

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Estación Experimental del Zaidín

UNIVERSIDAD DE GRANADA

**LA BASE DE DATOS DE REGULADORES TRANSCRIPCIONALES BAC-
TREGULATORS Y APLICACIONES BIOINFORMÁTICAS AL ANÁLISIS
DEL GENOMA DE PSEUDOMONAS**

TESIS DOCTORAL

Manuel Martínez Bueno

Granada, 2006

LA BASE DE DATOS DE REGULADORES TRANSCRIPCIONALES BACTREGU-
LATORS Y APLICACIONES BIOINFORMÁTICAS AL ANÁLISIS DEL GENOMA
DE PSEUDOMONAS

Memoria que presenta
Manuel Martínez Bueno
para aspirar la Título de Doctor

Fdo: Manuel Martínez Bueno

Vº Bº

Director

Fdo: Juan Luis Ramos Martín
Doctor en Biología
Profesor de Investigación del CSIC

Universidad de Granada

2006

Esta Tesis se ha realizado en la Unidad Estructural de Bioquímica y Biología Molecular y Celular de Plantas de la Estación Experimental del Zaidín (C.S.I.C.), Granada.

Queda todavía por hacer una gran parte del trabajo; aunque sólo podemos aportar una pequeña contribución, debe sin embargo, constituir una justificación suficiente para nuestra empresa.

Estrabón, I, 2, I

[...] pero nunca se le había permitido sentir la fuerza inmensa y la furia desatada, la furia que cesa cansada pero jamás apaciguada [...]. Del mismo modo que algunas personas pasan superficialmente por los años de su existencia hasta hundirse suavemente en una plácida tumba, ignorantes de la vida hasta el final de sus días, sin haberse visto nunca forzadas a comprobar toda la perfidia, la violencia y el terror que la vida puede albergar. En la tierra y en el mar hay gentes así de favorecidas..., o así de menospreciadas por el destino o por el mar.

“Tifón”
Joseph Conrad.

[..] espero que sin haber aburrido a mis lectores con detalladas explicaciones científicas, teorías y deducciones, únicamente añadiré que me he esforzado en relatar sólo la historia de la aventura misma. Habiéndolo hecho a mi modesta manera, amarro por el momento el barco, tomando bitadura con los cables de mal tiempo, y dejo al *Spray* bien seguro en puerto.

“Navegando en solitario alrededor del mundo”
Joshua Slocum.

Agradecimientos

En este momento de recapitulación que supone la conclusión de mi trabajo de Tesis Doctoral, vuelven a mi memoria los apoyos y trabas que finalmente la han conformado.

Si alguien ha hecho posible el que yo pueda plasmar estas reflexiones ha sido el Dr. Juan Luis Ramos Martín, que no sólo ha confiado en mí habiéndome permitido formar parte de su grupo, lo que ha constituido un orgullo que me acompañará el resto de mi vida, sino que si esta Tesis, al final ha visto la luz, ha sido gracias a su empeño que a menudo ha superado incluso al mío. Por ello mi más sincero agradecimiento.

En este apartado he de incluir a Antonio Jesús Molina, por su ayuda, en todo momento y por mostrarme esos recovecos de los programas que yo ni tan siquiera imaginaba. Gracias sobre todo por su actitud “siempre positiva” y el afrontar cualquier tarea con su sempiterno buen humor.

He de recordar Wilson Terán. Aunque en este momento se encuentre lejos, sólo en la distancia, no olvido nuestras discusiones que han constituido una rica fuente de reflexión científica, en el periodo que he pasado aquí. Esa visión global que posee, constituye una de las herramientas más poderosas del científico. Buena suerte.

Aunque la naturaleza de mi labor sea aparentemente ajena al clásico trabajo de laboratorio, los seminarios me han permitido conocer e intentar hacer propio el rigor científico que caracteriza la actividad del grupo. Además y no menos importante he disfrutado de la oportunidad de conocer a personas como Ana, Ana, Antonia, Antonio, Estrella, Eugenia, Isabel, Javier, Jesús, Manolo, Maria Jesús, Mari Trini, Maribel, M^a Mar, Nené, Patricia, Pieter, Silvia, Tino, Toñi, ...ha sido un placer.

En este apartado he de mencionar a Antonio Ramos, cuya contribución aunque no directamente científica, me ha proporcionado las condiciones propicias para finalizar esta Tesis.

1	Introducción.....	3
1.1.	Los factores sigma de la ARN polimerasa.....	6
1.2.	Reguladores transcripcionales.....	8
1.3.	Fases tempranas de la transcripción bacteriana.....	10
1.4.	Represión de la transcripción.....	13
2.	Objetivos.....	19
3.	Materiales y Métodos.....	23
4.	Capítulo 1.....	29
4.1	Resumen.....	29
4.2	Introducción.....	29
4.3	Métodos y sistemas de análisis.....	31
4.3.1	Selección de un conjunto no redundante de cada familia mediante análisis de las bases de datos SPTR y NCBI.....	31
4.3.2	La entrada en BacTregulators.....	31
4.4	Resultados.....	35
4.4.1	Información contenida y estructura de la base de datos BacTregulators.....	35
4.4.2	Acceso a la información en BacTregulators.....	36
4.5.	Discusión.....	37
4.5.1	Integración del conocimiento modificable en el paso de recuperación.....	37
5.	Anexo I (Organización de la base de datos BacTregulators).....	45
5.1.	Descripción de la Base de Datos de Proteínas.....	45
5.2	Descripción de la Base de Datos de Referencias Bibliográficas.....	57
5.3.	Descripción de la Base de Datos de Relaciones.....	60
5.4	Conexión entre las tres Bases de Datos de BacTregulators.....	61
6.	Capítulo 2.....	67
6.1.	Resumen.....	67
6.2.	Introducción.....	67
6.3.	Definiendo la familia TetR.....	77
6.3.1	El perfil de la familia TetR.....	77
6.3.2	Identificación de miembros de la familia TetR en las diferentes bases de datos de ADN y proteínas.....	80
6.4.	Proteínas de la familia TetR de estructura tridimensional conocida.....	85
6.4.1.	El regulador TetR.....	86
6.4.2.	El regulador QacR.....	91
6.4.3.	La estructura tridimensional de CprB.....	96
6.4.4.	La estructura tridimensional de EthR.....	98
6.4.5.	Estructura de los miembros de la familia TetR cristalizados y de funciones desconocidas.....	101
6.5.	Características estructurales de las proteínas de la familia TetR...	101
6.6.	Predicción de unión a ADN para las proteínas de la familia TetR basadas en el estudio de las estructuras de los cristales de TetR y QacR.....	102
6.6.1	Estableciendo relación entre las posiciones en el perfil y la posición estructural.....	102
6.6.2.	Predicción de los residuos de unión a ADN en los otros	

	miembros de la familia con estructura resuelta.....	105
6.6.3.	Predicción de los residuos de unión a ADN en los otros miembros de la familia sin estructura resuelta.....	107
6.7.	Algunos reguladores de la familia TetR forman parte de circuitos reguladores complejos.....	109
6.7.1.	El regulador AcrR es el regulador específico de la bomba de extrusión AcrAB.....	110
6.7.2.	El circuito Mtr de <i>Neisseria</i>	111
6.7.3.	BetI controla la ruta del osmorregulador colina-betaina en <i>Escherichia coli</i>	112
6.7.4.	El regulador ArpA de <i>Streptomyces</i>	113
6.7.5.	HapR regula los genes de virulencia en <i>Vibrio cholerae</i>	115
6.7.6.	Otros circuitos <i>quorum-sensing</i>	115
6.8.	Aplicaciones biotecnológicas y perspectivas futuras.....	118
	Anexo	
	II.....	123
7.	Capítulo 3.....	187
7.1.	Resumen.....	187
7.2.	Introducción.....	187
7.3.	Resultados y Discusión.....	189
7.3.1.	Identificación de los factores sigma del genoma de <i>Pseudomonas putida</i> KT2440.....	189
7.3.2.	Diecinueve factores sigma de <i>P. putida</i> KT2440 pertenecen a la familia ECF.....	193
7.3.3.	Análisis de los <i>loci</i> del genoma de <i>P. putida</i> KT2440 que codifican supuestos factores sigma-ECF con similitud de secuencia a <i>fecI</i> de <i>Escherichia coli</i>	194
7.3.4.	Comparación de factores sigma ECF de <i>P. putida</i> y <i>P. aeruginosa</i>	205
8.	Conclusiones.....	211
9.	Referencias.....	215

Relación de Tablas y Figuras

Relación de Tablas.

Tabla 1: Tabla de Proteínas.

Tabla 2: Código de función del NCBI.

Tabla 3: Tabla de Secuencias.

Tabla 4: Tabla de Elementos de Genoma.

Tabla 5: Tabla de Familias de Proteínas.

Tabla 6: Tabla de Tipos de Proteínas.

Tabla 7: Tabla de ActivadorRepresor.

Tabla 8: Tabla de Acciones.

Tabla 9: Tabla de Promotores.

Tabla 10: Tabla de Tipos de Acciones.

Tabla 11: Tabla de Sitios de unión al ADN.

Tabla 12: Tabla de unión ProteínaPromotor.

Tabla 13: Tabla de párrafos.

Tabla 14: Tabla de Referencias.

Tabla 15: Tabla de Evidencias.

Tabla 16: Tabla de Tipos de Evidencias.

Tabla 17: Tabla de unión PárrafoEvidencia.

Tabla 18: Tabla de unión PárrafoReferencia.

Tabla 19: Tabla de unión MantenedorPárrafo.

Tabla 20: Tabla de Relaciones entre Objetos.

Tabla 21: Tabla de Tipos de Relaciones.

Tabla 22: Tabla de Tipos de Objetos.

Tabla 23: Tabla de Condiciones.

Tabla 24: Tabla de unión RelaciónReferencia

Tabla 25: Tabla de unión RelaciónEvidencia.

Tabla 26: Tabla de Campos.

Tabla 27: Tabla de Tablas.

Tabla 28: Familias de reguladores de la transcripción en procariontes.

Tabla 29: Funciones específicas reguladas por los miembros de la familia TetR de represores.

Tabla 30: Distribución de proteínas TetR en procariontes.

Tabla 31: Frecuencia de aminoácidos en cada una de las posiciones críticas para el reconocimiento del operador en los miembros de la familia TetR.

Tabla 32: “Palabras de unión a ADN” de los miembros de la familia TetR.

Tabla 33: Factores sigma codificados en el genoma de *P. putida* KT2440.

Tabla 34: Factores sigma codificados en el genoma de *P. aeruginosa* PA01.

Tabla 35: Factores sigma en genomas los bacterianos anotados.

Tabla 36: Valores de similitud entre los factores sigma ECF de *P. putida* y la proteína FecI de *E. coli*.

Tabla 37: Receptores asociados con los agrupamientos de proteínas similares a FecIR en *P. putida*.

Tabla 38: Comparación de los factores sigma ECF de *P. putida* KT2440 y *P. aeruginosa* PA01.

Tabla 39: Comparación de los sensores transmembrana (TMS) asociados con los factores sigma ECF de *P. putida* KT2440 y *P. aeruginosa* PA01.

Tabla 40: Comparación de los receptores de hierro (*rec*) asociados a los factores ECF de *P. putida* KT2440 y *P. aeruginosa* PA01.

Relación de figuras.

Figura 1: Integración de las secuencias de proteínas de SPTR con las procedentes de los genomas disponibles en el NCBI en un conjunto no redundante para cada familia.

Figura 2: Definición de una entrada de BacTregulators.

Figura 3: Dos secuencias idénticas pueden pertenecer a entradas diferentes dependiendo de las diferencias en cualquiera de los cuatro niveles que definen a la entrada BacTregulators.

Figura 4: Las secciones de una entrada en BacTregulators.

Figura 5: Lista de entradas relacionadas con la entrada de ID BacTregulators.

Figura 6: Informe de los datos de conocimiento de las entradas con secuencia similar a la secuencia de la entrada 1 ($E < 10^{-36}$) en Gammaproteobacterias.

