR36 (JM101 derivative)	$\Delta ptsHlcr \Delta aroK \Delta aroL \Delta lacI \Delta pykF$ / (plasmid) $P_{trc} aroB tktA aroG^{fbr} aroE aroD zwf$	SHK (43, 0.42 ^b). 1 L batch reactors with 100 g/L of glucose and 30 g/L of yeast extract	[50]
SA116 (BW25113 derivative)	$\Delta aroK \Delta aroL P_{pps::P_{lacQ1}, P_{csrB::P_{lacQ1}}$ / (chromosome) $aroG^{fbr} tktA aroB aroE$, $P_{T5} ppsA csrB$, $5P_{tac} tktA nadK$	SHK (3, 0.33 ^b). Medium supplemented with 10 g/L of glucose, 1 g/L of peptone and 1 g/L of proline	[68]
W14/pR15BABKG (W3110 derivative)	$\Delta crr \Delta tyrA$ / (plasmid) $P_R aroG15 tyrB$, $P_L pheA^{fbr} ydiB aroK yddG$	L-PHE (47, 0.25 ^d). 15 L fed-batch reactors with glucose and 1 g/L of tyrosine	[132]
FUS4.11/pF81 _{kan} (W3110 derivative)	$\Delta pheA \Delta tyrA \Delta aroF \Delta lacIZYA \Delta pykA \Delta pykF$ / (chromosome) $P_{tac} aroF aroB aroL$, (plasmid) $P_{tac} pheA^{fbr} aroF aroB aroL$	L-PHE (13, 0.15 ^d). 15 L multi-phase fed-batch reactors with glycerol and lactic acid	[131]
BL21 (DE3)	(plasmid) containing the phenylalanine dehydrogenase gene of <i>Acinetobacter lwoffii</i>	L-PHE (5, 0.58 ^d) 2 L batch reactors with 10 g/L of glycerol	[130]
MG1655 derivative	(plasmid) $P_{lacUV5} aroE aroD aroB^{op}$, $P_{L-tetO1} aroG^{fbr} ppsA tktA$, (plasmid) $P_{lacUV5} tyrB tyrA^{fbr} aroC$, $P_{trc} aroA aroL$	L-TYR (2, 0.44 ^d). Shake flask cultures with 5 g/L of glucose	[58]
<i>rpoA14</i> ^R (K-12 derivative)	$\Delta pheA \Delta tyrR$ / (chromosome) $P_L tyrA^{fbr} aroG^{fbr}$, point mutations in <i>hisH</i> and <i>purF</i> , (plasmid) <i>rpoA</i>	L-TYR (14, 0.12 ^d) 2 L fed-batch reactors with glucose	[101]
MG1655 derivative	$\Delta pheA \Delta pheL$ / (chromosome) $P_{trc} tyrA$	L-TYR (55, 0.30 ^d). 200 L fed-batch reactors with glucose	[17]
FB-04/pSV03 (W3110 derivative)	$\Delta trpR \Delta tnaA \Delta pheA \Delta tyrA$ / (plasmid) $aroF^{fbr} trpE^{fbr} D$	L-TRP (13, 0.10 ^d). 3 L fed-batch reactors with glucose, 2 g/L of L-PHE and 3 g/L of L-TYR	[59]
GPT1017 (W3110 derivative)	$\Delta trpR \Delta tnaA \Delta ptsG \Delta aroP \Delta tnaB \Delta mtr$ / (chromosome) swapping of tryptophan attenuator and <i>trp</i> promoter by 5CP _{tacs} , (plasmid) $aroG^{fbr} trpE^{fbr} tktA$	L-TRP (16). 5 L fed-batch reactors with glucose and 1 g/L of yeast extract	[80]
TRTH0709/pMEL03 (MG1655 derivative)	$\Delta trpR \Delta tnaA \Delta pta \Delta mtr$ / (plasmid) $aroG^{fbr} trpE^{fbr} DCBA serA$, (plasmid) $tktA ppsA yddG$	L-TRP (49). 30 L fed-batch reactors with glucose and 1 g/L of yeast extract	[61]
Vio-4 (MG1655 derivative)	$\Delta trpR \Delta tnaA \Delta sdaA \Delta lac \Delta trpL \Delta gal \Delta xyl \Delta fuc$ / (chromosome) $P_{tac} aroF aroB aroL tktA serA^{fbr} vioD^f$, $trpE^{fbr}$, (plasmid) <i>viaABCE</i> ^g	Violacein (0.7). 0.7 L fed-batch reactors with arabinose, 12 g/L of tryptone and 24 g/L of yeast extract	[82]
BKD5 (BW25113 derivative)	$\Delta ptsG \Delta tyrR \Delta pykA \Delta pykF \Delta pheA$ / (plasmid) $P_{lacUV5} aroG^{fbr} tyrA^{fbr} aroE$, $P_{trc} ppsA tktA glk$, (plasmid) P_{lacUV5} T7 RNA polymerase, (plasmid) <i>hpaBC d-ldh</i> ^h	Salvianic acid A (7, 0.47 ^b). 0.5 L fed-batch flasks with glucose and 1 g/L of yeast extract	[94]
QH23 (ATCC 31884 derivative)	$\Delta pheLA \Delta tyrA$ / (plasmid) $P_{L-lacO1} tyrA^{fbr} ppsA tktA aroG^{fbr}$, (plasmid) $P_{L-lacO1} tal^i$ ^{op} <i>hpaBC</i>	Caffeic acid (0.8). Shake flask cultures with 2.5 g/L of glucose, 10 g/L of glycerol and phenylalanine	[105]
pAD-AG/ $\Delta tyrR$ (BL21 (DE3) derivative)	$\Delta tyrR$ / (plasmid) $aroG^{fbr} tyrA^{fbr}$, (plasmid) <i>tal</i> ^j	4-coumaric acid (1). Shake flask cultures with 15 g/L of glucose	[103]
VH33 $\Delta tyrR_DOPA$ (W3110 derivative)	$\Delta ptsHlcr$ / (chromosome) $P_{trc} galP$, (plasmid) $tktA P_{lacUV5} aroG^{fbr}$, (plasmid) $P_{trc} tyrC^c pheA$	L-DOPA (1.5, 0.05 ^d). 1 L batch reactors with LB and 50 g/L glucose	[84]
W3110 <i>trpD9923</i> /pS0 + pY + pAvnD	(plasmid) $P_{L-tetO1} ydiB aroD aroB aroG^{fbr} ppsA tktA$, (plasmid) $P_{lacUV5} tyrB tyrA^{fbr} aroC aroA aroL$, (plasmid) $P_{lacUV5} HCBT^k 4CL^l$ ¹ <i>tal</i> ^j	Avenanthramide D (27 ^e). Shake flask cultures with 10 g/L of glucose	[87]

^ag/L; ^bmol substrate/mol product; ^cgene from *Z. mobilis*; ^dg substrate/g product; ^eμM; ^fgene from *J. lividum*; ^ggenes from *C. violaceum*; ^hgene from *L. pentosus*; ⁱgene from *R. glutinis*; ^jgene from *S. espanaensis*; ^kgene from *D. caryophyllus*; ^lgene from *N. tabacum*; ^{op}codon-optimized variant.

mutase/prephenate dehydratase (PheA) also catalyzes the first step in the parallel biosynthetic pathways for L-TYR and L-PHE as well as the second step in L-PHE biosynthetic pathway (prephenate to phenylpyruvate) (Figure 1). The native enzyme is a dimer and each monomer contains a dehydratase active site, a mutase active site and an L-PHE binding site. PheA enzyme is inhibited by L-PHE (up to 90% of the prephenate dehydratase and 55% of the mutase activity) [1,2]. Feedback-resistant mutants of TyrA and PheA *E. coli* enzymes have been used for the efficient overproduction of L-TYR [11,17,57,67] and L-PHE [55,76] in combination with some of the previously described alterations in CCM and the SHK pathway (Table 1). An alternative approach to take advantage of the natural feedback-resistant diversity in the TyrA enzyme family was the expression of the TyrC^{fbr} enzyme (cyclohexadienyl dehydrogenase) from *Z. mobilis* and the CHA mutase domain of native PheA from *E. coli*, relieving rate-limiting steps and increasing the carbon flux towards L-TYR [57].

A strategy to minimize carbon loss to competing pathways was exemplified in the CHA node with the construction of L-PHE production strains expressing TyrA enzymes containing tags for increased proteolytic degradation, instead of completely removing the enzyme. The resultant strains have the advantage of not being auxotrophic to L-TYR while displaying a higher L-PHE/L-TYR production ratio than the strain containing the wild-type TyrA [77].

Additional modifications applied in L-TRP overproducers include the overexpression of exporter protein YddG [61,78,79], the inactivation of permeases AroP, Mtr and TnaB to avoid re-internalization [61,79,80], the deletion of gene *tnaA* coding for a tryptophanase to avoid product degradation [59,81,82] and expression of genes included in the tryptophan biosynthetic branch, including a feedback-resistant version of anthranilate synthase, TrpE^{fbr} [60,81].

Variations on the strategies described in this section have also been applied to the production of other valuable compounds derived from the AAA pathway such as phenyllactate, phenylacetate and phenylethanol [83], L-DOPA [84], mandelic acid [85], deoxyviolacein and violacein [82,86], avenanthramides [87], and resveratrol [88] (Table 2).

Increasing the genetic engineering repertoire: development and application of synthetic biology strategies and techniques

The field of synthetic biology has been continuously evolving and it is now acknowledged that this discipline is primarily concerned with the design and characterization of biological parts [89,90]. Indeed, modular and predictable parts find many applications in the modification of cellular metabolism, whether these alterations are direct

(modulation of the expression and function of enzymes comprised in metabolic pathways) or indirect (rewiring and repositioning of sensing components and cellular effectors). In this sense, the powerful recent advances in synthesis and assembly of macromolecules have changed the way to approach challenges in metabolic engineering. This has helped to generate a degree of biological diversity and reprogramming not previously reached with traditional biological controllers, promoting the merging of rational and combinatorial approaches to direct cellular design [91,92].

The aromatic biosynthetic pathway in *E. coli* was no exception to this paradigm shift, resulting in notable accomplishments over the last years. It is worth noting that even when the upregulation of a few genes can increase the carbon flux from CCM towards the aromatic biosynthetic pathway, the outcome is importantly influenced by a variety of factors, such as the combination of expression modules, genetic background and cultivation conditions. It is therefore ideal to design experiments to obtain a characterization of the contribution of each factor to the phenotype. Illustrative examples on this subject include the assessment of differences in the production of L-TYR by overexpressing various sets of genes in a stepwise approach (Table 1) [58,93].

The generation of synthetic parts in a faster, cheaper and more targeted way has also enabled metabolic engineers to reach unprecedented biochemical diversity, exemplified by the production of plant compounds using precursors present in the aromatic biosynthetic pathway in *E. coli*. In this way, combinations of simultaneous transcriptional modules and genetic platforms have resulted in strains with the ability to produce attractive compounds such as salvianic acid A [94], δ -tocotrienol and its intermediate 2-methyl-6-geranylgeranyl-benzoquinol [95,96] (Figure 2) and (*S*)-reticuline [97] (Figure 3).

Aside from the product titers reached so far, these approaches are appealing because the systematic evaluation of conditions permits a more precise identification of targets for future improvement. In this respect, more structurally complex compounds can be produced by the optimization of expression parameters, for example when approaching problems with the heterologous insertion of genes and pathways into *E. coli*. One successful case concerning a systematic analysis of heterologous expression is the production of (2*S*)-pinocembrin from glucose as the only carbon source [98] (Figure 2). In this report, the authors assembled gene expression modules, including genes from the SHK pathway as well as heterologous sources. With this arrangement it was possible to accumulate up to 40 mg/L of (2*S*)-pinocembrin, even when using four plasmids and enzymes with naturally low catalytic efficiencies. The same system was used to evaluate the capabilities for resveratrol production after

Table 2 Proposed applications of high-valued compounds derived from the aromatic pathway and synthesized by engineered *E. coli* strains

Compound	Summary of pharmaceutical and industrial applications	References
Shikimate ((3R,4S,5R)-3,4,5-trihydroxycyclohexene-1-carboxylic acid)	Antipyretic, antioxidant, anticoagulant, antithrombotic, anti-inflammatory, and analgesic agent. Has a key role in the synthesis of important pharmacological compounds such as anti-cancer and antibacterial agents, as well as hormones. Substrate in the chemical synthesis of the antiviral Tamiflu®.	[72,74]
Salvianic acid or danshensu (3,4-dihydroxyphenyllactic acid)	A naturally occurring plant polyphenolic acid, considered as a superior antioxidant. Its scavenging activities against free hydroxyl radicals and superoxide anion radicals are higher than vitamin C. Has a variety of other pharmacological effects, including improving cerebral blood flow, inhibiting platelet activation and arterial thrombosis, as well as anti-cancer and anti-inflammatory effects.	[94]
(2S)-pinocembrin (5,7-dihydroxyflavanone)	Flavonoid with demonstrated activity decreasing the neurological scores, alleviating brain edema, reducing the permeability of blood-brain barrier and alleviating cerebral ischemic injury in the middle cerebral artery occlusion in rats. Has been proposed as a novel therapeutic agent to reduce cerebral ischemia/reperfusion and blood-brain injury, useful for its antioxidant and anti-apoptotic effects.	[98]
Caffeic acid (3,4-dihydroxycinnamic acid)	Possesses various pharmacological activities including antioxidant, antitumoral, antiviral, antidepressive and antidiabetic functions.	[104]
Resveratrol (3,4',5-trihydroxystilbene)	Potential therapeutic effects in humans as antioxidant, anti-inflammatory, anticancer, and chemopreventive agent.	[99]
Violacein ((3E)-3-[5-(5-hydroxy-1H-indol-3-yl)-2-oxo-1H-pyrrol-3-ylidene]-1H-indol-2-one) and deoxyviolacein	Activity against herpes simplex virus and pathogenic bacteria such as <i>Staphylococcus aureus</i> and <i>Pseudomonas aeruginosa</i> . Violacein has shown successful activity against leukemia, lung cancer, human uveal melanoma and lymphoma cells, where it mediates apoptosis. It is also an interesting bio-dye showing attractive color tone and stability.	[82]
PDC (2-pyrone-4,6-dicarboxylic acid)	Proposed as a novel starting material for several useful synthetic polymers such as polyesters and polyamides.	[100]
(S)-reticuline ((1S)-1-[(3-hydroxy-4-methoxyphenyl)methyl]-6-methoxy-2-methyl-3,4-dihydro-1H-isoquinolin-7-ol)	Building block for benzyloisoquinoline alkaloids, including the analgesic compounds morphine and codeine, as well as the antibacterial agents berberine and palmatine. Useful in the development of novel antimalarial and anticancer drugs.	[97]
Hydroxytyrosol (3,4-dihydroxyphenylethanol)	Powerful antioxidant activity. Potential antitumoral, antiatherogenic, anti-inflammatory and antiplatelet aggregation agent.	[106]
Avenanthramides	Natural hydroxycinnamoyl anthranilates with antioxidant, anti-inflammatory, and antiproliferative effects, considered to contribute to the health benefits of oatmeal consumption. Potential antitumor activities.	[87]
δ-tocotrienol	Vitamin E component naturally produced by photosynthetic organisms. It has shown to induce apoptosis and inhibit proliferation of cancer cells. Possess to some extent neuroprotective, anticancer, and cholesterol lowering properties.	[95,96]