Figura 7: Estructura de la base de datos BacTregulators.

Figura 8: Alineamiento de 42 miembros de la familia TetR que empleado en la construcción del perfil de la familia TetR.

Figura 9: Diagrama de cintas del homodímero de TetR.

Figura 10: Superposición de los dominios de unión a ADN de los reguladores TetR, QacR, EthR y CprB

Figura 11: Unión de TetR a su operador.

Figura 12: Representación de la cavidad de TetR implicada en la unión a tetraciclina.

Figura 13: Superposición del monómero de TetR con tetraciclina.

Figura 14: Diagrama de cintas de los dos monómeros QacR unidos a su secuencia de ADN diana.

Figura 15: Unión de QacR a su operador.

Figura 16: Comparación de los contactos del par de residuos Lys39/Tyr40 con las bases nitrogenadas entre los monómeros distal y proximal.

Figura 17: Diagrama de cintas de los homodímero de otros miembros de la familia TetR cristalizados.

Figura 18: Diagrama de cintas de los monómeros de los miembros de la familia TetR cristalizados.

Figura 19: Superposición de los dominios de unión al operador de TetR y QacR mostrando los residuos de interacción con el ADN.

Figura 20: Superposición de los dominios de unión al operador de TetR y QacR mostrando los residuos conservados en las posiciones del perfil no implicadas en contactos con el ADN.

Figura 21. Dominios de unión a ADN de CprB, EthR, Cgl2612, YbiH, YcdC e YsiA

Figura 22: Ejemplos de redes de regulación que implican a miembros de la familia TetR.

Figura 23: Dendrograma de los 24 factores sigma detectados en el genoma de *Pseudomonas putida* KT2440.

Figura 24: Genes similares a *fecI* y sus regiones adyacentes en *P. putida* KT2440.

Figura 25: Alineamiento múltiple de los factores sigma ECF similares a *FecI* de *P. putida* KT2440 y *FecI* de *E. coli*.

Figura 26: Alineamiento múltiple de los 60 primeros residuos de las 11 proteínas similares a *FecR* ligadas a los factores sigma similares a *FecI* en *P. putida* KT2440.

Figura 27: Representación esquemática del alineamiento múltiple de la región N-terminal de los receptores similares a *FecR* de *P. putida* KT2440 ligados a grupos *fecIR* con un conjunto de receptores sin extensión N-terminal responsable de la interacción con *FecR*.

Figura 28: Genes de *Pseudomonas aeruginosa* que codifican factores sigma ECF y sus regiones adyacentes.

Introducción

La capacidad de un microorganismo para colonizar un nicho es el resultado de la interacción dinámica entre su repertorio genético y su aparato transcripcional (reguladores y ARN polimerasa) en respuesta a las condiciones ambientales del nicho (Babu and Teichmann 2003). Las respuestas más frecuentes de los organismos frente a las variaciones ambientales consisten en la activación y/o represión de genes para adaptar la fisiología celular o el metabolismo a las nuevas condiciones. Como consecuencia, las bacterias exhiben una amplia variedad de mecanismos reguladores de la expresión genética que pueden afectar a todos y cada uno de los genes de un microorganismo tal y como están poniendo de manifiesto los ensayos con “arrays” genómicos con múltiples microorganismos sometidos a distintos tipos de estreses ambientales (Rojo, 1999). Así, los genes se pueden "encender" o "apagar" mediante modificaciones en la ARN polimerasa, reordenaciones que conecten o desconecten a un gen de un determinado promotor, mediante la acción de proteínas reguladoras que activen o inhiban el inicio de la transcripción, o mediante la modulación de las fases de elongación y terminación en sitios específicos. Además, la estabilidad de los ARN mensajeros, la eficiencia de la traducción, el control de la actividad enzimática y la degradación de proteínas son también dianas de los mecanismos de regulación (Rojo, 1999). Una simplificación útil de esta cuestión es asumir que la transcripción constituye la componente principal en el control de la expresión génica y considerar que la expresión genética es la respuesta evolutiva al reto de sobrevivir en un medio ambiente cambiante (Cases y de Lorenzo, 2005).

El control de la transcripción requiere de la interacción de las vías sensoras en la señalización y de los reguladores que sensan y modulan la maquinaria molecular de la transcripción: la ARN polimerasa, el ácido nucleico a transcribir y las proteínas reguladoras que actúan sobre ella. La mayor parte de los conocimientos al respecto hasta la fecha se han centrado sobre el modo en que las interacciones proteína-proteína y proteína-ADN controlan la tasa de transcripción genética. Las técnicas de genética molecular bacteriana, en particular aplicadas a *Escherichia coli* y *Bacillus subtilis*, permitieron un progreso rápido en el esclarecimiento de los tipos de estrategias empleados en el control de la expresión genética a nivel transcripcional (Wigneshweraraj et al., 2005).

1.1. Los factores sigma de la ARN polimerasa.

El núcleo de la ARN polimerasa bacteriana consta de 5 subunidades ($\alpha_2\beta\beta'\omega$). Los componentes catalíticos de la ARN polimerasa residen en las subunidades $\beta\beta'$ que se ensamblan sobre el dímero α_2 . La subunidad ω posee actividad tipo chaperona, e interviene en el ensamblaje de la ARN polimerasa. La asociación de una sexta subuni-

dad, la subunidad factor σ , permite al complejo unirse a las secuencias específicas promotoras en el ADN y dirigir la transcripción de forma específica para cada promotor.

Por razones históricas la nomenclatura de los factores σ es compleja. En *E. coli* y otras muchas bacterias Gram-negativas, los genes de los factores σ se designan por *rpo* ("RNA polymerase subunit"), mientras que en la mayoría de las bacterias Gram-positivas los genes correspondientes se designan *sig*. Las proteínas se pueden designar por la letra griega σ con un superíndice que refleja o su peso molecular o el nombre del gen que lo codifica, o ser una letra designada arbitrariamente. Actualmente, con la disponibilidad creciente de genomas secuenciados, la situación se complica aún más al comprobar que algunos microorganismos contienen más tipos de factores σ que letras el alfabeto.

El número de tipos de factores σ varía de unas especies a otras siendo un reflejo del modo de vida de la bacteria tal como era de esperar, oscilando entre un único factor σ en *Mycoplasma genitalium* y los 63 en *Streptomyces coelicolor*, que es hasta la fecha el microorganismo con un mayor número de factores σ (Wigneshweraraj *et al.*, 2005).

Al complejo σ -ARN polimerasa se le denomina holoenzima. El holoenzima "prototipo" contiene un factor σ perteneciente a la clase σ^{70} . En *E. coli*, 6 de sus 7 factores sigma pertenecen a esta clase σ^{70} . Una forma alternativa de holoenzima contiene el factor σ^{54} , que no estando relacionado con σ^{70} ni por secuencia primaria ni por modo de acción, es el único representante de su clase (Merrick, 1993).

Los miembros de la familia σ^{70} se han dividido en cuatro grupos filogenéticos. El grupo 1 incluye a los factores σ esenciales, todos ellos estrechamente relacionados con σ^{70} de *E. coli*. El grupo 2 contiene proteínas también muy relacionadas con los factores σ primarios, pero que son dispensables para el crecimiento celular, tales como el producto del gen *rpoS* de *E. coli* y SigA de *B. subtilis*. Los factores σ del grupo 3 poseen una relación más distante con el σ^{70} tipo y suelen activar regulones en respuesta a señales específicas tales como control dependiente de desarrollo y choque térmico. Más en concreto, los miembros del grupo 3 se pueden subdividir en varios grupos implicados en: esporulación (SigE, SigF y SigG de *B. subtilis*), biosíntesis flagelar (FliA de *E. coli* o SigD de *B. subtilis*), respuesta a choque térmico (el producto del gen *rpoH* o σ^{32} de *E. coli*). Por último el grupo 4 incluye a la subfamilia numéricamente mayor pero a su vez la más divergente, la subfamilia de factores σ de función extracitoplásmica (ECF), cuyos miembros responden a señales procedentes del medio extracitoplásmico, como por ejemplo FecI y RpoE de *E. coli*.

Mientras que la función del factor σ^{70} es dirigir la expresión de la mayor parte de los genes durante la fase exponencial, los factores σ accesorios de los grupos 2 al 4,

suelen funcionar para conectar la expresión de conjuntos específicos de genes en respuesta a señales adecuadas. Sus funciones se pueden dividir en 3 grandes categorías: respuestas a estreses, desarrollo y metabolismo secundario. La amplia variedad de respuestas a estreses controlada por miembros de la familia σ^{70} incluye la fase estacionaria y la respuesta a estrés general (mediada por ejemplo por σ^S en *E. coli* y σ^B en *Bacillus subtilis*), plegamiento defectuoso de proteínas en el medio intracelular y extracitoplásmico (regulado en *E. coli* por σ^{32} y σ^E respectivamente), estrés oxidativo (por ejemplo σ^R en *Streptomyces coelicolor*), estrés osmótico (σ^M en *B. subtilis*) y estrés de la pared celular (controlado, por ejemplo por σ^E en *S. coelicolor* y σ^W en *B. subtilis*). Los programas de desarrollo bajo el control de miembros de la familia σ^{70} incluyen la biosíntesis de flagelos (que implica a σ^D en *B. subtilis* y σ^F en *Salmonella typhimurium*), formación de endosporas (mediada, por ejemplo, σ^E , σ^K , σ^F y σ^G en *B. subtilis*) y formación de exosporas (σ^{WhiG} en *S. coelicolor*). Las funciones de metabolismo auxiliar controladas por factores de la familia σ^{70} incluye por ejemplo la captación de hierro (σ^{FecI} en *E. coli* y σ^{PvdS} en *P. aeruginosa*) (Paget y Helmann, 2003).

El producto del gen *rpoN* (anteriormente conocido como *ntrA*) descrito durante el análisis de los genes implicados en el metabolismo nitrogenado en bacterias entéricas, codifica una subunidad σ pero de unas propiedades bastante diferentes a los factores σ previamente descritos (Merrick, 1993; Buck *et al.*, 2000). Esta subunidad en principio se denominó σ^N (por su implicación en el metabolismo nitrogenado) y posteriormente σ^{54} . En *E. coli*, σ^{54} representa casi el 8% del total de las moléculas de factores σ presentes. Generalmente sólo existe una copia por genoma, pero ocasionalmente puede haber dos genes altamente homólogos que codifican σ^{54} . Este factor σ no suele ser esencial para la supervivencia en condiciones favorables, excepto en *Myxococcus xanthus*. Sin embargo, el uso y distribución de σ^{54} a lo largo de bacterias filogenéticamente diversas subraya su significación biológica, sirva como ejemplo su presencia en la bacteria gram positiva *B. subtilis* y el termófilo extremo *Aquifex aeolicus*. Sin embargo está ausente en el gram positivo rico en C+G *Mycobacterium tuberculosis*, el extremófilo *Thermotoga maritima*, la cianobacteria *Synechocystis* sp. y los patógenos especializados *Rickettsia prowazekii* y *Mycoplasma genitalium* (Wigneshweraraj *et al.*, 2005).

Aunque la mayoría de los factores sigma alternativos participan en la transcripción de conjuntos de genes más o menos relacionados, no ocurre lo mismo con σ^{54} que aunque tal como hemos dicho se identificó en principio como un factor σ específico del nitrógeno ahora se conoce que participa en la expresión de una gran variedad de genes implicados en funciones diferentes que se transcriben desde promotores dependientes de σ^{54} . Entre estas funciones diversas hay genes implicados en el transporte y metabolismo de diferentes fuentes de carbono y nitrógeno, en patogenicidad en vegetales, en síntesis de diversos cofactores, producción de alginato, y síntesis de flagelos (Cases *et al.*, 2003;

Arora *et al.*, 1997), además de intervenir en el establecimiento de diazotrofia por bacterias fijadoras de nitrógeno (Wigneshweraraj *et al.*, 2005). Aunque para el caso de *E. coli* se propuso una hipótesis que unificaba la aparente disparidad de genes regulados por σ^{54} basada en las condiciones fisiológicas derivadas de la privación de nitrógeno, dicha hipótesis es sólo aplicable a *E. coli* y no explica otras muchas funciones en especies diferentes como por ejemplo su implicación en el catabolismo de compuestos aromáticos en *P. putida* (Cases *et al.*, 2003b).