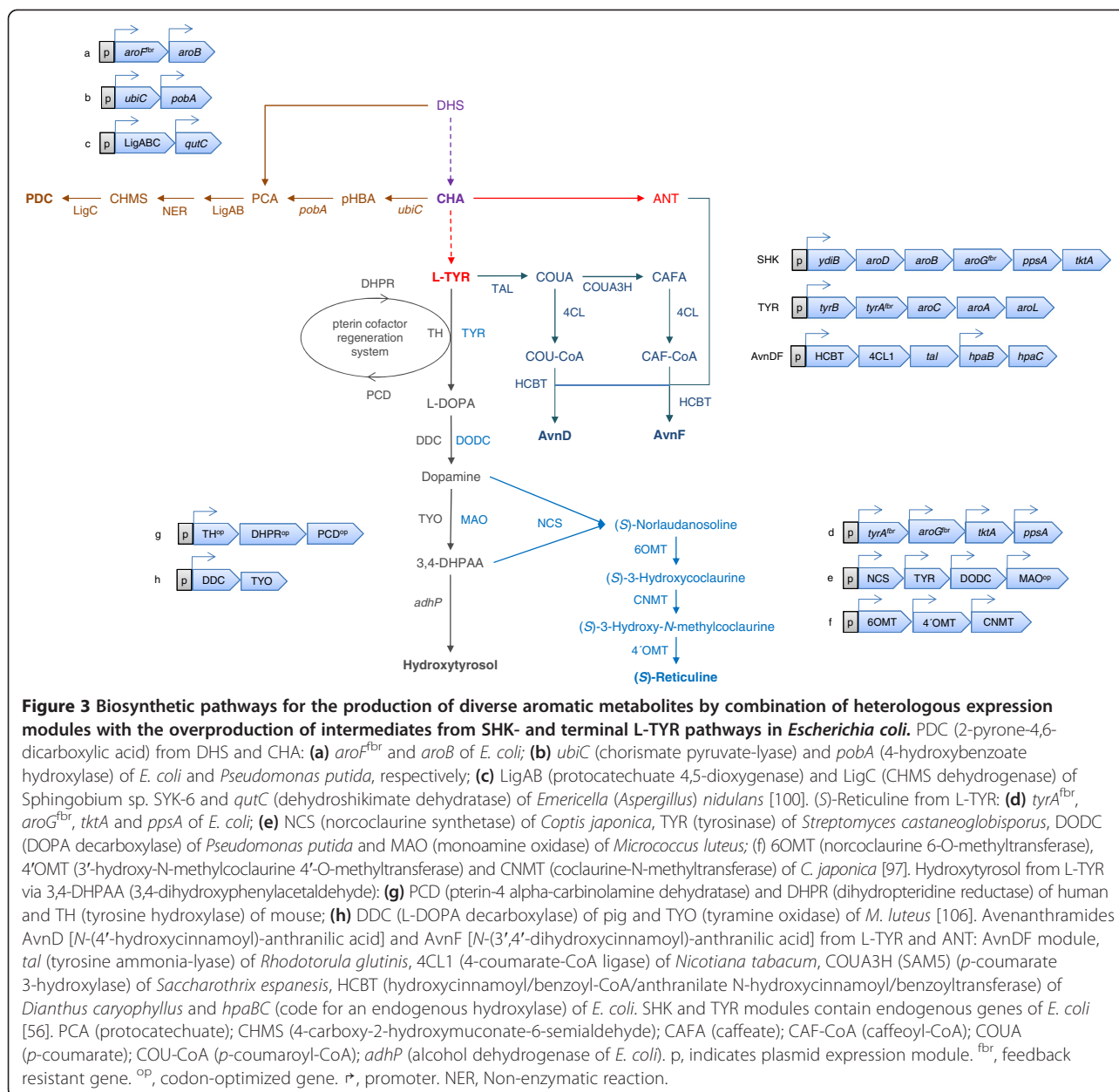
Recombinant pathways are presented in Figure 2 and Figure 3.

slight modifications were introduced to increase malonyl-CoA and L-TYR availability. This work revealed large variations in the concentration of produced metabolites with respect to small variations in the genetic constructs [99] (Figure 2). In a similar approach, the introduction of a foreign pathway succeeded in deviating carbon flow from 3-dehydroshikimate towards the synthesis of 2-pyrone-4,6-dicarboxylic acid (PDC) (Figure 3). Strains overexpressing six different genes from three different plasmids were able to produce the desired compound with a 17.3% yield from glucose [100].

It is also interesting to consider other strategies to generate and screen metabolic diversity, such as modifications of the global transcription machinery coupled to high-throughput screening for metabolite production

[101,102]. By merging these approaches with combinatorial techniques for gene overexpression, a 114% increase in L-TYR production from a previously engineered strain was reported [101].

Strains with the capability to overproduce L-TYR from simple carbon sources have been used as a backbone for production of more structurally complex compounds. For example, the construction of codon-optimized heterologous gene clusters with a wide span of strengths in promoter and ribosome binding sequences (RBS) has allowed the generation of *E. coli* strains capable of producing phenylpropanoic acids such as caffeic acid, coumaric acid and ferulic acid [103-105] (Figure 2), as well as hydroxycinnamoyl anthranilates [87] and other derivatives, such as hydroxytyrosol [106] (Figure 3). Another combinatorial



accumulate up to 21.9 g/L of L-TYR in high-density cultures [110].

Integration and application of data: systems-based approaches to the production of AAA

Even with the relative success obtained so far regarding the overproduction of aromatic compounds in *E. coli*, insights into the global metabolic state of engineered strains under production conditions are still scarce. Moreover, the effects of targeted strain modifications are typically underestimated, since they do not always result in significant differences in cell growth or production of specific metabolites. Combination of techniques such as

genomics, transcriptomics, proteomics, metabolomics and fluxomics can unravel the particular cellular state during a defined condition by providing snapshots of different levels of metabolism [111]. However, in order to turn this information into knowledge of new potential engineering targets, adequate comparisons must be established. Since it is not trivial to define the type and extent of data to be extracted and compared, systems biology approaches are needed to manage holistic information at different levels of cellular functions [112,113].

Although the systematic integration of -omics approaches have been applied to characterize and reverse engineer bacterial strains producing several amino acids [6,114-117],

there are still relatively few reports on the use of these techniques with AAA overproducers. For example, one study reports the effect of inactivating genes coding for PEP-consuming enzymes (PTS, *PykF* and *PykA*) over the flux distributions in the central carbon metabolism as an attempt to increase the availability of this AAA precursor [33]. The net result of either inactivation was a flux increase to biomass formation pathways, but several differences on important CCM nodes were also found between all conditions. Furthermore, PTS inactivation revealed a carbon recycling response between PEP and OAA combined with a reduced glycolytic flux. When these strains were transformed with plasmids encoding enzymes to promote the production of L-PHE, a 19-, 14-, and 25- fold increase on the yield of this amino acid was observed for the PTS, PTS⁻ *pykA*, and PTS⁻ *pykF* mutants, respectively [33].

Targeted proteomics and metabolite profiling analyses are also very valuable to provide feedback about expression systems used in the production of AAA. One report describes such approaches on a collection of L-TYR producing strains with different gene-expression arrays, allowing the authors to identify and improve sub-optimally expressed genes. After a second engineering round of the synthetic expression modules a strain was constructed which can produce L-TYR from glucose with 80% of the theoretical yield, estimated as 0.55 g/g in strains with a functional PTS [58]. A related work characterized the impact on SHK pathway enzyme levels resulting from the removal of TyrR regulator, along with the use of a feedback-resistant TyrA and deletion of the *pheA* gene on L-TYR producing strains. The results showed that small changes in protein levels caused by the genetic alterations can have a big impact on metabolite production, as a 250-fold span of L-TYR concentrations were detected [118]. A different work found many proteins differentially expressed as a response to the sole inactivation of the *pykF* gene, including DAHP synthase (*AroG*), SHK dehydrogenase (*AroE*), SHK kinase I (*AroK*), CHA synthase (*AroC*), prephenate dehydratase (*PheA*), anthranilate synthase (*TrpD*, *TrpE*) and L-TRP synthase (*TrpA*), as compared to the wild type strain [119].

In another example, transcriptional analysis and whole genome sequencing studies were performed on L-TYR producing strains obtained by combinatorial and targeted approaches, coupled to high-throughput screening, in an attempt to discover the changes that led to higher L-TYR production [101]. The transcriptional analysis revealed upregulation of genes related to acid stress resistance and global reductions in the expression of several pathways such as ribosomal protein and RNA formation, fatty acid elongation, de novo purine/pyrimidine biosynthesis and DNA replication, which imply a cellular shift from proliferation and growth to maintenance and stress

survival. Genomic analyses revealed differential single base-pair changes between the studied strains. When these mutations were reintroduced on a parental strain background higher L-TYR production was observed, showing their contributions to the overproduction phenotype. Finally, a reverse engineered strain was constructed, which gave a titer of 902 mg/L and an L-TYR yield on glucose of 0.18 g/g on a genetically-defined background [101]. Other works have also characterized the global transcriptional response to the presence of high levels of L-PHE or SHK in simple and complex media [120,121] or starvation conditions [122], revealing metabolic information that can be used for further improvement of the strains and cultivation conditions.

Along with data obtained by high-throughput systems, modeling of metabolism by mathematical approaches has become an important tool for analyzing cell responses and unravel the metabolic regulation between the cell information/control systems [111]. Moreover, genome-scale models of metabolism have been analyzed by constraint-based approaches [123]. Gene deletion effects over flux distributions have also been studied in order to find the combination that provides the best metabolic performance on a given condition. For example, the deletion impact of 1261 genes was modeled using a reconstruction of biochemical interactions, resulting in 195 genes exerting high impact on flux distributions in various metabolic subsystems [124]. A strategy developed to circumvent the need for kinetic parameters of enzymes present in a metabolic network is ensemble modeling, which uses phenotypic data obtained from overexpression and deletion of enzymes to screen out flux distributions from an initial ensemble of solutions derived from elementary reaction models [125]. This method has been used to model the AAA pathway for DAHP production with data obtained from the overexpression of CCM genes. A subset of flux distributions was found capable of describing the phenotypic characteristics of the strains and rendered information about the kinetic and stoichiometric limitations around PEP and E4P nodes [125]. As more genomic, transcriptomic and proteomic functional interactions continue to be unraveled, similar approaches will become powerful tools to model specific metabolic outcomes related to AAA biosynthesis.

Bioprocess engineering: optimization of AAA compound production

In order to create economically viable products, the processes developed and tested at laboratory scale have to be adapted to larger operational volumes. Although engineered strains should ideally perform equally in 1 L scale as in industrial scales (going from 500–10,000 L for fine chemicals to more than 100,000 L for commodity chemicals), a significant reduction in performance as a result of scale-up is often observed [126]. Therefore, it

is important to apply strategies to prevent physiological changes caused by heterogeneities of fermentation parameters during scale-up processes with *E. coli*. Stress-mediated cellular responses to chemical and physical factors can negatively impact as much as 60% over the productivity, the biomass and product yields when a strain is exposed to large-scale production conditions [127,128].

Fed-batch cultivations have been a popular method to produce aromatic compounds since they promote high cellular densities, offer tight control over the μ and substrate concentrations, and permit a better management of dissolved oxygen tension (DOT) to prevent the activation of fermentative pathways [16,129]. In one example, a fed-batch strategy improved violacein production from arabinose (through an expanded pathway from L-TRP) by adjusting the μ at 0.011 h^{-1} [82]. With this procedure, cellular concentrations with optical density values up to 70 were reached, producing 710 mg/L of violacein and avoiding acetate accumulation in the medium, a known inhibitor of growth and pigment production.

Another work studied the impact of different feeding strategies over the production of L-TRP in a recombinant strain [16]. An increase in the volumetric productivity of this compound was reached by a novel feeding strategy with a highly concentrated glucose solution (800 g/L) after the exhaustion of the initial glucose. By using a combined pseudo-exponential feeding at the exponential phase and a glucose-stat feeding after the exponential phase, an efficient control over the μ was achieved (below the acetic acid production threshold), reaching 38.8 g/L of L-TRP. This represented an increase of 19.9% due to reduced acetic acid accumulation [16].

Even if feeding strategies can cope to some extent with the problems derived from acetic acid production, a combination of these with genetic modifications has also been tested for the production of aromatic compounds. In a recent report, the effect of inactivating the gene coding for the enzyme phosphotransacetylase (Pta) over the production of L-TRP was assessed. By combining this modification with the use of a DO-stat for controlling inflow rate at a suitable DOT, the authors were able to increase the production of L-TRP and biomass while maintaining the growth rate and reducing the accumulation of acetic acid [129].

Substrate characteristics can also be optimization variables for the production of AAA pathway intermediates. One example is the evaluation of glycerol for L-PHE production [130,131]. The low cost of glycerol coupled with its higher degree of carbon reduction when compared to other sugars such as glucose, could result in high energy yield per carbon and hence be advantageous for AAA production processes. However, it is important to characterize the influence of fermentation parameters such as DOT, temperature and pH, as well as

the availability of substrates, over the growth and product formation rates. In one report, variations in oxygen supply (by changing aeration rates and impeller speeds) were tested over the L-PHE production capabilities of a recombinant strain growing on glycerol [130]. With this approach, a direct correlation between biomass and L-PHE production rates were found at impeller speeds up to 400 rpm, being this the maximum operational value before shear stress starts to diminish strain capabilities. After setting the impeller speed to 400 rpm, aeration optimization resulted in the highest product yield obtained, 0.58 g/g, which is 20% higher than the yields obtained before optimization of oxygen supply. Interestingly, the authors report this high yield with a strain in which the only recombinant measure taken is the heterologous expression of a phenylalanine dehydrogenase gene [130]. On the other hand, another group has recently reported the production of 13 g/L of L-PHE from glycerol and a yield of 0.15 g/g using a multi-phase fed-batch process with a strain containing several genetic modifications [131] (Table 1).

Product characteristics should also be taken into account when developing an efficient bioprocess. For example, L-TYR exhibits low solubility in typical fermentation conditions, triggering its precipitation when saturation is reached. This characteristic would normally be beneficial for a fermentation process, as a precipitated compound can be easily recovered and it is not expected to affect strain physiology and production capabilities. Interestingly, one report described that the L-TYR crystals can stabilize foam, causing operational problems during fermentation [17]. Consequently, this foaming process was studied on 10 and 200 L fed-batch fermentations for L-TYR production to assess the effect of pH, antifoam concentration, cooling rate, L-PHE concentration and seeding level on foam production. It was determined that high concentrations of L-PHE or antifoam, as well as low pH and low seeding, are the preferred conditions to avoid detrimental foaming production. With this approach it was possible to produce L-TYR from glucose with a yield of 0.3 g/g and titers as high as 55 g/L on a 200 L scale [17]. Moreover, this study revealed important data for the design of an economically feasible process for the production of L-TYR.