La asociación de diferentes tipos de factores σ al núcleo ARN polimerasa, cada uno de los cuales dirige al holoenzima a un conjunto de promotores específicos, es una estrategia que en sí misma constituye el primer nivel en la jerarquía del control de la expresión genética. Así por ejemplo, aunque el número total de genes en *E. coli* es de unos 4000, el de núcleos ARN polimerasa por célula es de 2000, de los que al menos un tercio no están ocupados en la transcripción y constituyen una reserva que se puede emplear como respuesta a cambios ambientales. Así los núcleos de ARN polimerasa libre pueden dirigirse hacia la transcripción de conjuntos de genes específicos (regulones) mediante su asociación a una variedad de factores σ alternativos, los cuales a su vez se expresan bajo condiciones específicas.

Además, la transcripción de la mayoría de los genes bacterianos requiere de la participación de uno o más factores accesorios adicionales, llamados genéricamente factores de transcripción. Así entre un grupo de genes bajo el control de un mismo tipo de factor σ , la transcripción desde cada promotor está auxiliada o inhibida por los factores de transcripción, que constituirían un segundo nivel en la jerarquía del control de la expresión génica (Ishihama, 2000; Vicente *et al.*, 1999).

1.2. Reguladores transcripcionales.

Los reguladores de la transcripción suelen ser proteínas de unión a ADN que presentan diferentes tipos de organización de dominios. La mayoría de los reguladores microbianos implicados en la transcripción son proteínas con dos dominios en un solo polipéptido, un dominio sensor de señal y un dominio de unión a ADN que transduce la señal (Abo-Amer *et al.*, 2004; Babu and Teichmann, 2003; Henikoff *et al.*, 1988; Hitomi *et al.*, 2003; Karniol and Vierstra, 2004; Lauble *et al.*, 1989; Nixon *et al.*, 1986; Pabo and Lewis, 1992; Pabo y Sauer, 1984; Pabo y Sauer, 1992; Pao *et al.*, 1994; Perron *et al.*, 2004; Schleif, 1987; Soisson *et al.*, 1997; Wintjens y Rooman, 1996; Wösten *et al.*, 2004).

En otros casos, la componente sensora y la de unión a ADN, no están en la misma proteína sino que implica a dos proteínas como en los sistemas reguladores de 2 componentes tales como CzcR/CzcS; DcuS/DcuR; NifL/NifA; NtrB/NtrC; PhoP/PhoQ; TodS/TodT, etc, (Derzelle *et al.*, 2004; Hayes *et al.*, 1997; Kustu *et al.*, 1991; Lau *et al.*, 1997; Martínez-Argudo *et al.*, 2002; Martínez-Argudo y Contreras, 2002; Morett y Segovia, 1993; Poole *et al.*, 1993; Pragman *et al.*, 2004; Redenbach *et al.*, 1998; Ventre *et al.*, 2004; Westbrook-Wadman *et al.*, 1999), en los que una de las proteínas suele ser una quinasa ligada a membrana que detecta la señal apropiada y que fosforila a una proteína de unión a ADN que media la transcripción desde el promotor correspondiente.

Los sistemas de un solo componente, en los que el dominio regulador de la transcripción y el dominio sensor están unidos en una sola proteína, como hemos dicho, son con diferencia los más abundantes en procariotas. En sus formas más simples combinan un dominio de unión a ADN con un dominio de unión de moléculas pequeñas, por ejemplo efector (dominios SMBD). Las organizaciones más complejas pueden implicar múltiples SMBDs con diferentes aportaciones alostéricas, o incluso dominios adicionales tales como el AAA+ATPasa (en los activadores ligados a σ^{54}). Los SMBDs se organizan a partir de un número relativamente pequeño de plegamientos ancestrales, los cuales se describen a continuación.

a. El plegamiento tipo PAS: que incluye al dominio PAS (une ligandos tales como hemo, flavín nucleótidos, cinamato), al dominio GAF (une NMPc, tetrapirroles y formato) y los dominios de unión a ligando tipo IclR (versiones divergentes del dominio PAS).

b. Dominios de proteínas periplásmicas tipo I y II que incluyen a los dominios de unión a ligando de la familia LysR.

c. Plegamiento tipo ferredoxina: incluye al dominio ACT y los dominios de unión a ligando de los factores de transcripción tipo Lrp y los de ferredoxinas clásicas unidos a dominios de unión de ADN de archeas y cianobacterias.

d. El plegamiento doble hebra β -hélice (CUPIN): presente en los dominios de unión a ligando tipo AraC, además del dominio de unión de cNMP (presente en la familia CAP).

e. El dominio GyrI que contiene dos copias del módulo estructural SHS2 ($\beta\alpha\beta$) uno de los principales dominios de unión a ligando de la familia MerR.

Algunos otros SMBDs comparten un plegamiento común con los dominios propios de enzimas pero carecen de actividad catalítica y simplemente actúan como sitios

de reconocimiento de efectores. En este grupo se pueden diferenciar varios subgrupos, por ejemplo:

- i. El dominio UTRA del grupo HutC/FarR de la familia GntR de factores de transcripcionales, que poseen el mismo plegamiento que la corismato liasa.
- ii. El dominio DeoR que posee el plegamiento α/β presente en la familia de isomerasas de monosacáridos fosfatados, (por ejemplo la ribosa fosfato isomerasa).
- iii. El dominio vinil-4 reductasa (V4R), capaz de unir efectores tipo hidrocarburo presente en los activadores XylR y DmpR.

Otro tipo de combinación que encontramos en los factores de transcripción de un solo componente es la fusión del dominio de unión a ADN con dominios enzimáticos dotados de actividad catalítica. Como ejemplo tenemos a la proteína LexA (dotada de un dominio de la familia peptidasa S24), el represor de los genes implicados en respuesta a daño del ADN (regulón SOS), que en presencia de ADN monocatenario y RecA, cataliza su rotura autocatalítica que elimina su capacidad de unión al ADN, permitiendo la transcripción del regulón SOS. Otros ejemplos son el represor del operón de la biotina, BirA que contiene un dominio biotín ligasa C terminal, que en presencia de biotina sintetiza biotinil-5'-adenilato que unido a BirA mantiene reprimida la transcripción de los genes de la biosíntesis de la biotina. Otros ejemplos notables son el dominio aminotransferasa dependiente de pirodoxal fosfato en las proteínas reguladoras de la familia GntR o el dominio treonín sintetasa restringido al género *Pyrococcus*.

Respecto de los dominios de unión a ADN los análisis estructurales revelan que casi el 95% de los factores transcripcionales descritos en procariotas usan el dominio hélice-giro-hélice en sus diversas versiones (Argos *et al.*, 1983; Bairoch y Rudd, 1995; Bateman *et al.*, 2000; Branden y Tooze, 1991; Brennan *et al.*; 1990; Fromknecht *et al.*, 2003; Harrison y Aggarwal, 1990; Harrison, 1991; Pérez-Rueda y Collado-Vides, 2001; Rousseau *et al.*, 2004; Schevitz *et al.*, 1985).

Destacamos por un lado a aquellos constituidos solo por hélices, desde el tipo básico de tres hélices (familia Fis, NtcR), al de cuatro hélices (AraC, TetR, LuxR, TrpR, DnaA), hasta versiones multihélices, muy infrecuentes (el regulador Spo0A de bacterias Gram positivas formadoras de esporas). Por otro lado tenemos a los dominios hélice-giro-hélice suplementados con una horquilla con dos hebras β en su región C-terminal, desde la versión básica, con 2 hebras β , (Lrp/AsnC, Arg, DeoR, LexA, MCM) o la que posee además una hélice C-terminal extra tras la horquilla (Fur, Mar, GntR, LysR), pasando por las versiones con 3 hebras β (ArsR), con 4 hebras (plegamiento tipo CAP, de la familia de mismo nombre).

Por último citar variantes muy modificadas de dominios hélice-giro-hélice como el de la familia MetJ-Arc en el que la primera hélice se sustituye por una hebra β , o el de la familia MerR igual al anterior respecto a la primera hélice, pero dotada además de la horquilla con 2 hebras β características y una hélice C-terminal que suple estructuralmente la falta de la primera (Aravind *et al.*, 2005).

Los activadores AAA+ son proteínas modulares generalmente constituidas por 3 dominios: el N terminal o dominio regulador, el dominio central o AAA+, catalítico, y el dominio C terminal o de unión a ADN. En algunos reguladores puede faltar el dominio N o el C terminal. Según su composición de dominios y principalmente su mecanismo regulador, estos activadores se clasifican en 5 grupos, del I al V.

El dominio de unión a ADN es de tipo hélice-giro-hélice y confiere la especificidad de unión al *enhacer*. Como las secuencias *enhacer* suelen tener una simetría binaria, la mayoría de los AAA+s se une al ADN inicialmente como dímeros.

Los activadores del holoenzima con σ^{54} comparten el tipo de dominio catalítico, muy conservado, que da nombre a la familia, el dominio AAA+ que porta la actividad ATPásica y contacta directamente con σ^{54} . La característica estructural unificadora de los activadores AAA+ es que tal como sugiere la estructura resuelta de PspF de *E. coli*, son dímeros en su estado "básico" pero que generalmente se deben ensamblar 3 dímeros para formar un oligómero hexamérico, de aspecto anillado, necesario para la actividad. Este requerimiento probablemente se deba a la organización del sitio catalítico ATPásico que se forma en la interfase entre protómeros adyacente. Este hecho justificaría la unión del activador a varias secuencias de ADN *enhacer* adyacentes (Morett y Segovia, 1993; Wigneshweraraj *et al.*, 2005).

Dependiendo del dominio regulador, N terminal, podemos encontrar varias vías de regulación en estas proteínas.

a) Regulación por fosforilación. La mayoría de los miembros del grupo I de activadores AAA+ pertenecen a los sistemas reguladores de dos componentes y poseen un dominio receptor que se fosforila por la correspondiente proteína quinasa sensora que capta la señal ambiental específica. El dominio receptor controla el rendimiento del dominio AAA+ ya sea positivamente como en el caso de NtrC (proteína C reguladora de nitrógeno), o negativamente como en el caso de DctD (proteína D de transporte de ácidos dicarboxílicos). Parece que el dominio receptor controla la actividad del activador mediante la modulación del estado de oligomerización del dominio AAA+. Por ejemplo, se ha demostrado que en NtrC la fosforilación de su dominio receptor estimula la oligomerización del dominio AAA+. Además, la estructura de NtrC1 de *A. aeolicus* muestra que se requieren reordenamientos mayores de los dominios receptores para

disparar la oligomerización de los dominios AAA+, lo que subsecuentemente configura el sitio activo para la hidrólisis del ATP. Otros activadores AAA+ del grupo I tales como LevR de *B. subtilis* portan un dominio de regulación fosfotransferasa (PRD) que se regula mediante fosforilación / defosforilación de residuos conservados de histidina.

b) Regulación por moléculas efectoras. Los activadores AAA+ del grupo II se regulan por la unión directa de un efector al dominio regulador que es un SMBD. Hay dos tipos de dominios reguladores de grupo II: el dominio PAS, por ejemplo la proteína RocR (proteína reguladora de los operones *roc*) de *B. subtilis* y el dominio V4R del que los ejemplos mejor estudiados son XylR (Xylene catabolism regulatory protein) y DmpR (Dimethylphenol catabolism regulatory protein) de *Pseudomonas putida*. El dominio V4R regula directamente la actividad ATPasa del dominio AAA+, pues proteínas mutantes XylR y DmpR sin dominio V4R mostraban actividad ATPasa constitutiva y eran activos para la activación de la transcripción (Wigneshweraraj *et al.*, 2005).