Process optimization could also be concerned with an enhancement of the strain ability to withstand high concentrations of aromatic compounds, not only for toxic final products but for harmful intermediates or byproducts, which often accumulate as a consequence of the suboptimal alleviation of control levels in the biosynthetic pathways. This is a commonly-encountered problem with many of the intermediates and final products in the AAA. For example, one group reported the optimization of L-TRP production by modifying the export

and import capabilities of a modified strain in order to minimize its intracellular concentration and avoid feedback control by product accumulation [79]. This group constructed a strain featuring the plasmid-based overexpression of the AAA exporter YddG, resulting in a production increase of 12.6% compared to the parental strain on a 30 L fermentation. Another example of the successful combination of genetic and fermentation procedures involves the construction of a strain for L-PHE overproduction with a PTS-independent glucose transport and expression of feedback-resistant versions of AroG and PheA. By overexpressing genes *ydiB*, *aroK* and *tyrB* with a temperature-dependent system, as well as *yddG* in a TyrA⁻ background, the authors were able to produce up to 47 g/L of L-PHE with a yield of 0.25 g/g from glucose in a 15 L fed-batch process [132].

Finally, bioprocess design is also an important factor in optimization of the production of aromatic pathway derivatives. The bioconversion of phenylpyruvate (PPN) to L-PHE was studied with an immobilized cell bioprocess [133]. This technique has several advantages such as the ability to reuse the immobilized cells, the capacity to utilize high cell densities and improved stability of the system. For example, a mixed-gel surface composed of k-carrageenan and gelatin, together with the optimization of its composition to enhance the mechanical strength and reduce the toxicity and solidification point was used as biomass carrier for the production of L-PHE [133]. Studies on the effect of pH, temperature, Mg²⁺ and trehalose presence resulted in the implementation of a process showing an improvement of 80% on the L-PHE conversion from PPN after 15 successive batch experiments.

Conclusions

The present review aims to provide a panorama of the current achievements and newly found goals related to the production of aromatic compounds in *E. coli*. The AAA pathway and the metabolic changes resulting from its deregulation have attracted the interest of metabolic engineers for many years and remain important research targets on several organisms. It is evident that the establishment of efficient bioprocesses on this topic requires the design and implementation of multidisciplinary strategies, taking advantage of the fast-paced developments coming from nearly all biotechnological fields but particularly from those related with information technologies, such as systems and synthetic biology. The works compiled here are a good example of the benefits obtained when new ideas and viewpoints are introduced to an established field in order to cope with long-known problems. From the comparisons presented, it is noticeable that the use of rational and combinatorial approaches powered by the ability to develop complex genetic circuits and high-throughput screenings of new producers has set new trends

when dealing with the production of aromatic compounds in *E. coli*. The benefits of the integral application of these technologies can already be observed, not only from the improved production processes for AAA and pathway intermediates with large and established markets, but also with the generation of novel derivative compounds with important pharmaceutical applications.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors participated in the preparation of this contribution. AR had a major role in writing and editing the manuscript. All authors have read and approved the final version.

Acknowledgements

This work was supported by CONACYT 105782, 177568 and DGAPA-PAPIIT UNAM IN206812 grants.

Received: 30 May 2014 Accepted: 17 August 2014

Published online: 09 September 2014

References

1. Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martinez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muniz-Rascado L, Ong Q, Paley S, Schroder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus RP, Paulsen I, Karp PD: **EcoCyc: fusing model organism databases with systems biology.** *Nucleic Acids Res* 2013, **41**:D605–D612.
2. Sprenger G: **Aromatic Amino Acids.** In *Amin Acid Biosynth - Pathways, Regul Metab Eng.* Edited by Wendisch VF. Berlin, Heidelberg: Springer; 2007:418 [Microbiology Monographs, vol. 5].
3. Herrmann KM, Weaver LM: **The shikimate pathway.** *Annu Rev Plant Biol* 1999, **50**:473–503.
4. Richards TA, Dacks JB, Campbell SA, Blanchard JL, Foster PG, McLeod R, Roberts CW: **Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements.** *Eukaryot Cell* 2006, **5**:1517–1531.
5. Lütke-Eversloh T, Santos CNS, Stephanopoulos G: **Perspectives of biotechnological production of L-tyrosine and its applications.** *Appl Microbiol Biotechnol* 2007, **77**:751–762.
6. Becker J, Wittmann C: **Systems and synthetic metabolic engineering for amino acid production - the heartbeat of industrial strain development.** *Curr Opin Biotechnol* 2012, **23**:718–726.
7. Ikeda M, Takeno S: **Amino acid production by *Corynebacterium glutamicum*.** In *Corynebacterium glutamicum. Volume 23.* Edited by Yukawa H, Inui M. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013:107–147.
8. Ajinomoto Co. Inc: **FY2013 Market and Other Information.** 2014.
9. Li Z, Ji X, Kan S, Qiao H, Jian M, Lu D, Wang J, Huang H, Jia H, Ouyang P, Ying H: **Past, Present and Future Industrial Biotechnology in China.** In *Biotechnol China II Chem Energy Environ.* Edited by Tsao GT, Ouyang P, Chen J. Berlin, Heidelberg: Springer; 2010:1–42.
10. Bongaerts J, Krämer M, Müller U, Raeven L, Wubbolts M: **Metabolic engineering for microbial production of aromatic amino acids and derived compounds.** *Metab Eng* 2001, **3**:289–300.
11. Sprenger GA: **From scratch to value: Engineering *Escherichia coli* wild type cells to the production of L-phenylalanine and other fine chemicals derived from chorismate.** *Appl Microbiol Biotechnol* 2007, **75**:739–749.
12. Pittard J, Yang J: **Biosynthesis of the Aromatic Amino Acids.** *Eco Sal Plus* 2008, **1**.
13. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, Palsson BØ: **A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism-2011.** *Mol Syst Biol* 2011, **7**.
14. Klein-Marcuschamer D, Santos CNS, Yu H, Stephanopoulos G: **Mutagenesis of the bacterial RNA polymerase alpha subunit for improvement of complex phenotypes.** *Appl Environ Microbiol* 2009, **75**:2705–2711.
15. Jeong J, Cho N, Jung D, Bang D: **Genome-scale genetic engineering in *Escherichia coli*.** *Biotechnol Adv* 2013, **31**:804–810.

16. Cheng L-K, Wang J, Xu Q-Y, Xie X-X, Zhang Y-J, Zhao C-G, Chen N: **Effect of feeding strategy on L-tryptophan production by recombinant *Escherichia coli***. *Ann Microbiol* 2012, **62**:1625–1634.
17. Patnaik R, Zolanz RR, Green DA, Kraynie DF: **L-Tyrosine production by recombinant *Escherichia coli*: Fermentation optimization and recovery**. *Biotechnol Bioeng* 2008, **99**:741–752.
18. Ikeda M: **Amino acid production processes**. In *Microb Prod L-amino acids*. Berlin, Heidelberg: Springer; 2003:1–35.
19. Gosset G: **Production of aromatic compounds in bacteria**. *Curr Opin Biotechnol* 2009, **20**:651–658.
20. Westerhoff HV, Palsson BO: **The evolution of molecular biology into systems biology**. *Nat Biotechnol* 2004, **22**:1249–1252.
21. Blazeck J, Alper H: **Systems metabolic engineering: Genome-scale models and beyond**. *Biotechnol J* 2010, **5**:647–659.
22. Postma PW, Lengeler JW, Jacobson GR: **Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria**. *Microbiol Rev* 1993, **57**:543–594.
23. Gosset G: **Improvement of *Escherichia coli* production strains by modification of the phosphoenolpyruvate: sugar phosphotransferase system**. *Microb Cell Fact* 2005, **4**:14.
24. Escalante A, Salinas-Cervantes A, Gosset G, Bolivar F: **Current knowledge of the *Escherichia coli* phosphoenolpyruvate-carbohydrate phosphotransferase system: peculiarities of regulation and impact on growth and product formation**. *Appl Microbiol Biotechnol* 2012, **94**:1483–1494.
25. Sauer U, Eikmanns BJ: **The PEP – pyruvate – oxaloacetate node as the switch point for carbon flux distribution in bacteria**. *FEMS Microbiol Rev* 2005, **29**:765–794.
26. Sprenger G, Siewe R, Martin K, Sonke T: *Microbial Preparation of Substances from Aromatic Metabolism/I*. WO patent 98/18936. 1998.
27. Frost JW, Knop DR: *Biocatalytic Synthesis of Shikimic Acid*. US patent 6,472,169 B1. 2002.
28. Chandran SS, Yi J, Draths KM, Von Daeniken R, Weber W, Frost JW: **Phosphoenolpyruvate availability and the biosynthesis of shikimic acid**. *Biotechnol Prog* 2003, **19**:808–814.
29. Yi J, Draths KM, Li K, Frost JW: **Altered glucose transport and shikimate pathway product yields in *E. coli***. *Biotechnol Prog* 2003, **19**:1450–1459.
30. Balderas-Hernandez VE, Sabido-Ramos A, Silva P, Cabrera-Valladares N, Hernandez-Chavez G, Baez-Viveros JL, Martinez A, Bolivar F, Gosset G: **Metabolic engineering for improving anthranilate synthesis from glucose in *Escherichia coli***. *Microb Cell Fact* 2009, **8**:19.
31. Flores N, Xiao J, Bolivar F, Valle F: **Pathway engineering for the production of aromatic compounds in *Escherichia coli***. *Nat Biotechnol* 1996, **14**:620–623.
32. Valle F, Mejia N, Berry A: *Application of Glucose Transport Mutants for Production of Aromatic Pathway Compounds*. WO patent 96/34961. 1996.
33. Meza E, Becker J, Bolivar F, Gosset G, Wittmann C: **Consequences of phosphoenolpyruvate: sugar phosphotransferase system and pyruvate kinase isozymes inactivation in central carbon metabolism flux distribution in *Escherichia coli***. *Microb Cell Fact* 2012, **11**:127.
34. Sabido A, Sigala JC, Hernández-Chávez G, Flores N, Gosset G, Bolivar F: **Physiological and transcriptional characterization of *Escherichia coli* strains lacking interconversion of phosphoenolpyruvate and pyruvate when glucose and acetate are coutilized**. *Biotechnol Bioeng* 2014, **111**:1150–1160.
35. Patnaik R, Liao JC: **Engineering of *Escherichia coli* central metabolism for aromatic metabolite production with near theoretical yield**. *Appl Environ Microbiol* 1994, **60**:3903–3908.
36. Liao JC: *Microorganisms and Methods for Overproduction of DAHP by Cloned pps Gene*. WO patent 96/08567. 1996.
37. Yi J, Li K, Draths KM, Frost JW: **Modulation of phosphoenolpyruvate synthase expression increases shikimate pathway product yields in *E. coli***. *Biotechnol Prog* 2002, **18**:1141–1148.
38. Liao JC, Hou SY, Chao YP: **Pathway analysis, engineering, and physiological considerations for redirecting central metabolism**. *Biotechnol Bioeng* 1996, **52**:129–140.
39. Gulevich AY, Biryukova IV, Zimenkov DV, Skorokhodova AY, Kivero AD, Belareva AV, Mashko SV: *Method for Producing An L-amino Acid Using A Bacterium Having Enhanced Expression of the pckA Gene*. US Patent 2006/0035348 A1. 2006.
40. Tatarko M, Romeo T: **Disruption of a global regulatory gene to enhance central carbon flux into phenylalanine biosynthesis in *Escherichia coli***. *Curr Microbiol* 2001, **43**:26–32.
41. Yakandawala N, Romeo T, Friesen AD, Madhyastha S: **Metabolic engineering of *Escherichia coli* to enhance phenylalanine production**. *Appl Microbiol Biotechnol* 2008, **78**:283–291.
42. Zhao G, Winkler ME: **An *Escherichia coli* K-12 tktA tktB mutant deficient in transketolase activity requires pyridoxine (vitamin B6) as well as the aromatic amino acids and vitamins for growth**. *J Bacteriol* 1994, **176**:6134–6138.
43. Nakahigashi K, Toya Y, Ishii N, Soga T, Hasegawa M, Watanabe H, Takai Y, Honma M, Mori H, Tomita M: **Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism**. *Mol Syst Biol* 2009, **5**:306.
44. Draths KM, Pompliano DL, Conley DL, Frost JW, Berry A, Disbrow GL, Staversky RJ, Lievens JC: **Biocatalytic synthesis of aromatics from D-Glucose: the role of transketolase**. *J Am Chem Soc* 1992, **114**:3956–3962.
45. Patnaik R, Spitzer RG, Liao JC: **Pathway engineering for production of aromatics in *Escherichia coli*: confirmation of stoichiometric analysis by independent modulation of AroG, TktA, and Pps activities**. *Biotechnol Bioeng* 1995, **46**:361–370.
46. Báez JL, Bolivar F, Gosset G: **Determination of 3-deoxy-D-arabino-heptulosonate 7-phosphate productivity and yield from glucose in *Escherichia coli* devoid of the glucose phosphotransferase transport system**. *Biotechnol Bioeng* 2001, **73**:530–535.
47. Lu J, Liao JC: **Metabolic engineering and control analysis for production of aromatics: role of transaldolase**. *Biotechnol Bioeng* 1997, **53**:132–138.
48. Mascarenhas D, Ashworth DJ, Chen CS: **Deletion of pgi alters tryptophan biosynthesis in a genetically engineered strain of *Escherichia coli***. *Appl Environ Microbiol* 1991, **57**:2995–2999.
49. Ahn J, Chung BKS, Lee D, Park M, Karimi IA, Jung J, Lee H: **NADPH-dependent pgi-gene knockout *Escherichia coli* metabolism producing shikimate on different carbon sources**. *FEMS Microbiol Lett* 2011, **324**:10–16.
50. Rodriguez A, Martínez JA, Báez-Viveros JL, Flores N, Hernández-Chávez G, Ramírez OT, Gosset G, Bolivar F: **Constitutive expression of selected genes from the pentose phosphate and aromatic pathways increases the shikimic acid yield in high-glucose batch cultures of an *Escherichia coli* strain lacking PTS and pykF**. *Microb Cell Fact* 2013, **12**:86.
51. Li K, Frost JW: **Microbial synthesis of 3-dehydroshikimic acid: a comparative analysis of D-xylose, L-arabinose, and D-glucose carbon sources**. *Biotechnol Prog* 1999, **15**:876–883.
52. Martínez K, De Anda R, Hernández G, Escalante A, Gosset G, Ramírez OT, Bolivar F: **Coutilization of glucose and glycerol enhances the production of aromatic compounds in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system**. *Microb Cell Fact* 2008, **7**:1.
53. Ahn JO, Lee HW, Saha R, Park MS, Jung JK, Lee DY: **Exploring the effects of carbon sources on the metabolic capacity for shikimic acid production in *Escherichia coli* using in silico metabolic predictions**. *J Microbiol Biotechnol* 2008, **18**:1773–1784.
54. Chen K, Dou J, Tang S, Yang Y, Wang H, Fang H, Zhou C: **Deletion of the aroK gene is essential for high shikimic acid accumulation through the shikimate pathway in *E. coli***. *Bioresour Technol* 2012, **119**:141–147.
55. Báez-Viveros JL, Osuna J, Hernández-Chávez G, Soberón X, Bolivar F, Gosset G: **Metabolic engineering and protein directed evolution increase the yield of L-phenylalanine synthesized from glucose in *Escherichia coli***. *Biotechnol Bioeng* 2004, **87**:516–524.
56. Liu S-P, Xiao M-R, Zhang L, Xu J, Ding Z-Y, Gu Z-H, Shi G-Y: **Production of L-phenylalanine from glucose by metabolic engineering of wild type *Escherichia coli* W3110**. *Process Biochem* 2013, **48**:413–419.
57. Chavez-Bejar MI, Lara AR, Lopez H, Hernandez-Chavez G, Martinez A, Ramirez OT, Bolivar F, Gosset G: **Metabolic engineering of *Escherichia coli* for L-tyrosine production by expression of genes coding for the chorismate mutase domain of the native chorismate mutase-prephenate dehydratase and a cyclohexadienyl dehydrogenase from *Zymomonas mobilis***. *Appl Environ Microbiol* 2008, **74**:3284–3290.
58. Juminaga D, Baidoo EEK, Redding-Johanson AM, Bath TS, Burd H, Mukhopadhyay A, Petzold CJ, Keasling JD: **Modular engineering of L-tyrosine production in *Escherichia coli***. *Appl Environ Microbiol* 2012, **78**:89–98.
59. Zhao Z-J, Zou C, Zhu Y-X, Dai J, Chen S, Wu D, Wu J, Chen J: **Development of L-tryptophan production strains by defined genetic modification in *Escherichia coli***. *J Ind Microbiol Biotechnol* 2011, **38**:1921–1929.
60. Shen T, Liu Q, Xie X, Xu Q, Chen N: **Improved production of tryptophan in genetically engineered *Escherichia coli* with TktA and PpsA overexpression**. *J Biomed Biotechnol* 2012, **2012**:1–8.

61. Wang J, Cheng L-K, Wang J, Liu Q, Shen T, Chen N: Genetic engineering of *Escherichia coli* to enhance production of L-tryptophan. *Appl Microbiol Biotechnol* 2013, **97**:7587–7596.
62. Doroshenko VG, Tsyrenzhapova IS, Krylov AA, Kiseleva EM, Ermishev VY, Kazakova SM, Biryukova IV, Mashko SV: Pho regulon promoter-mediated transcription of the key pathway gene *aroG*F_{br} improves the performance of an L-phenylalanine-producing *Escherichia coli* strain. *Appl Microbiol Biotechnol* 2010, **88**:1287–1295.
63. Dell KA, Frost JW: Identification and removal of impediments to biocatalytic synthesis of aromatics from D-Glucose: rate-limiting enzymes in the common pathway of aromatic amino acid biosynthesis. *J Am Chem Soc* 1993, **115**:11581–11589.
64. Krämer M, Bongaerts J, Bovenberg R, Kremer S, Müller U, Orf S, Wubbolts M, Raeven L: Metabolic engineering for microbial production of shikimic acid. *Metab Eng* 2003, **5**:277–283.
65. Oldiges M, Kunze M, Degenring D, Sprenger GA, Takors R: Stimulation, monitoring, and analysis of pathway dynamics by metabolic profiling in the aromatic amino acid pathway. *Biotechnol Prog* 2004, **20**:1623–1633.
66. Escalante A, Calderón R, Valdivia A, De Anda R, Hernández G, Ramírez OT, Gosset G, Bolívar F: Metabolic engineering for the production of shikimic acid in an evolved *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. *Microb Cell Fact* 2010, **9**:21.
67. Lutke-Eversloh T, Stephanopoulos G: Feedback inhibition of chorismate mutase/prephenate dehydrogenase (TyrA) of *Escherichia coli*: Generation and characterization of tyrosine-insensitive mutants. *Appl Environ Microbiol* 2005, **71**:7224–7228.
68. Cui Y-Y, Ling C, Zhang Y-Y, Huang J, Liu J-Z: Production of shikimic acid from *Escherichia coli* through chemically inducible chromosomal evolution and cofactor metabolic engineering. *Microb Cell Fact* 2014, **13**:21.
69. Johansson L, Lindskog A, Silfversparre G, Cimander C, Nielsen KF, Lidén G: Shikimic acid production by a modified strain of *E. coli* (W3110.shik1) under phosphate-limited and carbon-limited conditions. *Biotechnol Bioeng* 2005, **92**:541–552.
70. Chen X, Li M, Zhou L, Shen W, Algasan G, Fan Y, Wang Z: Metabolic engineering of *Escherichia coli* for improving shikimate synthesis from glucose. *Bioresour Technol* 2014, **166**:64–71.
71. Ghosh S, Chisti Y, Banerjee UC: Production of shikimic acid. *Biotechnol Adv* 2012, **30**:1425–1431.
72. Rawat G, Tripathi P, Saxena RK: Expanding horizons of shikimic acid: Recent progresses in production and its endless frontiers in application and market trends. *Appl Microbiol Biotechnol* 2013, **97**:4277–4287.
73. Tripathi P, Rawat G, Yadav S, Saxena RK: Fermentative production of shikimic acid: a paradigm shift of production concept from plant route to microbial route. *Bioprocess Biosyst Eng* 2013, **36**:1665–1673.
74. Estevez A, Estevez R: A short overview on the medicinal chemistry of (–)-shikimic acid. *Mini Rev Med Chem* 2012, **12**:1443–1454.
75. Salgado H, Peralta-Gil M, Gama-Castro S, Santos-Zavaleta A, Muñoz-Rascado L, García-Sotelo JS, Weiss V, Solano-Lira H, Martínez-Flores I, Medina-Rivera A, Salgado-Osorio G, Alquicira-Hernández S, Alquicira-Hernández K, López-Fuentes A, Porrón-Sotelo L, Huerta AM, Bonavides-Martínez C, Balderas-Martínez YI, Pannier L, Olvera M, Labastida A, Jiménez-Jacinto V, Vega-Alvarado L, Del Moral-Chávez V, Hernández-Alvarez A, Morett E, Collado-Vides J: RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res* 2013, **41**(Database issue):D203–D213.
76. Báez-Viveros J, Flores N, Juárez K, Castillo-España P, Bolívar F, Gosset G: Metabolic transcription analysis of engineered *Escherichia coli* strains that overproduce L-phenylalanine. *Microb Cell Fact* 2007, **6**:30.
77. Doroshenko VG, Shakulov RS, Kazakova SM, Kivero AD, Yampolskaya TA, Mashko SV: Construction of an L-phenylalanine-producing tyrosine-prototrophic *Escherichia coli* strain using *tyrA* *ssrA*-like tagged alleles. *Biotechnol Lett* 2010, **32**:1117–1121.
78. Doroshenko V, Airich L, Vitushkina M, Kolokolova A, Livshits V, Mashko S: YddG from *Escherichia coli* promotes export of aromatic amino acids. *FEMS Microbiol Lett* 2007, **275**:312–318.
79. Liu Q, Cheng Y, Xie X, Xu Q, Chen N: Modification of tryptophan transport system and its impact on production of L-tryptophan in *Escherichia coli*. *Bioresour Technol* 2012, **114**:549–554.
80. Gu P, Yang F, Li F, Liang Q, Qi Q: Knocking out analysis of tryptophan permeases in *Escherichia coli* for improving L-tryptophan production. *Appl Microbiol Biotechnol* 2013, **97**:6677–6683.
81. Gu P, Yang F, Kang J, Wang Q, Qi Q: One-step of tryptophan attenuator inactivation and promoter swapping to improve the production of L-tryptophan in *Escherichia coli*. *Microb Cell Fact* 2012, **11**:30.
82. Rodrigues AL, Trachtmann N, Becker J, Lohanatha AF, Blotenberg J, Bolten CJ, Korneli C, De Souza Lima AO, Porto LM, Sprenger GA, Wittmann C: Systems metabolic engineering of *Escherichia coli* for production of the antitumor drugs violacein and deoxyviolacein. *Metab Eng* 2013, **20**:29–41.
83. Koma D, Yamanaka H, Moriyoshi K, Ohmoto T, Sakai K: Production of aromatic compounds by metabolically engineered *Escherichia coli* with an expanded shikimate pathway. *Appl Environ Microbiol* 2012, **78**:6203–6216.
84. Muñoz AJ, Hernández-Chávez G, Anda R, Martínez A, Bolívar F, Gosset G: Metabolic engineering of *Escherichia coli* for improving L-3,4-dihydroxyphenylalanine (L-DOPA) synthesis from glucose. *J Ind Microbiol Biotechnol* 2011, **38**:1845–1852.
85. Sun Z, Ning Y, Liu L, Liu Y, Sun B, Jiang W, Yang C, Yang S: Metabolic engineering of the L-phenylalanine pathway in *Escherichia coli* for the production of S- or R-mandelic acid. *Microb Cell Fact* 2011, **10**:71.
86. Rodrigues AL, Becker J, De Souza Lima AO, Porto LM, Wittmann C: Systems metabolic engineering of *Escherichia coli* for gram scale production of the antitumor drug deoxyviolacein from glycerol. *Biotechnol Bioeng* 2014, **20**:1–31.
87. Eudes A, Juminaga D, Baidoo EE, Collins FW, Keasling JD, Loqué D: Production of hydroxycinnamoyl anthranilates from glucose in *Escherichia coli*. *Microb Cell Fact* 2013, **12**:62.
88. Lim CG, Fowler ZL, Hueller T, Schaffer S, Koffas MAG: High-yield resveratrol production in engineered *Escherichia coli*. *Appl Environ Microbiol* 2011, **77**:3451–3460.
89. Lynch SA, Gill RT: Synthetic biology: New strategies for directing design. *Metab Eng* 2012, **14**:205–211.
90. Luo Y, Lee J-K, Zhao H: Challenges and opportunities in synthetic biology for chemical engineers. *Chem Eng Sci* 2013, **103**:115–119.
91. Boyle PM, Silver PA: Parts plus pipes: Synthetic biology approaches to metabolic engineering. *Metab Eng* 2012, **14**:223–232.
92. Yadav VG, De Mey M, Giaw Lim C, Kumaran Ajithkumar P, Stephanopoulos G: The future of metabolic engineering and synthetic biology: Towards a systematic practice. *Metab Eng* 2012, **14**:233–241.
93. Lutke-Eversloh T, Stephanopoulos G: Combinatorial pathway analysis for improved L-tyrosine production in *Escherichia coli*: identification of enzymatic bottlenecks by systematic gene overexpression. *Metab Eng* 2008, **10**:69–77.
94. Yao Y-F, Wang C-S, Qiao J, Zhao G-R: Metabolic engineering of *Escherichia coli* for production of salvanic acid A via an artificial biosynthetic pathway. *Metab Eng* 2013, **19**:79–87.
95. Albermann C, Ghanegaonkar S, Lemuth K, Vallon T, Reuss M, Armbruster W, Sprenger G: Biosynthesis of the vitamin E compound delta-tocotrienol in recombinant *Escherichia coli* cells. *Chembiochem* 2008, **9**:2524–2533.
96. Ghanegaonkar S, Conrad J, Beifuss U, Sprenger G, Albermann C: Towards the in vivo production of tocotrienol compounds: engineering of a plasmid-free *Escherichia coli* strain for the heterologous synthesis of 2-methyl-6-geranylgeranyl-benzoquinol. *J Biotechnol* 2012, **164**:238–247.
97. Nakagawa A, Minami H, Kim J-S, Koyanagi T, Katayama T, Sato F, Kumagai H: A bacterial platform for fermentative production of plant alkaloids. *Nat Commun* 2011, **2**:326.
98. Wu J, Du G, Zhou J, Chen J: Metabolic engineering of *Escherichia coli* for (2S)-pinocembrin production from glucose by a modular metabolic strategy. *Metab Eng* 2013, **16**:48–55.
99. Wu J, Liu P, Fan Y, Bao H, Du G, Zhou J, Chen J: Multivariate modular metabolic engineering of *Escherichia coli* to produce resveratrol from L-tyrosine. *J Biotechnol* 2013, **167**:404–411.
100. Nakajima M, Nishino Y, Tamura M, Mase K, Masai E, Otsuka Y, Nakamura M, Sato K, Fukuda M, Shigehara K, Ohara S, Katayama Y, Kajita S: Microbial conversion of glucose to a novel chemical building block, 2-pyrone-4,6-dicarboxylic acid. *Metab Eng* 2009, **11**:213–220.
101. Santos CNS, Xiao W, Stephanopoulos G: Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in *Escherichia coli*. *Proc Natl Acad Sci* 2012, **109**:13538–13543.
102. Alper H, Stephanopoulos G: Global transcription machinery engineering: A new approach for improving cellular phenotype. *Metab Eng* 2007, **9**:258–267.