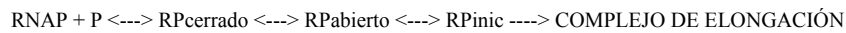
c) Regulación mediante interacciones proteína-proteína. El grupo III de activadores se regulan mediante interacciones directas proteína-proteína, vía dominio GAF. Como paradigma de este grupo podemos tomar a NifA que regula la fijación de nitrógeno y que se ha estudiado bien en numerosos grupos de gram negativos. En respuesta a señales ambientales, la actividad de NifA se regula negativamente en algunas bacterias por interacción directa con NifL.

El grupo IV de activadores AAA+ son una excepción pues carecen del dominio regulador N terminal. Los miembros del grupo IV como los del III se regulan mediante interacciones directas proteína-proteína. El miembro mejor caracterizado del grupo IV es PspF (phage shock protein E) de *E. coli*. PspF activa la transcripción de los genes *psp* que se relacionan con la adaptación celular a cambios en la integridad de la membrana. PspF es activo constitutivamente y se regula negativamente por la interacción con PspA. PspA actúa directamente con PspF inhibiendo la actividad ATPasa de éste último (Wigneshweraraj *et al.*, 2005).

1.3. Fases tempranas en la transcripción bacteriana.

El inicio de la transcripción es un proceso intrincado que implica varios pasos. Después de la unión de la ARN polimerasa al promotor, el complejo inicial que se ha formado sufre una serie de cambios antes de que la polimerasa pueda abandonar el promotor como un complejo de elongación que se desliza sobre el ADN transcribiéndolo. Brevemente, la ARN polimerasa inicialmente se une al promotor (P) como un "complejo cerrado" (RPC). A continuación, la separación de las cadenas del ADN en el centro activo del enzima, conduce a la formación de un "complejo abierto" (PRO) el cual, en

presencia de los cuatro nucleósidos trifosfatados, continúa hacia un "complejo inicial" (RP_{inic}). El complejo inicial puede verse temporalmente afectado por un proceso iterativo transcripcionalmente ineficaz o silencioso, durante el que genera y libera cadenas cortas de ARN. Este ciclo abortivo finaliza cuando la ARN polimerasa logra romper sus contactos con el promotor, liberando al factor sigma y abandonando al promotor progresando sobre el ADN como un complejo de elongación productivo. El proceso general puede esquematizarse como:



La eficiencia de la transición de un complejo al siguiente es diferente para los distintos promotores. Así, la fuerza de un promotor recae sobre la eficiencia combinada de cada uno de los pasos descritos, de modo que la menor eficiencia de uno de ellos será un paso limitante y actuará como un cuello de botella en el proceso. Como consecuencia el inicio de la transcripción puede modularse mediante reguladores que pueden afectar a cada una de las etapas de esta transición. Muchos activadores transcripcionales actúan acelerando uno o varios pasos limitantes, con más frecuencia en el paso inicial de unión de la ARN polimerasa al promotor o en la transición de complejo cerrado a complejo abierto. También se ha considerado que los represores actúan limitando el acceso de la ARN polimerasa al promotor (inhibición de la formación del complejo cerrado) y muchos represores usan de hecho esta vía, sin embargo existen represores que permiten la unión simultánea de ARN polimerasa al promotor aunque de un modo en el que no se alcanza la fase de elongación, tal como trataremos más adelante (Rojo, 1999).

Aunque σ^{54} se une al mismo tipo de núcleo ARN polimerasa que los miembros de la familia σ^{70} , cada clase de factor σ confiere propiedades diferentes al holoenzima. El holoenzima con σ^{70} transcribe desde promotores con motivos de reconocimiento conservados centrados en las posiciones -35 y -10 respectivamente, delante del sitio de inicio de la transcripción, en la posición +1; mientras que la forma alternativa del holoenzima que contiene a σ^{54} dirige la transcripción desde promotores con motivos conservados de reconocimiento en las posiciones -24 y -12 respecto del punto de inicio de la transcripción en posición +1.

El holoenzima σ^{70} -ARN polimerasa inicia la transcripción en ausencia de activadores y en caso de que participen, la activación de la transcripción suele ocurrir mediante interacciones entre el holoenzima y la proteína activadora, unida ésta última más o menos cerca de la zona del -35 del promotor. El ejemplo clásico y mejor conocido lo constituye CAP ("Catabolite Activator Protein") de *E. coli*. En la activación de la transcripción mediante CAP desde promotores de clase I, el sitio de unión de CAP (centrado

en la posición -61.5) está situado delante del sitio de la ARN polimerasa y la activación implica una interacción simple proteína-proteína entre CAP y σ^{70} -ARN polimerasa que conlleva el aumento de la tasa de formación de complejos cerrados sobre el promotor, es decir aumenta la afinidad por el promotor, lo que se traduce en un aumento de la tasa de transcripción basal. En la activación de la transcripción mediante CAP desde promotores de clase II, el sitio de unión de CAP sobre el ADN (centrado en posición -41.5) solapa con el elemento -35 del promotor y la activación de la transcripción implica tres tipos de interacciones proteína-proteína entre CAP y σ^{70} -ARN polimerasa que además de incrementar la afinidad de σ^{70} -ARN polimerasa por el promotor, incrementa también la tasa de isomerización de complejo cerrado a complejo abierto (Lawson *et al.*, 2004).

Por el contrario, el holoenzima σ^{54} -ARN polimerasa, en ausencia del correspondiente activador transcripcional, permanece en el promotor como complejo cerrado. La isomerización del complejo cerrado a un complejo abierto es estrictamente dependiente de una amplia clase de proteínas activadoras, dependientes de hidrólisis de ATP. Los activadores del holoenzima σ^{54} -ARN polimerasa no se unen al ADN cerca del promotor, como lo suelen hacer los activadores que afectan a σ^{70} -ARN polimerasa, sino que se unen a secuencias llamadas *enhancers* localizadas a unos 100 o 150 pares de bases del +1. A estas proteínas se les denomina proteínas de unión a *enhancers* (EBP). Estos activadores no actúan incrementando los niveles de transcripción basal, sino que son totalmente necesarios para la formación del complejo abierto y por tanto para que la transcripción tenga lugar.

Las proteínas de deformación del ADN, por ejemplo IHF o ArgR (represor de la Arginina en *E. coli*), participan en la transcripción mediada por σ^{54} permitiendo la formación de lazos en el ADN que permiten la aproximación y el contacto del activador unido al *enhancer* con σ^{54} -ARN polimerasa unida al promotor como complejo cerrado, separados por unos 100 a 150 pares de bases (Pérez-Martín y de Lorenzo, 1997). Una variación notable la encontramos precisamente en *Pseudomonas putida* en la que la transcripción desde el promotor Pu es dependiente del reclutamiento de σ^{54} -ARN polimerasa mediante la proteína IHF, pero la formación del complejo abierto es también totalmente dependiente del activador AAA+ XylR (Bertoni *et al.*, 1998). En algunos promotores, por ejemplo el del gen *nifL* en *Klebsiella pneumoniae*, las proteínas de deformación del ADN no son necesarios y es la curvatura intrínseca del ADN la que facilita la interacción entre el activador unido al *enhancer* y el complejo cerrado. El contacto del complejo cerrado con el activador resulta en una reordenación de las interacciones proteína-proteína y proteína-ácido nucleico, que conduce mediante la conversión de la energía química derivada de la hidrólisis del ATP a la remodelación del complejo cerrado capacitándolo para adoptar la conformación abierta, con la separación estable de las

cadena de ADN en las regiones detrás del motivo de reconocimiento del promotor -12 (Pérez-Martín y de Lorenzo, 1997; Wigneshweraraj *et al.*, 2005).

Además de los mecanismos descritos arriba se ha identificado una posible tercera vía de transcripción bacteriana, la que usa subunidades de tipo σ^{70} pero depende de secuencias *enhacer* y activadores que necesitan nucleótidos para la activación de la transcripción (Richard *et al.*, 2003) Tal es el caso de la proteína activadora de unión a *enhacer* NtcR de *Rhodobacter capsulatus*, que activa la transcripción de σ^{70} -ARN polimerasa en lugar de σ^{54} -ARN polimerasa.

Además los activadores que interaccionan con σ^{70} -ARN polimerasa lo hacen con las subunidades α y/o σ^{70} , mientras que parece ser que los activadores que contactan con σ^{54} -ARN polimerasa lo hacen a través de las subunidades β y σ^{54} . Del estudio de las interacciones entre σ^{70} -ARN polimerasa y NtcR en *Rhodobacter capsulatus* se desprende que la subunidad β' es necesaria y suficiente para la activación con NtcR, lo que en efecto estaría apoyando la tesis de una tercera vía transcripcional (Richard *et al.*, 2003).

Por tanto, parece que en bacterias la maquinaria de transcripción del núcleo ARN polimerasa actúa al menos en 3 formas diferentes sujetas a mecanismos reguladores diferentes (Wigneshweraraj *et al.*, 2005).

1.4. Represión de la transcripción.

Respecto de la represión, tal como anticipábamos, cualquier factor que inhiba el acceso de la ARN polimerasa al promotor puede considerarse como represor. Esta definición incluye no sólo a los represores de la transcripción clásicos, sino también a los factores antisigma. Un factor antisigma se une reversiblemente al factor σ correspondiente impidiendo su asociación con el núcleo ARN polimerasa. De este modo, los promotores que dependan de una forma de ARN polimerasa unida a ese factor sigma no se reconocerán adecuadamente y la expresión de los genes correspondientes será silenciada. Se han caracterizado numerosos factores antisigma. Algunos ejemplos son FlgM que inhibe al factor sigma flagelar σ^F (σ^{28}) en bacterias gram positivas y gram negativas; y SpoIIAB que inhibe a los factores sigma específicos de la esporulación σ^F y σ^G (Rojo F, 1999). Recordemos que los factores antisigma conocidos están ligados a la regulación de la actividad de σ^{70} y que hasta la fecha nada parecido se ha encontrado para σ^{54} (Wigneshweraraj *et al.*, 2005).

En lo que se ha dado en llamar “represión clásica”, la inhibición de la unión al promotor de la ARN polimerasa se efectuaría habitualmente mediante la unión de una proteína represora a secuencias específicas de ADN llamadas operadores (O) generalmente situadas en las proximidades del promotor o bien solapadas a él. Esta represión se eliminaría mediante la adición del inductor. El ejemplo clásico de represión es el ejercido por LacI sobre la expresión del operón *lac*. En ausencia de inductor, LacI se une al operador O1 que solapa con el promotor, inhibiendo la transcripción unas 20 veces. Existen otras dos secuencias, O2 y O3, situadas 401 pares de bases detrás y 92 pares de bases delante de O1 respectivamente. Como LacI es tetramérico, puede unirse simultáneamente a O1 y a uno u otro de los sitios O2 y O3 formando un lazo de ADN que modifica la geometría del promotor incrementando la represión unas 50 veces.

Para excluir a la ARN polimerasa del promotor, un represor no tiene necesariamente que unirse a un sitio que solape con la región de unión de la ARN polimerasa. Por ejemplo el represor CytR se une a un sitio en -70 en el promotor *deo* de *E. coli*, ayudado de dos dímeros CAP que se unen en -40 y -92 (flanqueando el sitio de CytR). El represor se estabiliza en el ADN por contactos directos con los dos dímeros CAP. El complejo inhibidor resultante no solapa con el lugar de unión de la ARN polimerasa. De hecho, en ausencia de CytR, CAP activa la transcripción cuando está unido a la posición -40 (tal como ya discutimos). Sin embargo el complejo formado por CytR y los dos dímeros CAP parecen doblar al ADN de manera que impide la unión de la ARN polimerasa en el promotor. Además CytR enmascara la porción activadora del CAP unido en -40, que de otro modo activaría la transcripción. Por su mecanismo de acción se propuso a CytR más que como represor como antiactivador en este promotor (Rojo, 1999).