103. Kang S-Y, Choi O, Lee JK, Hwang BY, Uhm T-B, Hong Y-S: **Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing *Escherichia coli* strain.** *Microb Cell Fact* 2012, **11**:153.
104. Zhang H, Stephanopoulos G: **Engineering *E. coli* for caffeic acid biosynthesis from renewable sugars.** *Appl Microbiol Biotechnol* 2013, **97**:3333–3341.
105. Huang Q, Lin Y, Yan Y: **Caffeic acid production enhancement by engineering a phenylalanine over-producing *Escherichia coli* strain.** *Biotechnol Bioeng* 2013, **110**:3188–3196.
106. Satoh Y, Tajima K, Munekata M, Keasling JD, Lee TS: **Engineering of L-tyrosine oxidation in *Escherichia coli* and microbial production of hydroxytyrosol.** *Metab Eng* 2012, **14**:603–610.
107. Wang HH, Isaacs FJ, Carr P a, Sun ZZ, Xu G, Forest CR, Church GM: **Programming cells by multiplex genome engineering and accelerated evolution.** *Nature* 2009, **460**:894–898.
108. Wang HH, Kim H, Cong L, Jeong J, Bang D, Church GM: **Genome-scale promoter engineering by coselection MAGE.** *Nat Methods* 2012, **9**:591–593.
109. Yang J, Seo SW, Jang S, Shin S-I, Lim CH, Roh T-Y, Jung GY: **Synthetic RNA devices to expedite the evolution of metabolite-producing microbes.** *Nat Commun* 2013, **4**:1413.
110. Na D, Yoo SM, Chung H, Park H, Park JH, Lee SY: **Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs.** *Nat Biotechnol* 2013, **31**:170–174.
111. Shimizu K: **Toward systematic metabolic engineering based on the analysis of metabolic regulation by the integration of different levels of information.** *Biochem Eng J* 2009, **46**:235–251.
112. Bro C, Nielsen J: **Impact of “ome” analyses on inverse metabolic engineering.** *Metab Eng* 2004, **6**:204–211.
113. Park JH, Lee SY, Kim TY, Kim HU: **Application of systems biology for bioprocess development.** *Trends Biotechnol* 2008, **26**:404–412.
114. Park JH, Lee SY: **Towards systems metabolic engineering of microorganisms for amino acid production.** *Curr Opin Biotechnol* 2008, **19**:454–460.
115. Bartek T, Blombach B, Lang S, Eikmanns BJ, Wiechert W, Oldiges M, Nöh K, Noack S: **Comparative ¹³C metabolic flux analysis of pyruvate dehydrogenase complex-deficient, L-valine-producing *Corynebacterium glutamicum*.** *Appl Environ Microbiol* 2011, **77**:6644–6652.
116. Becker J, Zelder O, Häfner S, Schröder H, Wittmann C: **From zero to hero — Design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production.** *Metab Eng* 2011, **13**:159–168.
117. Van Ooyen J, Noack S, Bott M, Reth A, Eggeling L: **Improved L-lysine production with *Corynebacterium glutamicum* and systemic insight into citrate synthase flux and activity.** *Biotechnol Bioeng* 2012, **109**:2070–2081.
118. Singh P, Batth TS, Juminaga D, Dahl RH, Keasling JD, Adams PD, Petzold CJ: **Application of targeted proteomics to metabolically engineered *Escherichia coli*.** *Proteomics* 2012, **12**:1289–1299.
119. Kedar P, Colah R, Shimizu K: **Proteomic investigation on the *pyk-F* gene knockout *Escherichia coli* for aromatic amino acid production.** *Enzyme Microb Technol* 2007, **41**:455–465.
120. Polen T, Krämer M, Bongaerts J, Wubbolts M, Wendisch VF: **The global gene expression response of *Escherichia coli* to L-phenylalanine.** *J Biotechnol* 2005, **115**:221–237.
121. Cortés-Talpa L, Gutiérrez-Ríos RM, Martínez LM, De Anda R, Gosset G, Bolívar F, Escalante A: **Global transcriptomic analysis of an engineered *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system during shikimic acid production in rich culture medium.** *Microb Cell Fact* 2014, **13**:28.
122. Johansson L, Lidén G: **Transcriptome analysis of a shikimic acid producing strain of *Escherichia coli* W3110 grown under carbon- and phosphate-limited conditions.** *J Biotechnol* 2006, **126**:528–545.
123. Joyce AR, Reed JL, White A, Edwards R, Osterman A, Baba T, Mori H, Lesely SA, Palsson BO, Agarwalla S: **Experimental and computational assessment of conditionally essential genes in *Escherichia coli*.** *J Bacteriol* 2006, **188**:8259–8271.
124. Xu Z, Sun X, Yu S: **Genome-scale analysis to the impact of gene deletion on the metabolism of *E. coli*: constraint-based simulation approach.** *BMC Bioinformatics* 2009, **10**(Suppl 1):S62.
125. Rizk ML, Liao JC: **Ensemble modeling for aromatic production in *Escherichia coli*.** *PLoS One* 2009, **4**:e6903.
126. Takors R: **Scale-up of microbial processes: Impacts, tools and open questions.** *J Biotechnol* 2012, **160**:3–9.
127. Lara AR, Galindo E, Ramirez OT, Palomares LA: **Living with heterogeneities in bioreactors. Understanding the effects of environmental gradients on cells.** *Mol Biotechnol* 2006, **34**:355–381.
128. Noorman H: **An industrial perspective on bioreactor scale-down: What we can learn from combined large-scale bioprocess and model fluid studies.** *Biotechnol J* 2011, **6**:934–943.
129. Wang J, Huang J, Shi J, Xu Q, Xie X, Chen N: **Fermentation characterization of an L-tryptophan producing *Escherichia coli* strain with inactivated phosphotransacetylase.** *Ann Microbiol* 2013, **63**:1219–1224.
130. Khamduang M, Packdibamrung K, Chutmanop J, Chisti Y, Srinophakun P: **Production of L-phenylalanine from glycerol by a recombinant *Escherichia coli*.** *J Ind Microbiol Biotechnol* 2009, **36**:1267–1274.
131. Weiner M, Albermann C, Gottlieb K, Sprenger GA, Weuster-Botz D: **Fed-batch production of L-phenylalanine from glycerol and ammonia with recombinant *Escherichia coli*.** *Biochem Eng J* 2014, **83**:62–69.
132. Liu SP, Liu RX, Xiao MR, Zhang L, Ding ZY, Gu ZH, Shi GY: **A systems level engineered *E. coli* capable of efficiently producing L-phenylalanine.** *Process Biochem* 2014, **49**:751–757.
133. Hu Y, Tang T, Yang W, Zhou H: **Bioconversion of phenylpyruvic acid to L-phenylalanine by mixed-gel immobilization of *Escherichia coli* EP8-10.** *Process Biochem* 2009, **44**:142–145.

doi:10.1186/s12934-014-0126-z

Cite this article as: Rodriguez et al.: Engineering *Escherichia coli* to overproduce aromatic amino acids and derived compounds. *Microbial Cell Factories* 2014 **13**:126.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Metabolic and transcriptional response of *Escherichia coli* with a NADP⁺-dependent glyceraldehyde 3-phosphate dehydrogenase from *Streptococcus mutans*

Sara Centeno-Leija · José Utrilla ·
Noemí Flores · Alberto Rodríguez ·
Guillermo Gosset · Alfredo Martínez

Received: 16 July 2013 / Accepted: 17 August 2013 / Published online: 29 August 2013
© Springer Science+Business Media Dordrecht 2013

Abstract The NAD⁺-dependent glyceraldehyde-3-phosphate-dehydrogenase (NAD⁺-GAPDH) is a key enzyme to sustain the glycolytic function in *Escherichia coli* and to generate NADH. In the absence of NAD⁺-GAPDH activity, the glycolytic function can be restored through NADP⁺-dependent GAPDH heterologous expression. Here, some metabolic and transcriptional effects are described when the NAD⁺-GAPDH gene from *E. coli* (*gapA*) is replaced with the NADP⁺-GAPDH gene from *Streptococcus mutans* (*gapN*). Expression of *gapN* was controlled by the native *gapA* promoter (*E. coli*Δ*gapA*::*gapN*) or by the constitutive *trc* promoter in a multicopy plasmid (*E. coli*Δ*gapA*::*gapN*/pTrc*gapN*). The specific NADP⁺-GAPDH activity was 4.7 times higher in *E. coli*Δ*gapA*::*gapN*/pTrc*gapN* than *E. coli*Δ*gapA*::*gapN*. Growth, glucose consumption and acetic acid production rates increased in agreement with the NADP⁺-GAPDH activity level. Analysis of *E. coli*Δ*gapA*::*gapN*/pTrc*gapN* showed that although *gapN* expression complemented NAD⁺-GAPDH activity, the resulting low NADH levels decreased the expression of the respiratory chain and oxidative phosphorylation genes (*ndh*, *cydA*, *cyoB* and

atpA). In comparison with the wild type strain, *E. coli*Δ*gapA*::*gapN*/pTrc*gapN* decreased the percentage of mole of oxygen consumed per mole of glucose metabolized by 40 % with a concomitant reduction of 54 % in the ATP/ADP ratio. The cellular response to avoid NADPH excess led to the overexpression of the transhydrogenase coded by *udhA* and the down-regulation of the pentose-phosphate and Krebs cycle genes, which reduced the CO₂ production and increased the acetic acid synthesis. The *E. coli* strains obtained in this work can be useful for future metabolic engineering efforts aiming for the production of metabolites which biosynthesis depends on NADPH.

Keywords *Escherichia coli* · NAD(P)⁺-dependent glyceraldehyde-3-phosphate-dehydrogenase · Respiration · ATP · NAD(P)(H)

Introduction

During glucose consumption, NADH generation in *Escherichia coli* is conducted through the Embden-Meyerhof-Parnas pathway (EMP) by the NAD⁺-dependent glyceraldehyde-3-phosphate-dehydrogenase (NAD⁺-GAPDH) (E.C. 1.2.1.12) enzyme encoded by *gapA* (D'Alessio and Josse 1971; Charpentier and Branlant 1994). NAD⁺-GAPDH activity is essential to growth and is known to have higher relative activity compared to other enzymes in the EMP pathway (Charpentier and Branlant 1994; Thouvenot et al.

S. Centeno-Leija · J. Utrilla · N. Flores ·
A. Rodríguez · G. Gosset · A. Martínez (✉)
Departamento de Ingeniería Celular y Biotecnología,
Instituto de Biotecnología, Universidad Nacional
Autónoma de México, Apartado postal 510-3,
62210 Cuernavaca, MOR, Mexico
e-mail: alfredo@ibt.unam.mx

2004). In fact, it has been reported that controlling the expression of *gapA* allows modulation of the NAD⁺-GAPDH activity and therefore the carbon flux through the EMP pathway (Cho et al. 2012). Under aerobic conditions the main role of NADH is to generate ATP by oxidative phosphorylation in the respiratory chain and it is also required for more than 300 reduction reactions (Foster et al. 1990). On the other hand, in anaerobiosis NADH leads to fermentation reactions to maintain the redox balance in absence of oxygen as a final electron acceptor (Moat et al. 2002). Therefore, negative growth effects are observed when the NADH levels are altered, mainly by redox stress (Kim et al. 2009, 2011). Furthermore, *E. coli* mutants without NAD⁺-GAPDH activity are unable to grow in the presence of hexoses, presumably due to the toxic accumulation of intermediates (Hillman and Fraenkel 1975; Irani and Maitra 1977).

Several genera of microorganisms, such as *Bacillus*, *Streptococcus* and *Clostridium*, have a NAD⁺-GAPDH enzyme (E.C. 1.2.1.9) in addition to a NAD⁺-GAPDH enzyme, allowing the production of NADH and NADPH through the EMP (Iddar et al. 2005). Since some species of *Streptococcus* lack the oxidative branch of the pentose phosphate pathway (PPP), the NAD⁺-GAPDH activity is used as an alternative mechanism to generate NADPH (Brown and Wittenberger 1971). Unlike NAD⁺-GAPDH, NADP⁺-GAPDH does not require inorganic phosphate and this reaction is cooperative and irreversible. Specifically, the NADP⁺-GAPDH enzyme from *Streptococcus mutans* coded by the *gapN* gene (Boyd et al. 1995) has a substrate affinity 10 times higher towards glyceraldehyde-3-P than the NAD⁺-GAPDH enzyme GapA from *E. coli*, nevertheless NADP⁺-GAPDH displays a strong inhibition by substrate and NADPH (Crow and Wittenberger 1979; Habenicht 1997; Marchal and Branlant 2002).

It has been demonstrated that some genes encoding NADP⁺-GAPDH enzymes can complement an *E. coli* mutant deficient in native NAD⁺-GAPDH (Fillinger et al. 2000; Iddar et al. 2002, 2003). These mutants are able to grow under aerobic conditions despite the decrease in the NADH levels necessary for energy generation, but not in anaerobic conditions, possibly due to the inability of the cell to sustain an efficient redox balance (Valverde et al. 1999; Martínez et al. 2008). Martínez and co-workers reported that the replacement of NAD⁺-GAPDH from *E. coli* with a

NADP⁺-GAPDH enzyme from *Clostridium acetobutyricum* (GapC) increased the NADPH yield on glucose 1.7 times (Martínez et al. 2008) and a metabolic flux analysis revealed that the oxidative PPP branch and tricarboxylic acid cycle (TCA) fluxes were significantly reduced in the mutant $\Delta gapA::gapC$, presumably to avoid high NADPH levels. Accordingly, heterologous lycopene production, a NADPH-dependent metabolite, increased 2.5 times in rich medium (Martínez et al. 2008). In fact, because *E. coli* has a limited capacity to generate NADPH through catabolic reactions (Csonka and Fraenkel 1977; Sauer et al. 2004; Fuhrer and Sauer 2009) the production of NADPH through the EMP could be exploited to produce NADPH-dependent metabolites of industrial interest, such as lycopene, caprolactone, L-lysine and polyhydroxybutyrate (Martínez et al. 2008; Takeno et al. 2010; Kim et al. 2011; Kocharin et al. 2013). However, the physiological effects resulting from the replacement of GAPDH activity and the concomitant production of NADPH through the EMP pathway has not been extensively studied in *E. coli*. In this work, a quantitative analysis was performed with data obtained from kinetic parameters, respiration rates, ATP/ADP and NAD(P)/NAD(P)H ratios and transcriptional data, providing further insights on the physiological response of *E. coli* to the replacement of NAD⁺-GAPDH with a NADP⁺-GAPDH from *S. mutans* when growing under aerobic conditions on glucose mineral media.

Materials and methods

Bacterial strains and plasmids

The wild-type *E. coli* MG1655 (GenBank NC 000913) was used as the parental strain and *S. mutans* (provided by the Culture Collection of the Faculty of Chemistry of the Universidad Nacional Autónoma de México) was used as a source of DNA for *gapN* amplification.

All plasmids and primers used in this work are listed in Table 1. Standard procedures were employed for plasmid preparations, restriction-enzyme digestions, transformations and gel electrophoresis (Sambrook and Russell 2001). Each plasmid construction and mutants were verified by their restriction pattern in agarose gels and by sequencing. The chromosomal DNA was isolated with the Ultra Clean Microbial

Table 1 Plasmids and primers used in this work

Plasmid	Description	Source
pTrc99A	Designed for IPTG-inducible expression of proteins under the control of the hybrid <i>trp/lac</i> promoter. Amp ^r	Amann et al. (1988)
pACYC184	Designed with p15A origin of replication to coexist in cells with plasmids of the ColE1 compatibility group (e.g., pBR322, pUC19). Tc ^r and Cm ^r	Chang and Cohen (1978)
pTgapN/cat- <i>frt</i>	pTrc99A derivative with the <i>gapN</i> gene fused with the <i>cat-<i>frt</i></i> fragment	This study
pGapN	pTrc99A derivative containing the <i>gapN</i> gene	This study
pTrcgapN	pACYC184 derivative with <i>trc</i> promoter and <i>gapN</i> gene	This study
pKD46	Designed to inactivate chromosomal gene through phage λ Red recombinase	Datsenko and Wanner (2000)
pKD3	Designed to contain an FRT-flanked chloramphenicol resistance (<i>cat</i>) gene	Datsenko and Wanner (2000)
pCP20	Designed to remove FRT-flanked resistance markers	Cherepanov and Wackernagel (1995)
Primer	Sequence ^a	
gapN1	5' <u>CATGCCATGGCAATGACAAAACAATATAAAAATTA</u> 3'	
gapN2	5' <u>CGGGGTACCCCGTTATTTGATATCAAATACGACGG</u> 3'	
Trc1	5' <u>GCCGACATCATAACGGTTCTGG</u> 3'	
cm1	5' <u>CGGGGTACCTGTAGGCTGGAGCTGCTTCG</u> 3'	
cm2	5' <u>TCCCCCGGG</u> CATATGAATATCCTCCTTA 3'	
D1 ^b	5'GTAATTTTACAGGCAACCTTTTATTCACTAACAAATAGCTGGTGAATATATG ACAAAACAATATAAAAATTATG 3'	
D2 ^b	5'AAAAAAGAGCGACCGAAGTCGCTCTTTTATAGATCACAGTG TCATCTCAACCATATGAATATCCTCCTTAGTTCCT 3'	

^a Restriction sites employed during plasmid construction are underlined

^b Sequences that flank ~580 pb upstream and downstream of *gapA* gen are indicated in bold

DNA Isolation kit (MO BIO laboratories, Inc, Carlsbad, CA, USA). Gene amplifications were carried out by PCR on a C1000 Touch Thermal Cycler Manual (Bio-Rad Laboratories, Inc., Berkeley, CA, USA) using Expand High Fidelity polymerase (Roche Diagnostics GmbH, Mannheim, Germany). DNA fragments were isolated from agarose gels with the Roche Pure PCR Product Purification kit (Roche Diagnostics GmbH, Mannheim, Germany) according to the manufacturer's instructions.

The *gapA* gene (GenBank gene ID947679) from *E. coli* MG1655 was replaced by the gene *gapN* from *S. mutans* (GenBank Accession No. L38521). The MG1655 Δ *gapA::gapN* mutant strain was made using a variation of the method reported by Datsenko and Wanner (2000) through the fusion of the *gapN* gene with a chloramphenicol resistance cassette. Using *S. mutans* chromosomal DNA as a template, the *gapN* gene was amplified by PCR with gapN1 and gapN2

primers to add NcoI and KpnI restriction sites at the 5' and 3' ends, respectively. The Chloramphenicol resistance cassette flanked by FRT sites (*cat-*frt** fragment) was amplified by PCR with cm1 and cm2 primers using the pKD3 plasmid as a template (Datsenko and Wanner 2000). The *cat-*frt** fragment contained KpnI and XmaI restriction sites at the 5' and 3' ends respectively. The 3' end of the *gapN* gene and 5' end of the *cat-*frt** fragment was fused through the KpnI restriction site and this construction was named *gapN-cat-*frt** cassette. The *gapN-cat-*frt** cassette was inserted into the pTrc99A vector (Amann et al. 1988), which was digested with NcoI and XmaI, to generate pTgapN/cat-*frt* plasmid. The *gapN-cat-*frt** cassette was integrated into the *gapA* locus in *E. coli* MG1655 chromosome to allow transcriptional control of *gapN* by the *gapA* promoter region. The *gapN-cat-*frt** cassette was amplified from the pTgapN/cat-*frt* plasmid, using D1 and D2 primers with 50 nucleotides of homology to the *gapA* promoter

sequence and the beginning and end of the *gapN-cat-frt* cassette. *E. coli* MG1655 was transformed with the pKD46 plasmid (Datsenko and Wanner 2000), which promotes homologous recombination at the designed sites of the *gapN-cat-frt* cassette. Recombinant strains were selected on Luria–Bertani (LB) solid medium supplemented with 10 $\mu\text{g mL}^{-1}$ of chloramphenicol. The MG1655 Δ *gapA::gapN* mutant selected strain was verified by DNA sequencing with primers S1 (TAACGAATGGATTCTTCACTTACCGGTTTCG) and S2 (TATGCAGGGCAGAGGTGGTTTCAAATTCGC), which hybridize at \sim 580 nucleotides upstream and downstream positions, respectively, of *gapA* gene. The chloramphenicol marker was removed in the MG Δ *gapA::gapN* strain, using the procedure reported by Datsenko and Wanner through pCP20 plasmid (Che-repanov and Wackernagel 1995).

Plasmid pACYC184 (Chang and Cohen 1978) was used to overexpress the *gapN* gene. First, this gene was amplified by PCR with oligonucleotides *gapN1* and *gapN2* that introduce *NcoI* and *KpnI* flanking sites to facilitate cloning into plasmid pTrc99A. The resulting plasmid pGapN was used as a PCR template to amplify the *trcgapN* fragment with oligonucleotides *Trc1* and *gapN2*. This fragment was blunt-ligated into pACYC184 previously digested with *ScaI* to generate plasmid pTrcgapN (Table 1). The construction was verified by digestion with *EcoRI* and visualized by agarose gel electrophoresis.

Medium and growth conditions

The routine cultivation and genetic manipulation of *E. coli* MG1655 and its mutant were performed in Luria–Bertani medium (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, and 10 g L^{-1} NaCl, pH 7.2) supplemented with the appropriate antibiotics, when necessary. *S. mutans* strain was cultivated in blood agar (Merk KGaA, Darmstadt Germany) at 37 °C. Cells were harvested in a saline solution (0.9 %) in order to isolate the chromosomal DNA. For characterization of the MG Δ *gapA::gapN* harbouring pTrcgapN plasmid, *E. coli* cultivations were performed in 500 mL shake flasks containing 100 mL of M9 medium, (per liter): 6 g Na_2HPO_4 ; 3 g KH_2PO_4 ; 0.5 g NaCl; 1 g NH_4Cl ; 2 mM MgSO_4 ; 0.1 mM CaCl_2 ; 0.01 g Vit B1, and 8 g glucose at 37 °C with an agitation of 300 rpm.

To characterize the strains, *E. coli* MG1655 Δ *gapA::gapN*/pTrcgapN was compared with *E. coli*

MG1655 wild-type strain and MG Δ *gapA::gapN* carrying an empty pACYC184 plasmid. Cultures were performed in pH-controlled fermenters. Seed cells were prepared in a 2.8 L Fermbach flask containing 500 mL of M9 medium with 5 g L^{-1} glucose. The cells from 350 mL of culture were harvested by centrifugation at an OD_{600} of 1 and resuspended with 10 mL of fresh M9 medium; these cells were used as inoculum for the bioreactor. Fermentations were conducted in a 1-L bioreactor (Applikon Biotechnology, Netherlands) with a working volume of 0.7 L, 600 rpm, pH controlled at 7 with NaOH (2N), and an air flow rate of 0.84 L min^{-1} . The samples were collected in the mid-log phase at an OD_{600} of 3 to determine: cofactors, GAPDH enzymatic activity, and RNA extraction for transcriptional analysis. All experiments and analysis were performed in triplicates, figures and tables show averages and standard deviations.

Metabolite analysis and kinetic and stoichiometric parameters

The samples were centrifuged (4 °C), and the cell-free culture broth was frozen for subsequent analysis. The concentration of glucose and acetic acid in the culture broths was quantified using an HPLC (1200, Agilent, CA, USA) equipped with a refractive index detector and an ion-exchange column (50 °C) ($300 \times 78 \text{ mm}^2$; Aminex HPX-87H; BioRad, USA) using a mobile phase of 0.005 mol L^{-1} H_2SO_4 and a flow rate of 0.5 mL min^{-1} during elution. Respiration rate (q_{O_2}) and CO_2 production rate (q_{CO_2}) were calculated from a mass balance and the composition of inlet and outlet gases from the fermenter and measured by O_2 - and CO_2 -gas sensors (BlueSens, gas sensor GmbH, Herthen, Germany).

The data in figures and tables is the average of at least three different cultures. Cell growth was measured by monitoring the optical density at 600 nm (OD_{600}) in a spectrophotometer (DU-70, Beckman Instruments, Inc. Fullerton, CA, USA). OD_{600} was converted into dry cellular weight (cell concentration) using a calibration curve (1 $\text{OD}_{600} = 0.42 \text{ g L}^{-1}$ of dry cellular weight). The specific growth rates (μ) were determined by fitting the biomass data versus time to exponential regressions. The cell mass yield on glucose ($Y_{X/\text{GLC}}$), yield of acetate on glucose ($Y_{\text{ACE}/\text{GLC}}$), yield of CO_2 ($Y_{\text{CO}_2/\text{GLC}}$) and O_2 ($Y_{\text{O}_2/\text{GLC}}$) on glucose were estimated as the coefficient of the linear

regression of the respective metabolite concentration versus the concentration of glucose consumed during the exponential growth phase, in grams of biomass/ mmol_{GLC} or $\text{mmol}_{\text{PRODUCT}}/\text{mmol}_{\text{GLC}}$. The specific glucose consumption rate (q_s) was determined as the ratio of μ to $Y_{x/s}$. Respiration rate (q_{O_2}) and CO_2 production rate (q_{CO_2}) were calculated from mass balances and the specific growth rate.

Measurements of the intracellular cofactors concentrations during exponential growth phase

The intracellular cofactors, NAD^+ , NAD(H) , NAD(P)H and NAD(P)^+ , were extracted and assayed using the EnzyChromTM assay kit following the supplier's instructions (BioAssay Systems, Hayward, CA, USA). Briefly, three samples of approximately 12.6 mg of wet cells from cultures were immediately received in methanol (70 % v/v) at -50°C (Lee et al. 2010), for rapid inactivation of the cellular metabolism. The cell pellet was washed with cold PBS and resuspended with base or acid buffer (BioAssay Systems, Hayward, CA, USA) to extract the reduced or oxidized pyridine nucleotides (Bergmeyer 1985). Relative amounts of NAD^+ , NADH , NADP^+ , and NADPH were quantified by enzymatic methods (Bergmeyer 1985) and using NADP^+ -glucose-6-phosphate dehydrogenase and NAD^+ -lactate dehydrogenase (BioAssay Systems, Hayward, CA, USA).

To measure the $[\text{ATP}]/[\text{ADP}]$ ratio, three samples of ~ 5 mg of wet cells were received in 2 mL of phenol (equilibrated with 10 mM Tris–1 mM EDTA to pH 8) at 80°C , samples were treated as reported by Koebmann (Koebmann et al. 2002). The ATP concentration was then measured by using a luciferin-luciferase kit (ENLITEN[®] ATP Assay System Bioluminescence Detection kit; Promega Corporation, USA) following the supplier's instructions. Subsequently, the ADP was converted to ATP adding one enzymatic unit of pyruvate kinase and 1 mM phosphoenolpyruvate, then the total ATP was quantified as described above and the ADP concentration was calculated from the difference of both measurements.

Enzymatic assays

The equivalent to 6.3 mg of cells was collected by centrifugation at $4,000\times g$ (10 min at 4°C) and then

washed once with cold buffer A (20 mM Tricine buffer, pH 8.5, containing 3 mM 2-mercapto-ethanol). The cells were suspended in 500 μL of the same buffer and sonicated with 3 pulses of 45 s and pauses of 30 s at 4°C using a sonicator (Soniprep 150 ultrasonic disintegrator, MSE Ltd, London, U.K.). Cell debris was removed by centrifugation ($4,000\times g$, 10 min, 4°C). The resulting supernatant was used for the enzymatic assays. The GAPDH activity was measured as described in (Iddar et al. 2003). The NADP^+ -GAPDH reaction was started by adding the cell extract to the assay mixture containing buffer A, 1 mM NADP^+ , and 1 mM D-glyceraldehyde-3-phosphate (G-3-P) at 25°C . The NAD^+ -GAPDH reaction was measured using the same procedure with 1 mM NAD^+ and 10 mM AsO_4^{3-} . The absorbance variation at 340 nm was followed in a spectrophotometer (DU-70, Beckman Instruments, Inc. Fullerton, CA, USA). One specific unit of activity is defined as 1 μmol of NAD(P)H formed per min per mg of protein ($\text{IU}/\text{mg}_{\text{PROT}}$). The protein concentration was determined by the Bradford method (Bradford 1976) using the Bio-Rad reagent and BSA as standard (Bio-Rad).

RNA extraction, DNase treatment of RNA and cDNA synthesis for RT-qPCR analysis

Total RNA was isolated and purified using a previously reported method (Aguilar et al. 2012) with the equivalent to 11.1 mg of cells growing logarithmically in the fermenter at an $\text{OD}_{600\text{nm}}$ of 3. Samples were then suspended in 300 μL of DNase and RNase-free water (Ambion Inc, Austin, TX, USA) with RNase inhibitor (Fermentas Life Sciences, USA). RNA was analyzed on a formaldehyde agarose gel for integrity. The RNA concentrations were quantified using a Nanodrop (2000c, Thermo Scientific); the 260/280 and 260/230 ratios were examined for protein and solvent contamination. For all samples the 260/280 nm absorbance values were from 1.9 to 2.0 and in the range of 2.0–2.3 for the 260/230 nm ratio. RNA samples were stored at -70°C . Three RNA extractions and purifications were carried out from three independent fermentations for each strain. For DNase treatment, total RNA samples were treated with the TURBO DNA-free kit (Ambion Inc, Austin, TX, USA) at 37°C for 30 min, following the manufacturer's instructions. The cDNA was synthesized using RevertAidTM H minus First Strand cDNA Synthesis kit following the

manufacturer's conditions (Fermentas LifeSciences, USA). For each reaction approximately 5 µg of RNA and a mixture of 10 pmol µL⁻¹ of specific DNA reverse primers (b) were used. This cDNA was used as template for RT-qPCR assays. Nucleotide sequences of these primers have been previously published (Flores et al. 2005; Chávez-Béjar et al. 2008; Sigala et al. 2009; Aguilar et al. 2012).

RT-qPCR

RT-qPCR was performed with the ABI Prism 7000 Sequence Detection System and 7300 Real Time PCR System (Perkin Elmer/Applied Biosystems, USA) using the Maxima^R SYBR Green/ROX qPCR Master Mix (2X) kit (Fermentas LifeSciences, USA). MicroAmp Optical 96-well reaction plates (Applied Biosystems, USA) and Plate Max ultraclear sealing films (Axygen Inc, USA) were used in these experiments. All RT-qPCR quantifications were performed as previously reported (Aguilar et al. 2012) and complied with the MIQE guidelines (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) (Bustin et al. 2009; Taylor et al. 2010). The quantification technique used to analyze the data was the 2^{-ΔΔC_q} method described by Livak and Schmittgen (2001). For each analyzed gene the transcription levels of the corresponding MG1655 gene were considered equal to one and used as controls to normalize the data. Therefore, data are reported as relative expression levels compared to the expression level of the same gene in strain MG1655, using *ihfB* as housekeeping gene.

Results and discussion

Metabolic responses to *gapN* expression

In order to study the metabolic effects when the NAD⁺-GAPDH gene from *E. coli* (*gapA*) is replaced with NADP⁺-GAPDH gene from *S. mutans* (*gapN*), two derivative strains were generated using *E. coli* strain MG1655. Strain MG1655Δ*gapA*::*gapN* was constructed in one step by inserting the *gapN* gene at the same locus as the native *gapA* gene into the chromosome of strain MG1655. In this strain, the *gapN* gene expression was controlled by the *gapA* promoter region (Charpentier and Branlant 1994; Thouvenot et al. 2004). Also, the plasmid pTrc*gapN* was constructed to

overexpress the *gapN* gene from the *trc* promoter (see “Materials and methods” section) and the mutant strain MG1655Δ*gapA*::*gapN* was transformed with plasmid pTrc*gapN* (MG1655Δ*gapA*::*gapN*/pTrc*gapN*) to increase the *gapN* expression level.

Fermenter batch cultures with mineral medium were carried out supplemented with 8 g L⁻¹ of glucose under aerobic conditions. Kinetic data (Fig. 1a, b) showed that the MG1655Δ*gapA*::*gapN* strain had a specific growth rate (μ) of 0.11 ± 0.03 h⁻¹ and a glucose consumption rate (q_s) of 2.06 ± 0.17 mmol_{glc} g_{DCW}⁻¹h⁻¹ (Table 2). The μ and q_s of MG1655Δ*gapA*::*gapN*/pTrc*gapN* were 4 and 2.5 times higher than those parameters in MG1655Δ*gapA*::*gapN*, respectively. Since *E. coli* mutants of the *gapA* gene are unable to grow with hexoses as carbon source (Hillman and Fraenkel 1975; Irani and Maitra 1977; Baba et al. 2006) these data indicate that *gapN* expression can replace the native *gapA* function. Furthermore, the specific NADP⁺-GAPDH activity of the strain MG1655Δ*gapA*::*gapN* was 4.6 times lower relative to strain MG1655Δ*gapA*::*gapN*/pTrc*gapN* (Table 2) and no NAD⁺-GAPDH activity was detected (Table 2). The specific NADP⁺-GAPDH activity levels in the derivative strain reveal that growth rate and glucose uptake differences are attributable to *gapN* expression level. Indeed, it has been demonstrated in *E. coli* that carbon flux through the EMP pathway can be controlled through the glyceraldehyde 3-phosphate node by NAD⁺-GAPDH activity (Cho et al. 2012) and our results show that NADP⁺-GAPDH heterologous activity can also exert a similar response.