IcIR inhibe en *E. coli* la transcripción del operón del acetato, *aceBAK*, en una segunda vía mediante su unión a la caja I, situada entre -125 y -99 del promotor de *aceB*. La unión del represor a la caja I tiene lugar incluso después de la formación del complejo abierto, induciendo el desensamblaje de dicho complejo y por tanto la represión de la transcripción. En este caso, la localización sobre el ADN del dominio C terminal de la subunidad α de la ARN polimerasa (α CTD), se desliza cerca de la caja I, indicando que la represión estaría causada por una interacción directa entre el α CTD e IcIR asociado a la caja I (Yamamoto y Ishihama, 2003).

Finalmente hemos de considerar el efecto represor de la transcripción ejercido por proteínas que se unen al ADN en número relativamente elevado y con baja especificidad de secuencia. Un ejemplo de este tipo lo constituye la represión del gen *dnaA* de *E. coli* mediante la oligomerización sobre él de la proteína DnaA que obstruye el acceso de la ARN polimerasa al promotor. Un caso interesante es el de la proteína asociada a nucleoide, H-NS de *E. coli* la cual aparte de su papel estructural en la cromatina bacteriana, resultante de su capacidad para obligar al ADN a formar superhélices, se sabe que

disminuye la expresión de muchos genes. A pesar de su poca especificidad de secuencia, se une preferentemente a regiones de ADN que muestran una curvatura intrínseca, en las que pueden formar complejos nucleoproteicos multiméricos en los que el ADN se dobla alrededor de la H-NS. Estos complejos pueden reprimir la transcripción mediante dos vías: bien bloqueando físicamente el acceso de la ARN polimerasa al promotor o bien alterando la topología del ADN en la vecindad del promotor, dificultando de este modo el inicio de la transcripción. Aunque es difícil diferenciar entre ambos mecanismos, parece que el silenciado del promotor *bgl* de *E. coli* es un ejemplo de la primera posibilidad (el sitio de unión de la H-NS cubre al promotor), mientras que la represión del promotor *proU* de *E. coli* es un ejemplo de la segunda. Cuando la represión se ejerce a través de modificaciones en la topología del ADN, el paso del inicio de la transcripción se impide no tanto porque la ARN polimerasa no pueda unirse como por la transición a complejo abierto.

Numerosos represores se unen al ADN en un modo que permite la unión simultánea de la ARN polimerasa al promotor, al menos *in vitro*. En algunos casos se ha establecido que en tales complejos ternarios la ARN polimerasa es incapaz de abrir las cadenas de ADN no pudiendo proseguir hacia un complejo abierto. Sin embargo, como la formación de estos complejos ternarios se observó siempre *in vitro*, con ensayos de improntas *in vitro* con DNasa I que se efectúan a concentraciones iónicas significativamente más bajas que las celulares, estas conclusiones hay que tomarlas con precaución. De todos modos, de lo dicho podemos distinguir dos tipos de estos represores: (1) los que se unen a sitios solapantes (al menos parcialmente) con el sitio de unión de la ARN polimerasa (-40 a +10 respecto al sitio de inicio de la transcripción) que podrían o no formar complejos ternarios estables *in vivo*, a pesar de formarlos *in vitro*, y (2) los que se unen a sitios que no solapan con los de la ARN polimerasa y que no presentarían problemas para unirse cerca o adyacentes a la ARN polimerasa bajo condiciones de concentración iónica elevada (puesto que el modelo de impedimento estérico sería improbable en estos represores). Entre el primer grupo tendríamos como ejemplos a los represores Spo0A del promotor *abrB* y a MerR en el promotor *merT* (en ausencia de mercurio) que se ha mostrado dificultan la transición de complejo cerrado a abierto. (En el caso de MerR se ha comprobado por improntas *in vivo* que represor y ARN polimerasa se unen simultáneamente al promotor *metT*. En el segundo grupo hay también algunos ejemplos como el represor GalR que reprime dos promotores del operón *gal*, P₁ y P₂ (separados 5 pares de bases), mediante la unión a dos operadores, O_E y O_I. El operador O_E se localiza en -60.5 y O_I en +53.5. En presencia de la proteína asociada a cromatina HU, GalR se une a ambos operadores y genera un lazo de ADN que inhibe la transcripción desde ambos promotores. Cuando no se forma el lazo, porque HU esté ausente o porque sólo esté presente el operador O_E, GalR no inhibe la unión de la ARN polime-

rasa a los promotores, aunque reprime a P_1 (no a P_2) por dificultar la transición desde complejo cerrado a complejo abierto. Hay que señalar que en este caso, GalR está unido en -60.5, posición propia de la unión de numerosos activadores, sugiere que probablemente no inhibe la unión de ARN polimerasa incluso en condiciones de alta fuerza iónica (Rojo, 1999).

En resumen se puede decir que el proceso de la regulación de la transcripción es sumamente complicado por cuanto no sólo la actividad y ciclo catalítico de la ARN polimerasa es muy complejo, sino porque multitud de proteínas interfieren con la transcripción, estimulándola o reprimiéndola, pero en definitiva alterando el ciclo transcripcional. No deja por ello de ser sorprendente el grado de especialización alcanzado por los reguladores transcripcionales para modular la expresión génica, sino también cómo estos procesos se engloban en circuitos complejos que modulando la expresión de los genes, permiten a las bacterias responder a los cambios ambientales. El interés del grupo por el conocimiento global de estas cuestiones se plasmó en la construcción de la base de datos AraC/XylS (Tobes y Ramos, 2000), que reunía el conocimiento disponible acerca de los reguladores pertenecientes a esta familia de activadores de la transcripción.

En este ámbito de interés, en la presente Tesis Doctoral se desarrolla una herramienta informática en el campo de las bases de datos, BacTregulators (Martínez-Bueno *et. al*, 2003), que incluye a la familia AraC/XylS, a la familia TetR, objetivo concreto de esta Tesis, y actualmente a la nueva aportación la familia IclR.

Simultáneamente a la realización de esta tarea, la conclusión del Proyecto de Secuenciación del genoma de *Pseudomonas putida* KT2440 nos permitió, detectar, describir y estudiar los factores σ presentes en el mismo, también mediante aproximaciones bioinformáticas, trabajo que se incluye como último capítulo de la presente Tesis Doctoral.

Objetivos

Objetivos

En este trabajo podemos diferenciar 3 objetivos principales:

1. Desarrollo de la base de datos BacTregulators.
2. Definición y estudio de la familia TetR de reguladores de la transcripción.
3. Detección, descripción y estudio de los factores sigma presentes en el genoma de *Pseudomonas putida* KT2440.

Materiales y Métodos

1. Búsquedas por similitud de secuencia.

- BLAST, del NCBI, usado localmente:

<http://www.ncbi.nlm.nih.gov/Sitemap/index.html>

- BLAST y PSI-BLAST, <http://www.ncbi.nlm.nih.gov>, usados a través de la web.

2. Bases de datos de secuencias de proteínas.

- SWISSPROT, TrEMBL y TrEMBL New del Instituto Suizo de Bioinformática. Actualmente UniProt:

ftp://ftp.expasy.org/databases/unipro/current_release/Knowledgebase/taxonomic_divisions/ (sólo las categorías de archea y bacteria).

3. Genomas.

- Genoma de *P. putida* KT2440 (TIGR; Julio de 2002): <http://www.tigr.org>.

- Genoma de *P. aeruginosa* PA01 (TIGR; Julio de 2002):

<http://www.tigr.org>

- Genomas anotados depositados en el NCBI:

<ftp://ftp.ncbi.nih.gov/genomes/Bacteria>.

4. Base de datos de Taxonomía del NCBI.

Hemos empleado su jerarquía taxonómica en BacTregulators:

<ftp://ftp.ncbi.nih.gov/pub/taxonomy>

5. Código funcional (COG) del NCBI.

<ftp://ftp.ncbi.nih.gov/pub/COG/COG>, archivo "fun.txt".

6. Alineamientos múltiples y construcción de árboles.

- CLUSTALW, <http://clustalw.genome.ad.jp/>

- MULTIALIN, <http://protein.toulouse.inra.fr/>

7. Recursos sobre familias de proteínas.

- INTERPRO, <http://www.ebi.ac.uk/interpro>

- PROSITE, <http://www.expasy.org/prosite>

- Pfam, <http://www.sanger.ac.uk/Software/Pfam>

- TIGRfam, <http://www.tigr.org/TIGRFAMs>

- Prodom, del INRA, <http://protein.toulouse.inra.fr/prodom>

8. Predicción de estructura secundaria de proteínas.

Método de predicción de estructura secundaria HNN del “Pôle BioInformatique Lyonnais”, <http://npsa-pbil.ibcp.fr/>

9. pftools: pfmake, pfsearch y pfscale.

<http://www.isrec.isb-sib.ch/ftp-server/pftools>

"pfmake", herramienta de construcción del perfil.

“pfsearch”, herramienta que enfrenta un perfil contra bases de datos.

“pfscale”, herramienta de calibración del perfil.

10. Estudio tridimensional de moléculas.

-Los archivos “.pdb” del *Protein Data Bank* (PDB):

<http://www.rcbs.org/pdb>

- Los visores de estructura tridimensional, WebLab Viewer Lite, (ya no disponible) y Swiss-PDB-Viewer, <http://us.expasy.org/spdbv/>

-Programa CE (*Combinatorial Extension*), para superposición de estructuras tridimensionales, <http://cl.sdsc.edu/ce.html>

11. Programas propios.

Para afrontar los problemas concretos que iban surgiendo durante la realización del trabajo hemos construido numerosas aplicaciones. Para la escritura del código de estos programas se han usado lenguaje C y como compilador Dev-C++. Las aplicaciones más destacables son:

- "*Sequniq*" que integra un conjunto de secuencias en un grupo no redundante.

- "*Provalidator*" automatiza el proceso de validación de un perfil.

- "*Calculavectores*" que permite realizar cálculos geométricos con los átomos de una molécula a partir de su archivo ".pdb".

- Programa de cálculo de la frecuencia de aminoácidos en cada una de las posiciones de contacto con el ADN en el alineamiento múltiple del dominio de unión a ADN de los miembros de la familia TetR (Tabla 31).

- Programa para identificar los factores σ potenciales de *P. putida* KT2440 a partir de las salidas del BLAST local para las secuencias anotadas como factores σ en SwissProt y en *P. aeruginosa* PA01 contra el genoma de *P. putida* KT2440.

12. SQL como gestor de la Base de Datos BacTregulators, Apache como programa Servidor y PHP para programar las Entradas de BacTregulators.

Capítulo 1:

La Base de Datos de reguladores transcripcionales BacTregulators

4.1. Resumen

Cuando se inició la elaboración de esta parte de la Tesis Doctoral ya existía la Base de datos AraC-XylS (Tobes y Ramos, 2002). Se decidió la construcción de la Base de datos BacTregulators (www.bactregulators.org) como herramienta para recoger e integrar la información disponible acerca de las proteínas pertenecientes a distintas familias de reguladores transcripcionales de procariotas. Se convirtió al nuevo formato una versión actualizada de AraC-XylS y como trabajo de esta Tesis Doctoral, se decidió la inclusión de la familia TetR de reguladores de la transcripción.

Las proteínas incluidas en BacTregulators se identificaron mediante la exploración de los 123 genomas disponibles en las bases de datos SWISS-PROT y TrEMBL en 2004 utilizando perfiles que definían a estas familias. Como resultado del proceso de integración, a finales de 2004 la base de datos incluía 1933 proteínas diferentes de la familia AraC-XylS y 2353 de la familia TetR. La definición de una entrada en BacTregulators se basa en la secuencia de la proteína, el organismo de origen, el elemento de genoma en el que se localiza el gen y la posición del gen en el genoma. La base de datos BacTregulators proporciona una herramienta innovadora que permite obtener un informe con los datos correspondientes a las entradas relacionadas por similitud de secuencia con la entrada de la petición. BacTregulators detecta y clasifica los reguladores pertenecientes a las familias AraC-XylS y TetR presentes en los genomas procariotas contribuyendo así a una anotación más segura de los reguladores en los genomas. La información recogida sobre cada proteína puede ser útil para caracterizar un nuevo regulador o compilar información acerca de las propiedades biológicas de un regulador ya conocido.