On the other hand, very low acetic acid was produced by strain MG1655Δ*gapA*::*gapN* (0.04 ± 0.00 mol_{ACE} mol_{GLC}⁻¹) (Fig. 1c), which is expected since low glucose flux through the EMP pathway prevents the acetic acid synthesis (Hollywood and Doelle 1976; Andersen and von Meyenburg 1980). Moreover, this characteristic is desirable to improve the production of recombinant protein and increase the carbon availability to produce several metabolites (Dittrich et al. 2005; Kern et al. 2007; Waegeman et al. 2013). In contrast, MG1655Δ*gapA*::*gapN*/pTrc*gapN* directed ~40 % of substrate to acetic acid production (0.78 ± 0.02 mol_{ACE} mol_{GLC}⁻¹) (Table 2; Fig. 1c). High acetic acid levels were produced when a NADP⁺-GAPDH from *C. acetobutyricum* replaced a NAD⁺-GAPDH from *E. coli* to produce lycopene

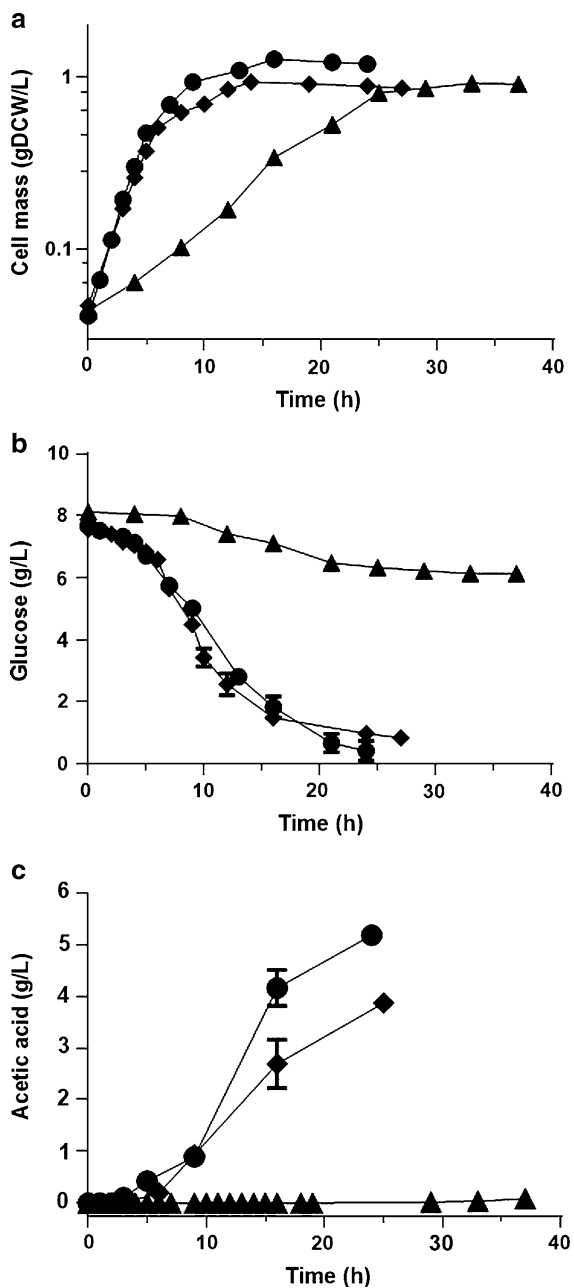


Fig. 1 Kinetic of cell mass formation (a), glucose consumption (b) and acetic acid production (c), strains: MG1655 (circles), MG1655ΔgapA::gapN (triangles) and MG1655ΔgapA::gapN/pTrcgapN (diamonds)

(Martínez et al. 2008), which could affect the lycopene yield. However, a decrease in acetic acid production is possible by deleting the *ackA-pta* pathway (Phue et al. 2010) or by using a strain with low glycolytic flux (Hollywood and Doelle 1976; Andersen and von

Meyenburg 1980) as MG1655ΔgapA::gapN. These results show that the heterologous NADP⁺-GAPDH enzyme from *S. mutans* not only replaced the native NAD⁺-GAPDH function, but also, as cited above, the carbon flux through the EMP pathway could be controlled by *gapN* expression level, which modulated the glucose uptake and also the acetic acid production rates.

In comparison with the parental strain MG1655, the specific growth and glucose consumption rates in strain MG1655ΔgapA::gapN decreased 5 and 3.4 times, respectively (Table 2; Fig. 1a, b). In agreement with that, the specific NADP⁺-GAPDH activity in strain MG1655ΔgapA::gapN was 2.6 times lower than the NAD⁺-GAPDH activity in wild-type strain (Table 2). Although the *gapN* gene expression was controlled by the same *gapA* native promoter, the differences in GAPDH activity levels could be attributable to inherent differences in the catalytic properties of each enzyme (D'Alessio and Josse 1971; Crow and Wittenberg 1979). In contrast, MG1655ΔgapA::gapN/pTrcgapN, has a μ and a q_s 15 and 25 % lower when compared with the wild-type strain (Table 2; Fig. 1a, b). However, the NADP⁺-GAPDH activity was 1.8 times higher than the NAD⁺-GAPDH activity in the parental strain (Table 2). This means that the increase of NADP⁺-GAPDH heterologous activity can partially restore growth and glucose uptake of the MG1655 strain. In order to elucidate the metabolic effects when NADPH is produced instead of NADH via glyceraldehyde 3-phosphate oxidation, the cofactor levels, transcriptional response and respiration rate were determined in MG1655ΔgapA::gapN/pTrcgapN.

Metabolic and transcriptional responses to replacement of NAD⁺-GAPDH by NADP⁺-GAPDH activity

As shown in Fig. 2, the NADH/NAD⁺ ratio in the strain MG1655ΔgapA::gapN/pTrcgapN decreased 25 % as compared to wild-type strain. In contrast, the NADPH/NADP⁺ ratio increased 2 times indicating that the alteration in the turnover of NAD(P)H via glyceraldehyde 3-phosphate oxidation affected the redox levels of the strain MG1655ΔgapA::gapN/pTrcgapN, which increased 2.8-times the NADPH/NADH ratio (Fig. 2). Concomitantly, the ATP/ADP

Table 2 Growth parameters of the strains MG1655 Δ gapA::*gapN*, MG1655 Δ gapA::*gapN*/pTrc*gapN* and MG1655 from batch cultures in glucose-mineral medium

Strain	μ (h ⁻¹)	q_s (mmol _{GLC} g _{DCW} ⁻¹ h ⁻¹)	$Y_{ACE/GLC}$ (mol _{ACE} mol _{GLC} ⁻¹)	NAD ⁺ -GAPDH activity IU/mg _{PROT}	NADP ⁺ -GAPDH activity IU/mg _{PROT}
MG1655 Δ gapA:: <i>gapN</i>	0.11 (0.03)	2.06 (0.17)	0.04 (0.00)	ND	0.201 (0.13)
MG1655 Δ gapA:: <i>gapN</i> /pTrc <i>gapN</i>	0.44 (0.01)	5.24 (0.13)	0.78 (0.02)	ND	0.937 (0.04)
MG1655	0.52 (0.01)	7.04 (0.05)	0.61 (0.07)	0.522 (0.21)	ND

Values in parenthesis indicate the standard deviations

ND Not detected

ratio in the MG1655 Δ gapA::*gapN*/pTrc*gapN* strain decreased 2 times when compared to that of the parental strain (Fig. 2). Moreover, the transcriptional analysis revealed that the expression level of the genes related to the respiratory chain and oxidative phosphorylation such as *nuo* (NADH:ubiquinone oxidoreductase I), *ndh* (ubiquinone oxidoreductase II), *cydA* (cytochrome *bd* oxidase subunit I), *cyoB* (cytochrome *bo* oxidase subunit I) and *atpA* (ATP syntase F₁ complex subunit I) decreased about 50 % in relation to the expression of these genes in the control strain (Fig. 3b). In agreement, several studies have demonstrated that the alteration in the energy levels disturbs the expression of respiration and oxidative phosphorylation genes (Jensen and Michelsen 1992; Noda et al. 2006). Accordingly, the specific respiration rate (q_{O_2}) and yield of consumed oxygen on glucose decreased ($Y_{O_2/GLC}$) ~20 and 40 % respectively (Table 3). Since the production of NADH is coupled to the respiratory chain to provide ATP through oxidative phosphorylation (Andersen and von Meyenburg 1980; Koebmann et al. 2002; Vemuri et al. 2006; Holm et al. 2010), these data suggest that the decreased respiration and energy levels were caused by a decrease in the expression of the respiratory and oxidative phosphorylation genes as a response to low NADH production.

Interestingly, the expression of the genes that code for the succinate dehydrogenase A and C (*sdhA* and *sdhC* respectively) and fumarase A (*fumA*) and fumarase B (*fumB*) were also increased more than 60 % when compared with the control strain (Fig. 3a, b). SdhA and SdhB are involved in complex II of the respiratory chain using FADH₂ as an electron donor (Brandsch 1989), while FumA and FumB produce NADH across the cytoplasmic membrane (Yagi and Matsuno-yagi 2003). The expression of both systems

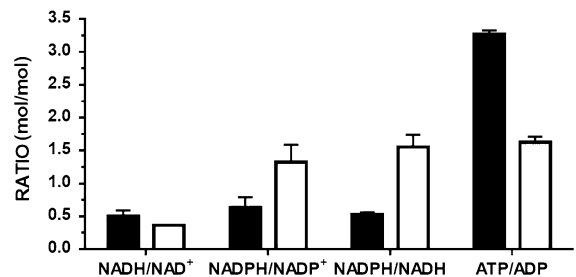


Fig. 2 NADPH/NADP⁺, NADH/NAD⁺, NADPH/NADH and ATP/ADP ratios (mol/mol) of strains MG1655 (*fill*) and MG1655 Δ gapA::*gapN*/pTrc*gapN* (*empty*)

depends on the intracellular NADH state, determined by the type of carbon source or by oxygen presence (Kalman and Gunsalus 1988; Bogaerts et al. 1995). Because the NADH levels were altered in MG1655 Δ gapA::*gapN*/pTrc*gapN*, it is possible that the expression of these oxidoreductases acts as a response to the low NADH levels.

On the other hand, transcriptional analysis indicated that the expression of genes from the EMP in MG1655 Δ gapA::*gapN*/pTrc*gapN* did not significantly change (>0.5 relative to parental strain) (Fig. 3a). This was expected since the expression of glycolytic genes depends on global transcriptional regulators such as Crp, which responds when the carbon uptake is limited (Takahashi et al. 1998) and as was shown, is not the case of MG1655 Δ gapA::*gapN*/pTrc*gapN*.

The expression of the *udhA* gene, encoding transhydrogenase UdhA, increased 2.4 times in the MG1655 Δ gapA::*gapN*/pTrc*gapN* strain (Fig. 3c) presumably to generate NADH from an excess of NADPH (Canonaco et al. 2001; Sauer et al. 2004; Holm et al. 2010; Charusanti et al. 2010). Likewise,

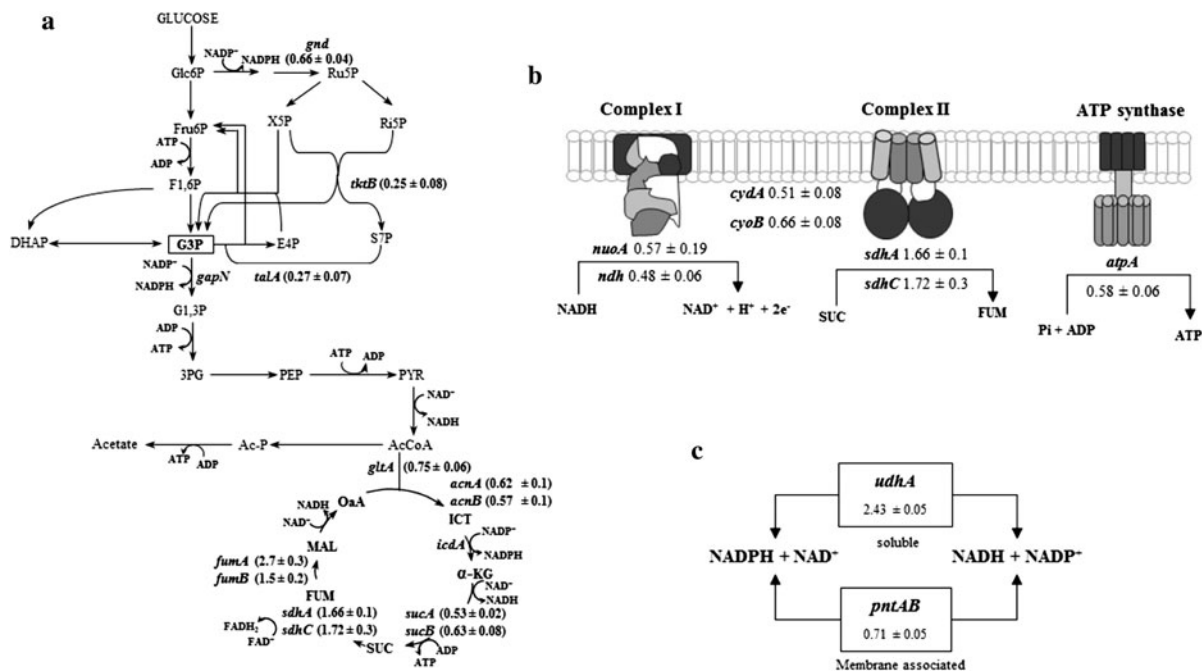


Fig. 3 Transcript levels of genes involved in: carbon central metabolism (a), respiratory chain-oxidative phosphorylation (b) and transhydrogenase system (c) in *MGΔgapA::gapN/pTrcgapN* strain. The values are expressed relative to the wild-type strain MG1655. According to the significance criterion, only those relative gene transcription values ≥ 1.5 (up-regulation) or ≤ 0.5 (down-regulation), when compared to the MG1655 reference strain, are shown. Metabolite abbreviations: Embden-Meyerhof-Parnas Pathway: *Glc6P* glucose-6-phosphate, *Fru6P* fructose-6-phosphate, *F1,6P*, fructose-1,6-

biphosphate, *DHAP* dihydroxyacetone phosphate, *G3P* glyceraldehyde 3-phosphate, *G1,3P* 1,3-biphosphoglycerate, *3PG* 3-phosphoglycerate, *PEP* phosphoenolpyruvate, *PYR* pyruvate. Pentoses Phosphate Pathway: *Ru5P* ribulose-5-phosphate, *Ri5P* ribose-5-phosphate, *X5P* xylulose-5-phosphate, *S7P* sedoheptulose-7-phosphate, *E4P* erythrose-4-phosphate. Tricarboxylic Acid Cycle: *AcCoA* acetyl coenzyme A, *ICT* isocitrate, α -*KG* α -ketoglutarate, *SUC* succinate, *FUM* fumarate, *MAL* malate, *OaA* oxaloacetate. Acetic Acid Pathway: *Ac-P* acetyl phosphate

Table 3 Kinetic parameters of strains MG1655 and *MG1655ΔgapA::gapN/pTrcgapN* from batch cultures in glucose mineral medium

Strain	q_{O_2} (mmol_{O_2} $\text{g}_{DCW}^{-1} \text{h}^{-1}$)	q_{CO_2} (mmol_{CO_2} $\text{g}_{DCW}^{-1} \text{h}^{-1}$)	q_{ACE} (mmol_{ACE} $\text{g}_{DCW}^{-1} \text{h}^{-1}$)	$Y_{CO_2/GLC}$ (mol mol^{-1})	$Y_{O_2/GLC}$ (mol mol^{-1})	$Y_{ACE/GLC}$ ($\text{mol}_{ACE} \text{mol}_{GLC}^{-1}$)
MG1655	16.98 (0.35)	16.84 (0.25)	4.97 (0.32)	0.88 (0.02)	0.92 (0.03)	0.61 (0.07)
MG1655	14.01	13.89	4.85	0.54	0.55	0.78
<i>ΔgapA::gapN/pTrcgapN</i>	(0.26)	(0.16)	(0.45)	(0.03)	(0.03)	(0.02)

Values in parenthesis indicate the standard deviations

the expression of key genes from the oxidative branch of PPP pathway such as *gnd* (6-phosphogluconate dehydrogenase), *tktB* (transketolase II) and transaldolase A (*talA*) decreased around 34 and 75 % when compared to the parental strain (Fig. 3a), suggesting a low activity of the oxidative branch of PPP. Because

PPP is one of the most important routes to produce NADPH (Csonka and Fraenkel 1977; Fuhrer and Sauer 2009), these data concur with the previous observation that high NADPH levels decrease carbon flux through PPP (Kabir and Shimizu 2003; Martínez et al. 2008).