4.2. Introducción

El mecanismo más importante que las bacterias usan para adaptar su fisiología a las variaciones de las condiciones ambientales se basa en la regulación de la transcripción (Ramos *et al.*, 1997, 2001; Ishihama, 2000). Los reguladores de la transcripción proporcionan un "ajuste fino" de la expresión génica para afrontar cambios ambientales específicos y condiciones de estrés, el cual se logra mediante la acción conjunta de factores sigma que confieren selectividad de promotor y de reguladores transcripcionales que modulan la actividad de la ARN polimerasa (Ramos *et al.*, 1997; Gallegos *et al.*, 1997; Rojo, 1999; Vicente *et al.*, 1999; Ishihama, 2000; Wigneshweraraj *et al.*, 2005). Desde 1990 cuando Ramos y colaboradores describieron la incipiente familia AraC/XylS, ésta se ha actualizado y ha sido objeto de revisiones sucesivas (Gallegos *et al.*, 1997; Egan, 2002). La información más reciente se ha recopilado en la Base de

datos AraC-XylS (Tobes y Ramos, 2002). En un esfuerzo por comprender las redes de regulación que implican a reguladores pertenecientes a diferentes familias, se decidió extender la base de datos incorporando a otras familias de reguladores. Este nuevo enfoque pretende ayudar a definir las redes de regulación que controlan las respuestas de los microorganismos a los diferentes desafíos ambientales. Además, la irrupción de genomas secuenciados de procariotas está dando lugar a un crecimiento exponencial en la cantidad de secuencias de proteínas disponibles, pero por el contrario, la adquisición experimental de información sobre estas proteínas es relativamente limitada. La información recogida en la base de datos BacTregulators puede ser útil para diseñar ensayos de identificación de función, así como para facilitar la anotación de genes en los nuevos genomas disponibles.

Por otro lado, los datos procedentes de los genomas secuenciados, a menudo no están integrados, requiriéndose un esfuerzo adicional para evitar la redundancia. BacTregulators se ha creado con la intención de integrar la información sobre reguladores procariotas del NCBI y de las bases de datos de proteínas SWISS-PROT y TrEMBL.

La base de datos BacTregulators recopila datos de tres familias de reguladores: AraC-XylS, TetR e IclR. La información de los miembros de la familia AraC-XylS se ha actualizado y en este momento (Noviembre 2005) abarca 1933. Dichas secuencias se han obtenido explorando los genomas de bacterias y arqueas incluidos en la base de datos RefSeq NCBI y de las bases de datos de proteínas SWISS-PROT y TrEMBL (SPTR). Las secuencias se han integrado en un conjunto no redundante.

La familia a la que ha contribuido esta Tesis Doctoral es TetR, mientras que el trabajo de A.J. Molina-Henares está contribuyendo a la de IclR. En esta Tesis Doctoral se ha desarrollado el perfil para TetR, que se ha validado y se ha usado para detectar las proteínas TetR tanto en genomas como en la base de datos SPTR. Además las hemos reunido en un conjunto no redundante de 2353 secuencias pertenecientes a la familia TetR.

4.3. Métodos y sistemas de análisis

4.3.1. Selección de un conjunto de proteínas no redundante de cada familia mediante análisis de las bases de datos SPTR y NCBI.

El primer paso en el desarrollo de BacTregulators fue seleccionar los conjuntos de reguladores de la transcripción con los perfiles que definían a cada familia (Gallegos *et al.*, 1997; Tobes *et al.*, 2004). El examen incluyó búsquedas en SPTR y en todos los genomas completos de proyectos de secuenciación (disponibles en el sitio de recursos de genomas microbianos del NCBI). El proyecto HAMAP (High-quality Automated and Manual Annotation of Microbial Proteomes) (Gattiker *et al.*, 2003) tiene como objetivo la anotación automática en SWISS-PROT de un porcentaje significativo de proteínas procedentes de los proyectos de secuenciación de genomas de bacterias y arqueas. El proyecto HAMAP ha iniciado un proceso de incorporación de los reguladores de la transcripción desde la secuencia de genomas secuenciados a SPTR. En la actualidad, el solapamiento entre conjuntos de secuencias de proteínas reguladoras detectado en SPTR y en los proyectos de secuenciación de genomas es sólo parcial. Por esa razón fue necesario desarrollar una herramienta que integrase las secuencias de proteínas de SPTR con las procedentes de los genomas disponibles en el NCBI en un conjunto no redundante para cada familia (Figura 1). Esta información está disponible en el sitio web de BacTregulators (www.bactregulators.org).

4.3.2. La entrada en BacTregulators.

La extracción de información sobre reguladores es una tarea lenta y compleja. Cuando se extraen datos acerca de un regulador siempre es posible asignarlos a una secuencia de proteína y a un microorganismo. Además, si el microorganismo se ha secuenciado, podemos asignar la información a una proteína codificada por un marco de lectura abierto concreto localizado en una posición precisa en un elemento de genoma.

La expansión en la disponibilidad de genomas secuenciados enfatiza la importancia de los datos genómicos en la definición de una proteína. Por otro lado, para asegurar flexibilidad y precisión en la base de datos es necesario mantener la información al máximo nivel de dispersión, considerando que es sencillo combinar información atendiendo a características comunes pero que es muy complejo separar información una vez mezclada.

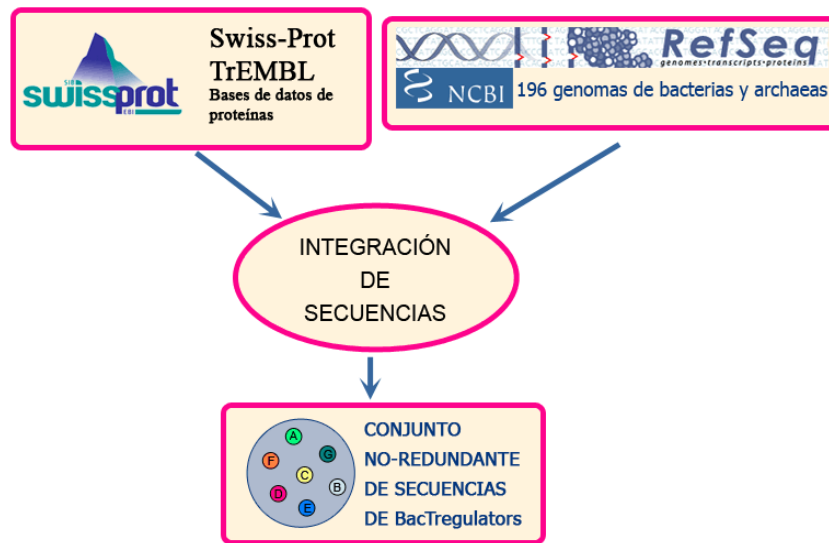


Figura 1. Integración de las secuencias de proteínas de SPTR con las procedentes de los genomas disponibles en el NCBI en un conjunto no redundante para cada familia.

Así, la definición de una entrada en BacTregulators se fundamenta en la secuencia de la proteína, el organismo de origen, el elemento de genoma (cromosoma, plásmido, fago, etc) y la localización específica del gen en su elemento de genoma correspondiente (Figura 2).

El microorganismo origen se registra de acuerdo al criterio taxonómico de la base de datos taxonómica del NCBI (Wheeler *et al.*, 2000) (<ftp://ftp.ncbi.nih.gov/pub/taxonomy/>).

Además, la base de datos BacTregulators almacena la estructura del árbol taxonómico haciendo así posible acceder al linaje de cada microorganismo origen. Las proteínas con diferencias en alguno de estos cuatro niveles se cargan como entradas diferentes (Figura 3).

De esta manera el número de entradas en la base de datos BacTregulators (8557) es mayor que el número de secuencias (4286) puesto que varias entradas pueden compartir la misma secuencia. Sin embargo todas las entradas con la misma secuencia están interconectadas.

La mayoría de las entradas incluidas en BacTregulators se han generado automáticamente tras la búsqueda con los perfiles para las familias AraC-XylS y TetR en las bases de datos SWISS-PROT, TrEMBL y TrEMBL-new (Boeckmann *et al.*, 2003) y los proteomas completos disponibles en el Refseq NCBI “*Microbial Genomes Resources*” (Pruitt and Maglott, 2001).

Definición de nuevo registro:

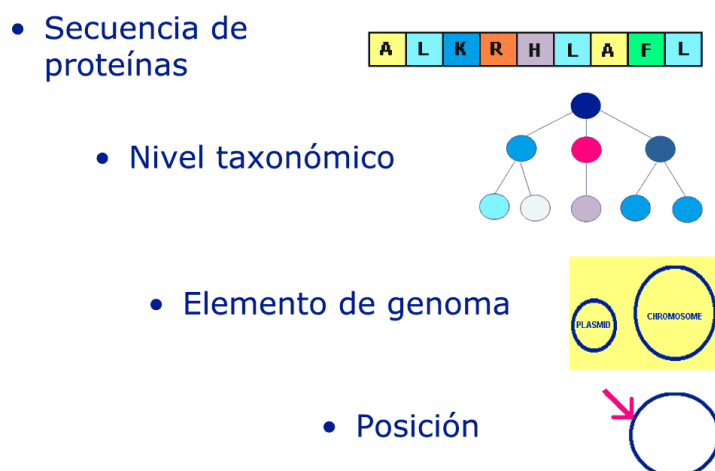


Figura 2: Cada entrada BacRegulators se define atendiendo a la secuencia de la proteína, la clasificación taxonómica del organismo origen, el elemento de genoma y la posición en la que se sitúa el gen correspondiente.

La generación automática de entradas fue diferente dependiendo de la base de datos origen. Cuando la proteína procedía de la base de datos de genomas microbianos del NCBI, el proceso fue simple: cada regulador detectado en un proteoma y correspondiente a un genoma completo originó una entrada con información concerniente a cada uno de los 4 niveles de definición (secuencia, organismo origen, elemento de genoma y posición en el elemento de genoma).

Sin embargo si la base de datos origen fue SPTR, el proceso fue más complejo. La primera diferencia fue que la definición de una entrada en la base de datos SPTR sólo se basa en secuencia de proteína (Boeckmann *et al.*, 2003), en consecuencia una misma entrada SPTR puede corresponder a más de un organismo y más de un elemento de genoma. Otra diferencia a tener en cuenta es la posibilidad de redundancia, especialmente entre las entradas de TrEMBL y TrEMBL-new (Gattiker *et al.*, 2003). Así para la generación automática de entradas de la base de datos BacRegulators desde la base de datos SPTR hemos usado la siguiente estrategia:

(1) Si la entrada SPTR correspondía a un solo organismo y a un solo elemento de genoma, se generaba una sola entrada en BacRegulators.

(2) Si la entrada SPTR correspondía a más de un organismo o a más de un elemento de genoma, se originaron tantas entradas en BacRegulators como fueron necesarias. Para evitar la redundancia, la presencia de cada entrada generada en la base de datos se examinó antes de su inclusión como una entrada nueva en BacRegulators.

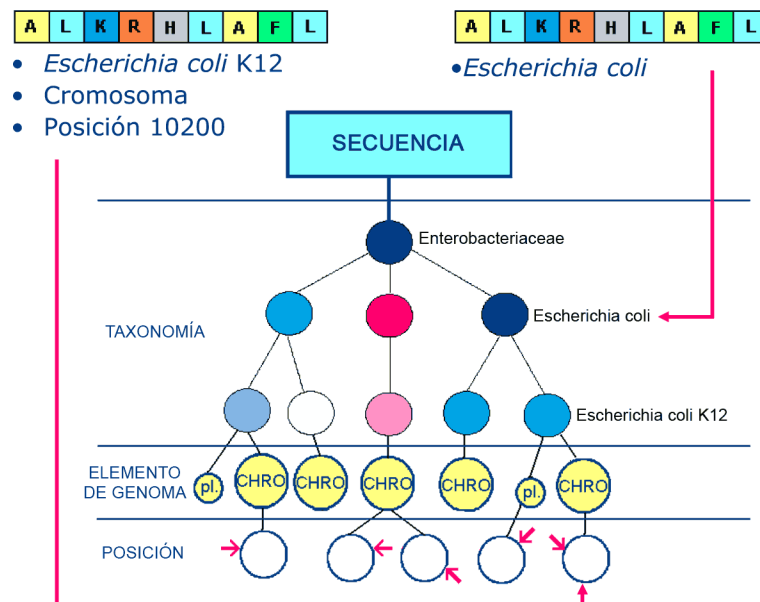


Figura 3: Dos secuencias idénticas pueden pertenecer a entradas diferentes dependiendo de las diferencias en cualquiera de los cuatro niveles que definen a la entrada BacTregulators: secuencia de la proteína, organismo origen, elemento de genoma y posición. En esta figura, por ejemplo, las entradas 1 y 2 comparten la misma secuencia pero difieren en los otros 3 niveles, el organismo origen, el elemento de genoma y la posición en la que el gen que codifica a la proteína se sitúa.

Durante la fase de extracción de datos de los trabajos publicados, el mantenedor de la base de datos asigna la información a la entrada apropiada, usando herramientas diseñadas específicamente para esta tarea. Si ocasionalmente la entrada apropiada no está incluida en la base de datos, el mantenedor puede crear entradas adicionales para asignar la información extraída de las referencias. Aunque el trabajo de inclusión de nuevos datos de conocimiento y secuencias es continuo, se incorporarían a BacTregulators mediante actualizaciones periódicas con el fin de asegurar la integridad y calidad de los datos.

4.4. Resultados

4.4.1. Información contenida y estructura de la base de datos BacTregulators.

La estructura de la base de datos BacTregulators está especialmente diseñada para incluir datos procedentes de genomas. La información se almacena en una forma altamente estructurada, organizada en 158 campos agrupados en 26 tablas en una base de datos administrada mediante el sistema MySQL y el lenguaje PHP. En el Anexo I de esta Tesis Doctoral se presenta al organización de Tablas y Relaciones en BacTregulators.

La base de datos contiene 3 tipos de datos: (1) secuencias, (2) conocimiento y (3) referencias.

(1) La información concerniente a las secuencias se ha tratado mediante integración, filtrado y examen de cada secuencia de proteína. Se proporcionan tres conjuntos no redundantes de secuencias de proteínas para cada familia de reguladores:

- i. El conjunto completo de secuencias de proteínas obtenido de las bases de datos SPTR y genomas procariotas del NCBI.
- ii. El conjunto de secuencias de proteínas procedentes de las bases de datos SWISS-PROT, TrEMBL y TrEMBL-new.
- iii. El conjunto de secuencias de proteínas procedentes de los genomas procariotas del NCBI.

Como en la base de datos AraC-XylS descrita previamente (Tobes y Ramos, 2002), también se proporciona estos tres conjuntos de secuencias separadas por dominios (unión a ADN, N-terminal y C-terminal), resultando un total de 9 subconjuntos de secuencias para cada familia de reguladores.

(2) En BacTregulators el proceso de extracción de conocimiento combina procedimientos automáticos y manuales. Algunos datos asignados a cada entrada se han obtenido automáticamente de las bases de datos SPTR y del NCBI *Microbial Resources* (ver más adelante). El resto del conocimiento se ha extraído manualmente de referencias bibliográficas. Para el conocimiento extraído manualmente hemos adoptado una estructura especial correspondiente a datos de texto, usando el párrafo como unidad de información, estando cada párrafo referenciado individualmente. Esta estructura identifica el origen de cada porción de conocimiento. Durante la extracción manual de conocimiento, hemos diseccionado los datos experimentales que apoyan lo dicho en cada párrafo. Este tipo de datos de conocimiento se ha llamado "evidencia experimental" y reciben

un tratamiento especial en la base de datos. Cuando hay evidencia experimental que apoya el conocimiento expresado en un párrafo, se muestra un anclaje a dicha evidencia. Estos datos son útiles para evaluar la fiabilidad del dato de conocimiento.

(3) Cada párrafo se referencia y se ancla al Abstract de Medline correspondiente. Actualmente la base de datos de referencias contiene 457 referencias correspondientes a la familia TetR.

El sitio web de BacTregulators también proporciona tutoriales e información gráfica de la estructura tridimensional de las proteínas que se hayan cristalizado.

4.4.2. Acceso a la información en BacTregulators.

La información contenida en la base de datos BacTregulators es accesible por varias vías. Los datos se pueden buscar mediante el identificador BacTregulators o bien por el correspondiente identificador SWISS-PROT-TrEMBL o NCBI. La búsqueda mediante familia, microorganismo, nombre y COG permite al usuario obtener conjuntos específicos de proteínas. Por último, el uso de búsquedas de texto en todos los campos de la base de datos permite resultados más flexibles. La base de datos de referencias se puede consultar independientemente mediante búsqueda de texto.

La información acerca de una entrada concreta se muestra en cuatro secciones (Figura 4).

(1) La primera sección (barra azul vertical) muestra los datos que definen a la entrada: identificador de secuencia, organismo origen, elemento de genoma y posición en dicho elemento.

(2) La segunda sección (barra verde azulada vertical) contiene los datos obtenidos automáticamente de otras bases de datos: número de acceso SPTR o NCBI, base de datos origen, nombre completo de la proteína y nombre corto, nombre del gen, orientación del gen, código COG, código funcional del NCBI y última actualización de esos datos.

(3) La tercera sección (barra verde vertical) muestra el conocimiento extraído manualmente de artículos publicados. Este conocimiento se estructura en los siguientes campos: función, genes regulados, redes de regulación, efectores, posición y orientación del gen, promotores de los genes regulados, promotores del regulador, dependencia de factores sigma, patogenicidad, aplicaciones, datos sobre mutaciones, estructura 3D, oli-

gomerización, similitudes, comentarios y última actualización de estos datos obtenidos manualmente (Tobes y Ramos, 2002). Para cada uno de estos campos de conocimiento se proporciona archivos de texto descargables con la información correspondiente a cada campo.

(4) La última sección (barra naranja vertical) proporciona acceso a la información relacionada con la secuencia de la proteína, tanto a la secuencia completa como a los 3 fragmentos de la misma, N-terminal, secuencia conservada (normalmente la del fragmento de unión a ADN) y secuencia C-terminal. Por último la subsección de similitud por BLAST permite al usuario acceder a un archivo con el resultado de la comparación mediante BLAST de la secuencia de la proteína de la entrada con la del resto de los miembros de la familia, y una lista de entradas de secuencia similar ordenada mediante el valor E de BLAST.

4.5. Discusión

4.5.1. Integración del conocimiento modificable en el paso de recuperación.

Cuando los usuarios acceden a una entrada concreta, el conocimiento asociado con las entradas relacionadas puede ser útil para inferir características funcionales. Este concepto fue la base para diseñar una herramienta innovadora que presenta al usuario un informe integrado de un conjunto de varias entradas relacionadas con una entrada inicial. Aunque la información se sigue manteniendo al máximo grado de dispersión asociada con cada entrada, la información correspondiente a un conjunto dado de entradas relacionadas se puede combinar en el paso de la recuperación. Se ha considerado la similitud entre secuencias como el criterio clave para definir un grupo de entradas relacionadas. De este modo, el nivel de integración del conocimiento puede ajustarse por el usuario actuando sobre los parámetros de similitud y de nivel taxonómico.

Para que el usuario seleccione el conjunto de entradas relacionadas, se le muestra una lista con todas las entradas de BacTregulators con una similitud significativa (valor E) con la secuencia inicial. Todas las entradas que comparten la misma secuencia que la inicial se muestran con fondo amarillo.

En un primer paso, el usuario selecciona un valor de E como el umbral para definir el conjunto de secuencias relacionadas. En un segundo paso, este conjunto puede limitarse mediante criterio taxonómico: el usuario elige el nivel taxonómico requerido para la inclusión de una entrada en el conjunto de entradas relacionadas. El linaje co-

rrespondiente al microorganismo origen de la entrada inicial se muestra para permitir al usuario afinar los límites taxonómicos (Figura 5). Una vez definidos el umbral de similitud y los límites taxonómicos, se muestra un informe integrado del conjunto de entradas seleccionadas. Los datos se muestran en campos en secciones diferentes. Cada sección incluye los datos de un campo en todas las entradas del conjunto. El identificador de entrada de BacTregulators se indica al inicio de cada párrafo. Una lista de las entradas integradas también es accesible en la página del informe (Figura 6).

Es incuestionable que las proteínas ortólogas a menudo comparten muchas características funcionales. De ahí que el inferir una función a partir de la similitud de secuencia (Gattiker *et al*, 2003) adquiere un sentido particular para las proteínas ortólogas. Sin embargo en la práctica, es extremadamente difícil definir proteínas ortólogas. Hemos evitado este problema permitiendo al usuario elegir un conjunto de entradas del que se puede inferir la función. Esta herramienta puede servir de ayuda al usuario a la hora de decidir la posible función de un regulador de función desconocida. Así un conjunto de entradas con una secuencia similar pueden incluir algunas entradas con función definida y otras entradas con función desconocida. Si la función es similar en todas las proteínas caracterizadas del grupo, sería probable que las proteínas no caracterizadas posean también una función similar.

La herramienta que describimos aquí, facilita la selección de agrupamientos funcionales de entradas para las que el conocimiento puede ser propagado a entradas con función no publicada.

BacTregulators está especialmente ideada para detectar los reguladores en los genomas procariotas y clasificarlos en familias. BacTregulators facilita la anotación segura de los reguladores en los genomas. Además la información recogida para cada proteína de una familia puede ser útil para los investigadores que deseen caracterizar un nuevo regulador o investigar las propiedades biológicas de un regulador conocido.

BacTregulators

Haciendo clic aquí el usuario puede obtener un informe integrado con datos de proteínas cuya secuencia es similar al registro nº 1

Click here to retrieve data from entries whose sequence is similar to entry 1

BacTregulators Entry No. 1 **TetR family**

Sequence:
1
Taxonomy: Escherichia coli O6
Genome Element: Undefined
Position: Undefined

Los datos que definen un registro en BTR están en la sección con barra vertical azul oscuro

Accession Number: P34000
Source Database: Swiss-Prot
Protein Name: Potential acrAB operon repressor
Protein Short Name: ACRR_ECOLI
Gene Name: ACRR OR B0464 OR C0582 OR SF0409
Automatically obtained data last update: 06-21-2003

Los datos que se obtienen automáticamente están en la sección con barra azul claro

Activator or repressor: Repressor
Function: AcrR represses its own expression. [References: 1]
AcrR represses the expression of the acrAB operon. [References: 152]
The expression of the gene micF could be under the AcrR control. [References: 152]

Los datos de conocimiento obtenidos manualmente por el experto están en la sección con barra verde

Regulated genes:
AcrR represses the expression of the acrAB operon. [References: 1]
The operon acrAB encodes the AcrAB multidrug efflux pump. [References: 2]
The acrAB operon does not contain a gene coding for an outer membrane channel. [References: 152]
TolC is required as a component of the AcrAB efflux machinery. TolC is upregulated by MarA. [References: 146]
The TolC channel-tunnel spans the bacterial outer membrane multidrug efflux pump. [References: 153]
AcrA is a periplasmic lipoprotein whose amino-terminus is anchored to the inner membrane. AcrA belongs to the membrane fusion proteins family (MFP). [References: 2]
In the mature AcrA sequence, two hydrophobic regions exist near the N and C termini. These regions are universally present among MFP members and may serve as regions for interaction with cytoplasmic and outer membrane components of the pump. [References: 152]

Cada párrafo es referenciado y enlazado a un artículo de Medline

Evidences: 143
MarA seems not to be essential for the transcriptional activation of acrAB. [References: 123] [Evidences: 148]
The quorum sensing regulator SdiA also controls the expression of the AcrAB multidrug efflux pump conferring resistance against quinolones. [References: 145]
The acrAB expression is influenced by the growth rate. In slow-growth conditions, the expression of acrAB is upregulated, even in the absence of antimicrobial compounds. Up-regulation of acrAB is observed in a slow-growing, antibiotic-resistant bacterium. [References: 148] [Evidences: 144]
Glucose limitation increased the level of acrAB expression, at least in part due to increased transcription, at all growth rates. This suggests that cAMP may act, via CRP, as a positive regulator of acrAB expression. [References: 148]
Transcription of acrAB, upregulated by slow growth, is typical of a gearbox regulated gene and it is interesting to note that the -10 region of the acrAB promoter sequence (-CGAGGTTT-) resembles a gearbox-type sequence homology rather than a sigma-70 consensus sequence. [References: 148]
Slow growth regulation of acrAB is not controlled by RpoS. [References: 148] [Evidences: 145]