The expression of genes involved in directing acetyl-CoA through TCA cycle such as *gltA* (citrate synthase), *acnA* (aconitate hydratase 1), *acnB* (bifunctional aconitate hydratase 2 and 2-methylisocitrate dehydratase), *sucA* (2-oxoglutarate decarboxylase) and *sucB* (dihydrolipoyltranssuccinylase) were $\sim 50\%$ downexpressed in the MG1655 $\Delta gapA::gapN/pTrcgapN$ strain (Fig. 3a) suggesting that the TCA cycle was less active than in the parental strain. Furthermore, the moles of CO₂ produced per mole of glucose consumed decreased 39% in MG1655 $\Delta gapA::gapN/pTrcgapN$ strain (Table 3). Concomitantly, the acetic acid yield on glucose ($Y_{Acce/Glc}$) was 22% higher in MG1655 $\Delta gapA::gapN/pTrcgapN$ when compared to the parental strain. Such data indicate that in strain MG1655 $\Delta gapA::gapN/pTrcgapN$ more acetyl-CoA was directed to produce acetic acid instead of entering the TCA cycle, this could be a response to increase ATP levels via AckA-Pta and reduce NADPH production by isocitrate dehydrogenase activity (*IcdA*). A similar observation was reported by Martínez et al., showing that an *E. coli* $\Delta gapA::gapC$ mutant also exhibited low flux through the acetyl-CoA node (Martínez et al. 2008).

Conclusions

The NADP⁺-GAPDH from *S. mutans* replaced the NAD⁺-GAPDH native function in *E. coli*. The growth, glucose uptake and acetic acid production rates were controlled by the NADP⁺-GAPDH heterologous activity level. Although the overexpression of the NADP⁺-GAPDH gene can restore the growth in the NAD⁺-GAPDH mutant to similar levels as those shown by the wild type strain, the replacement increased the NADPH levels with a concomitant decrease in the NADH levels. The metabolic and transcriptional analysis indicated that the redox perturbation restricted NADPH production through other native pathways and limited respiratory capabilities, decreasing the energy levels.

Acknowledgments We gratefully acknowledge Georgina Hernández-Chávez, Luz María Martínez, Mercedes Enzaldo, Ramón de Anda and Adolfo Escalante for technical assistance. We thank Martha Giles (Facultad de Química—Universidad Nacional Autónoma de México) for providing the *Streptococcus mutans* strain. This work was supported by the Mexican Council of Science and Technology (CONACyT). SC-L held a scholarship from CONACyT.

References

- Aguilar C, Escalante A, Flores N et al (2012) Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system. *BMC Genom* 13:385
- Amann E, Ochs B, Abel K (1988) Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* 69:301–315
- Andersen KB, von Meyenburg K (1980) Are growth rates of *Escherichia coli* in batch cultures limited by respiration? *J Bacteriol* 144:114–123
- Baba T, Ara T, Hasegawa M et al (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2:2006.0008
- Bergmeyer H (1985) *Methods of enzymatic analysis*, 3rd edn. VCH Verlagsgesellschaft mbH, Weinheim
- Bogaerts J, Zoske S, Weidner U, Linden G (1995) Transcriptional regulation of the proton translocating NADH dehydrogenase (*nuoA-N*) of *Escherichia coli* by electron acceptors, electron donors and gene regulators. *Mol Microbiol* 16:521–534
- Boyd DA, Cvitkovitch DG, Hamilton IANR (1995) Sequence, expression, and function of the gene for the phosphate dehydrogenase of *Streptococcus mutans*. *J Bacteriol* 177:2622–2627
- Bradford 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
- Brandsch R (1989) Covalent cofactor binding to flavoenzymes requires specific effectors. *Eur J Biochem* 182:125–128
- Brown AT, Wittenberger CL (1971) Mechanism for regulating the distribution of glucose carbon between the Embden-Meyerhof and hexose-monophosphate pathways in *Streptococcus faecalis*. *J Bacteriol* 106:456–467
- Bustin SA, Benes V, Garson JA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622
- Canonaco F, Hess TA, Heri S et al (2001) Metabolic flux response to phosphoglucose isomerase knock-out in *Escherichia coli* and impact of overexpression of the soluble transhydrogenase UdhA. *FEMS Microbiol Lett* 204:247–252
- Chang ACY, Cohen SN (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* 134:1141–1156
- Charpentier B, Branlant C (1994) The *Escherichia coli gapA* gene is transcribed by the vegetative RNA polymerase holoenzyme Eσ70 and by the heat shock RNA polymerase Eσ32. *J Bacteriol* 176:830–839
- Charusanti P, Conrad TM, Knight EM et al (2010) Genetic basis of growth adaptation of *Escherichia coli* after deletion of *pgi*, a major metabolic gene. *PLoS Genet* 6(11):e1001186
- Chávez-Béjar M, Lara AR, López H et al (2008) Metabolic engineering of *Escherichia coli* for L-tyrosine production by expression of genes coding for the chorismate mutase domain of the native chorismate mutase-prephenate dehydratase and a cyclohexadienyl dehydrogenase from

- Zymomonas mobilis*. Appl Environ Microbiol 74:3284–3290
- Cherepanov PP, Wackernagel W (1995) Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene 158:9–14
- Cho H-S, Seo SW, Kim YM et al (2012) Engineering glyceraldehyde-3-phosphate dehydrogenase for switching control of glycolysis in *Escherichia coli*. Biotechnol Bioeng 109:2612–2619
- Crow L, Wittenberg CL (1979) Separation and properties of NAD⁺- and NADP⁺-dependent glyceraldehyde-3-phosphate dehydrogenases from *Streptococcus mutans*. J Biol Chem 254:1134–1142
- Csonka LN, Fraenkel DG (1977) Pathways of NADPH formation in *Escherichia coli*. J Biol Chem 252:3382–3391
- D'Alessio G, Josse J (1971) Glyceraldehyde phosphate dehydrogenase of *Escherichia coli*. J Biol Chem 246:4326–4333
- Datsenko K, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97:6640–6645
- Dittrich CR, Vadali RV, Bennett GN, San K (2005) Redistribution of metabolic fluxes in the central aerobic metabolic pathway of *E. coli* mutant strains with deletion of the *ackA-pt* and *poxB* pathways for the synthesis of isoamyl acetate. Biotechnol Prog 21:627–631
- Fillinger S, Boschi-Muller S, Azza S et al (2000) Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. J Biol Chem 275:14031–14037
- Flores N, Flores S, Escalante A et al (2005) Adaptation for fast growth on glucose by differential expression of central carbon metabolism and gal regulon genes in an *Escherichia coli* strain lacking the phosphoenolpyruvate:carbohydrate phosphotransferase system. Metab Eng 7:70–87
- Foster JW, Park YK, Penfound T et al (1990) Regulation of NAD metabolism in *Salmonella typhimurium*: molecular sequence analysis of the bifunctional *nadR* regulator and the *nadA-pnuC* operon. J Bacteriol 172:4187–4196
- Fuhrer T, Sauer U (2009) Different biochemical mechanisms ensure network-wide balancing of reducing equivalents in microbial metabolism. J Bacteriol 191:2112–2121
- Habenicht A (1997) The non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase: biochemistry, structure, occurrence and evolution. Biol Chem 378:1413–1419
- Hillman JD, Fraenkel DG (1975) Glyceraldehyde 3-phosphate dehydrogenase mutants of *Escherichia coli*. J Bacteriol 122:1175–1179
- Hollywood N, Doelle H (1976) Effect of specific growth rate and glucose concentration on growth and glucose metabolism of *Escherichia coli* K-12. Microbios 17: 23–33
- Holm AK, Blank LM, Oldiges M et al (2010) Metabolic and transcriptional response to cofactor perturbations in *Escherichia coli*. J Biol Chem 285:17498–17506
- Iddar A, Valverde F, Serrano A, Soukri A (2002) Expression, purification, and characterization of recombinant non-phosphorylating NAP-dependent glyceraldehyde-3-phosphate dehydrogenase from *Clostridium acetobutylicum*. Protein Expr Purif 25:519–526
- Iddar A, Valverde F, Serrano A, Soukri A (2003) Purification of recombinant non-phosphorylating NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from *Streptococcus pyogenes* expressed in *E. coli*. Mol Cell Biochem 247:195–203
- Iddar A, Valverde F, Assobhei O et al (2005) Widespread occurrence of non-phosphorylating dehydrogenase among Gram-positive bacteria. Int Microbiol 8:251–258
- Irani MH, Maitra PK (1977) Properties of *Escherichia coli* mutants deficient in enzymes of glycolysis. J Bacteriol 132:398–410
- Jensen PR, Michelsen OLE (1992) Carbon and energy metabolism of *atp* mutants of *Escherichia coli*. J Bacteriol 174:7635–7641
- Kabir MM, Shimizu K (2003) Fermentation characteristics and protein expression patterns in a recombinant *Escherichia coli* mutant lacking phosphoglucose isomerase for poly(3-hydroxybutyrate) production. Appl Environ Microbiol 62:244–255
- Kalman LV, Gunsalus RP (1988) The *frdR* Gene of *Escherichia coli* globally regulates several operons involved in anaerobic growth in response to nitrate. J Bacteriol 170:623–629
- Kern A, Tilley E, Hunter IS et al (2007) Engineering primary metabolic pathways of industrial micro-organisms. J Biotechnol 129:6–29
- Kim S, Moon D-B, Lee C-H et al (2009) Comparison of the effects of NADH- and NADPH-perturbation stresses on the growth of *Escherichia coli*. Curr Microbiol 58:159–163
- Kim S, Lee CH, Nam SW, Kim P (2011) Alteration of reducing powers in an isogenic phosphoglucose isomerase (*pgi*)-disrupted *Escherichia coli* expressing NAD(P)-dependent malic enzymes and NADP-dependent glyceraldehyde 3-phosphate dehydrogenase. Lett Appl Microbiol 52: 433–440
- Kocharin K, Siewers V, Nielsen J (2013) Improved polyhydroxybutyrate production by *Saccharomyces cerevisiae* through the use of the phosphoketolase pathway. Biotechnol Bioeng 110:2216–2224
- Koebmann BJ, Westerhoff HV, Snoep JL et al (2002) The glycolytic flux in *Escherichia coli* is controlled by the demand for ATP. J Bacteriol 184:3909–3916
- Lee HC, Kim JS, Jang W, Kim SY (2010) High NADPH/NADP⁺ ratio improves thymidine production by a metabolically engineered *Escherichia coli* strain. J Biotechnol 149:24–32
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} method. Methods 25:402–408
- Marchal S, Branlant G (2002) Characterization of the amino acids involved in substrate specificity of nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase from *Streptococcus mutans*. J Biol Chem 277:39235–39242
- Martínez I, Zhu J, Lin H et al (2008) Replacing *Escherichia coli* NAD-dependent glyceraldehyde 3-phosphate dehydrogenase (GAPDH) with a NADP-dependent enzyme from *Clostridium acetobutylicum* facilitates NADPH dependent pathways. Metab Eng 10:352–359
- Moat AG, Foster JW, Spector MP (2002) Fermentation pathways. Microbial Physiology p 412–433
- Noda S, Takezawa Y, Mizutani T et al (2006) Alterations of cellular physiology in *Escherichia coli* in response to

- oxidative phosphorylation impaired by defective F1-ATPase. *J Bacteriol* 188:6869–6876
- Phue J-N, Lee SJ, Kaufman JB et al (2010) Acetate accumulation through alternative metabolic pathways in *ackA⁻pta⁻poxB⁻* triple mutant in *E. coli* B (BL21). *Biotechnol Lett* 32:1897–1903
- Sambrook J, Russell D (2001) *Molecular cloning a laboratory manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Sauer U, Canonaco F, Heri S et al (2004) The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. *J Biol Chem* 279:6613–6619
- Sigala JC, Flores S, Flores N et al (2009) Acetate metabolism in *Escherichia coli* strains lacking phosphoenolpyruvate: carbohydrate phosphotransferase system; evidence of carbon recycling strategies and futile cycles. *J Mol Microbiol Biotechnol* 16:224–235
- Takahashi H, Inada T, Postma P, Aiba H (1998) CRP down-regulates adenylate cyclase activity by reducing the level of phosphorylated IIA(Glc), the glucose-specific phosphotransferase protein, in *Escherichia coli*. *Mol Gen Genet* 259:317–326
- Takeno S, Murata R, Kobayashi R et al (2010) Engineering of *Corynebacterium glutamicum* with an NADPH-generating glycolytic pathway for L-lysine production. *Appl Environ Microbiol* 76:7154–7160
- Taylor S, Wakem M, Dijkman G et al (2010) A practical approach to RT-qPCR-publishing data that conform to the MIQE guidelines. *Methods* 50:S1–S5
- Thouvenot B, Charpentier B, Branlant C (2004) The strong efficiency of the *Escherichia coli* *gapA* P1 promoter depends on a complex combination of functional determinants. *Biochem J* 382:371–382
- Valverde F, Losada M, Serrano A (1999) Engineering a central metabolic pathway: glycolysis with no net phosphorylation in an *Escherichia coli* *gap* mutant complemented with a plant *gapN* gene. *FEBS Lett* 449:153–158
- Vemuri GN, Altman E, Sangurdekar DP et al (2006) Overflow metabolism in *Escherichia coli* during steady-state growth: transcriptional regulation and effect of the redox ratio. *Appl Environ Microbiol* 72:3653–3661
- Waegeman H, De Lausnay S, Beauprez J et al (2013) Increasing recombinant protein production in *Escherichia coli* K12 through metabolic engineering. *N Biotechnol* 30:255–261
- Yagi T, Matsuno-yagi A (2003) The proton-translocating NADH-quinone oxidoreductase in the respiratory chain: the secret unlocked. *Biochemistry* 42:2266–2274