Algunos párrafos están enlazados a evidencias experimentales que apoyan estos datos de conocimiento

Effectors: molecules and conditions (temperature, pH...): General stress conditions increased transcription of acrR even more strongly than that of acrAB. [References: 1]
Orientation and position: AcrR is situated 142 bp downstream and divergently transcribed from the acrAB operon. [References: 296]
Promoters of the gene coding the protein: The non coding region between acrR and acrA has 128 bp and contain the acrR and acrAB promoters. [References: 151]
The upstream region of the acrAB operon contains a candidate AcrR operator. This binding site overlaps with the acrAB promoter. [References: 151]

La sección con barra naranja contiene información sobre las secuencias de proteínas

Los datos de conocimiento están clasificados en campos

Involvement of the protein in other processes: Transcription of acrAB, upregulated by slow growth, is typical of a gearbox regulated gene and it is interesting to note that the -10 region of the acrAB promoter sequence (-CGAGGTTT-) resembles a gearbox-type sequence homology rather than a sigma-70 consensus sequence. [References: 151]

Aquí el usuario puede acceder a los resultados de BLAST de la secuencia correspondiente al registro nº 1 contra todas las secuencias de la familia TetR

Knowledge last update: 07-31-2003

PROTEIN SEQUENCES
Complete sequence
N-terminal sequence
Conserved region sequence
C-terminal sequence
BLAST similarities
BLAST result file
Table of entries with similar sequence

El usuario puede acceder a la secuencia completa o de los diferentes dominios de la proteína. El alineamiento de cada secuencia con el perfil de la familia nos permite definir el dominio conservado o dominio de unión a ADN. Los fragmentos restantes de la secuencia de la proteína se llaman dominio N-terminal y C-terminal

Figura 4. Las secciones de una entrada en BacTregulators.

BacTregulators

El usuario ha seleccionado Gammaproteobacteria como límite taxonómico para el informe integrado

To retrieve data from entries whose sequence is similar to entry 1, please select a similarity threshold (BLAST E value) in the list and a taxonomic limit to restrict the selected set of entries.

Prokaryote (default)	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia	Escherichia coli	Escherichia coli O6
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

E BLAST	Id	AN	Long name	Organism
1e-123	1	P34000	Potential <i>acrAB</i> operon repressor	Escherichia coli O6
	2	P34000	Potential <i>acrAB</i> operon repressor	Shigella flexneri
	1172	AAP15941	<i>acrAB</i> operon repressor	Shigella flexneri 2a str. 2457T
	1790	26246477	Potential <i>acrAB</i> operon repressor	Escherichia coli CFT073
	1806	16128448	<i>acrAB</i> operon repressor	Escherichia coli K12
	2324	24111847	<i>acrAB</i> operon repressor	Shigella flexneri 2a str. 301
	2335	30061964	<i>acrAB</i> operon repressor	Shigella flexneri 2a str. 2457T
1e-122	570	Q8X5D3	<i>acrAB</i> operon repressor	Escherichia coli O157:H7
	1818	15829771	<i>acrAB</i> operon repressor	
	1831	15800193	<i>acrAB</i> operon repressor	
1e-107	561	Q8XEW7	<i>acrAB</i> operon repressor (TetR/AcrR family) (Potential operon repressor)	
	562	Q8XEW7	<i>acrAB</i> operon repressor (TetR/AcrR family) (Potential operon repressor)	
	2269	1759457	potential <i>acrAB</i> operon repressor	
	2290	29142771	potential <i>acrAB</i> operon repressor	
	2294	16763857	<i>acrAB</i> operon repressor (TetR/AcrR family)	Salmonella typhimurium LT2
3e-094	154	Q93K42	AcrR protein	Klebsiella pneumoniae
6e-094	222	Q9AEG3	Putative AcrR protein	
4e-074	953	Q8ZC89	Putative tetR-family transcriptional regulatory operon repressor)	
	2808	16123293	putative tetR-family transcriptional regulatory	
	2815	22124958	<i>acrAB</i> operon repressor	
3e-067	206	Q8VPB0	Regulatory protein AcrR	
9e-055	1			Salmonella typhimurium
2e-043	1			Pseudomonas syringae pv. tomato str. DC3000
3e-040	1			Pseudomonas putida
7e-040	192	Q9KJC4	AcrR	Pseudomonas putida
	2102	16988121	transcriptional regulator TtgR	
1e-036	518	Q8ZLN6	Transcriptional repressor for <i>envCD</i> (<i>acrE</i>)	
	2303	16766684	transcriptional repressor for <i>envCD</i> (<i>acrE</i>)	
2e-033	6	P31676	Potential <i>acrEF/envCD</i> operon repressor	
	1813	16131152	putative transcriptional regulator	
	1826	15833390	putative transcriptional regulator	
	1839	15803798	putative transcriptional regulator	Escherichia coli O157:H7 EDL933

Registros con la misma secuencia que el registro nº 1 se muestran en amarillo

Los registros que comparten secuencia se muestran en el mismo tono de gris: los registros 570, 1818 y 1831 comparten secuencia. Los registros 561, 562, 2269, 2290 y 2294 también comparten secuencia

Todos los registros de organismos incluidos en el grupo de las Gammaproteobacteria con un valor de E menor que 1e-36 se incluirán en el informe integrado

El usuario ha seleccionado un valor de umbral E=1e-36 para definir el conjunto de secuencias relacionadas con la nº 1

Un cuadro verde en esta columna indica que se han asociado a este registro datos de conocimiento extraídos manualmente por expertos

Figura 5. Lista de entradas relacionadas con la entrada de ID BacTregulators 1. El usuario puede seleccionar un valor umbral de E y los límites taxonómicos para la recuperación de los datos en el informe integrado.

BacTregulators

Knowledge data from entries with sequence similar to entry 1 (E major than 1e-036) in Gammaproteobacteria (total of 28 entries): [List of entries](#)

Function

[BTR-1] AcrR represses its own expression. [References: 1]

[BTR-1] AcrR represses the expression of the acrAB operon.

[BTR-1] The expression of the gene micF could be under the

[BTR-192] ArpR seems to be a repressor for expression of the

[BTR-192] ArpABC in *Pseudomonas putida* S12 is only involved in multidrug efflux pumps. [References: 198]

RegulGenes

[BTR-1] AcrR represses the expression of the acrAB operon.

[BTR-1] The operon acrAB encodes the AcrAB multidrug efflux pump.

[BTR-1] The acrAB operon does not contain a gene coding for an outer membrane channel. [References: 152]

[BTR-1] TolC is required as a component of the AcrAB efflux machinery. TolC is upregulated by MarA. [References: 146]

[BTR-1] The TolC channel-tunnel is a large exit duct for the AcrAB multidrug efflux pump. [Reference: 146]

[BTR-1] AcrA is a periplasmic lipoprotein whose amino-terminus is anchored to the inner membrane. AcrA belongs to the membrane fusion proteins family (MFP). [References: 2]

[BTR-1] In the mature AcrA sequence, two hydrophobic regions exist near the N and C termini. These regions are universally present among MFP members and may serve as regions for interaction with cytoplasmic and outer membrane components of the host cell.

[BTR-1] Slow growth regulation of acrAB is not controlled by RpoS. [References: 148] [Evidences: 145]

Effectors

[BTR-1] General stress conditions increased transcription of acrR even more strongly than that of acrAB. [References: 1]

OrientPos

[BTR-1] AcrR is situated 142 bp downstream and divergently transcribed from the acrAB operon. [References: 296]

[BTR-192] arpR is situated upstream.

[BTR-192] arpR is 248 bp and divergently transcribed from the acrAB operon.

PromotersText

[BTR-1] The non coding region between acrR and acrA has 128 bp and contain the acrR and acrAB promoters. [References: 151]

[BTR-1] The upstream region of the acrAB operon contains a candidate AcrR operator sequence. It is a perfect palindrome of 24 bp that overlaps with the acrAB promoter. [References: 151]

[BTR-1] There is a MarA binding site located in the coding region of acrR. [References: 151]

[BTR-1] There is a MarA binding site upstream of tolC. [References: 151]

[BTR-192] The intergenic region arpR/arpA contains a putative sigma70-dependent promoter motif [tttacaAACAACCAAGAATGTAAgtatatt]. [References: 198]

RegGenProm

[BTR-1] Transcription of acrAB, upregulated by slow growth, is typical of a gearbox regulated gene and it is interesting to note that the -10 region of the acrAB promoter sequence (-CGAGGTTT-) resembles a gearbox-type sequence homology rather than a sigma-70 consensus sequence. [References: 148]

Pathogen

[BTR-1] AcrR represses the expression of the operon acrAB, which confers multidrug resistance. [References: 1]

Annotations:

- Here the user can access to the list of entries integrated in this report and see their E value, BTR identifier, accession number, name and organism. In this list the entries with curator extracted-knowledge are marked with a green box. These entries with a green box are the source of the knowledge data that appear integrated in this report.
- A link to the specific entry is available
- In each field the data appears ordered by the similarity with entry 1.
- The knowledge data are integrated by fields

Figura 6. Informe de los datos de conocimiento de las entradas con secuencia similar a la secuencia de la entrada 1 ($E < 1e-036$) en Gammaproteobacterias.

ANEXO I

5. Organización de la base de datos BacTregulators

La base de datos BacTregulators es el resultado de la unión de tres bases de datos:

(1) Base de Datos de Proteínas (secuencias de aminoácidos y de ADN tales como, promotores, sitios de unión al ADN, genomas, cuando estén secuenciados, etc).

(2) Base de Datos de Referencias Bibliográficas (incluye no solo los datos de las referencias de cada proteína, sino párrafos extraídos, y evidencias experimentales de dichas referencias y del “mantenedor” que los extrajo e incluyó en BacTregulators).

(3) Base de Datos de Relaciones entre los “objetos”, proteínas, de BacTregulators.

La Figura 7 muestra la organización en tablas y relaciones entre ellas de la Base de Datos BacTregulators.

5.1. Descripción de la Base de Datos de Proteínas.

Se organiza alrededor de la **Tabla de Proteínas**. Dicha tabla se relaciona, directa o indirectamente, con el resto de las tablas. Así pues, la tabla de proteínas se relaciona directamente con las siguientes tablas maestras (relaciones de 1 a muchos):

- Tabla de Elementos de Genomas.
- Tabla de Familias.
- Tabla de Tipos de Proteínas.
- Tabla de Activador/Represor.
- Tabla de Códigos de Función.

Además, la tabla de Proteínas se relaciona directamente con las siguientes tablas “no maestras”:

- Tabla de Secuencias, de 1 a muchos, actuando la Tabla de Proteínas como maestra en esta relación (varias proteínas pueden tener una misma secuencia). A su vez esta tabla se relaciona con la de Promotores, de uno a muchos (una secuencia puede actuar sobre varios promotores).

- Tabla de Acciones, de 1 a muchos (una proteína puede tener muchas acciones).

- Tabla de Promotores, a través de una tabla de unión (explicita una relación de muchos a muchos). A su vez, está relacionada con la tabla de Elementos de Genoma (de uno a muchos, un genoma contiene muchos promotores), con la de Secuencias y con la

Acciones (también de uno a muchos, pues desde un promotor puede haber muchas acciones).

- Tabla de Promotores, a través de una tabla de unión (explicita una relación de muchos a muchos). A su vez, está relacionada con la tabla de Elementos de Genoma (de uno a muchos, un genoma contiene muchos promotores), con la de Secuencias y con la Acciones (también de uno a muchos, pues desde un promotor puede haber muchas acciones).

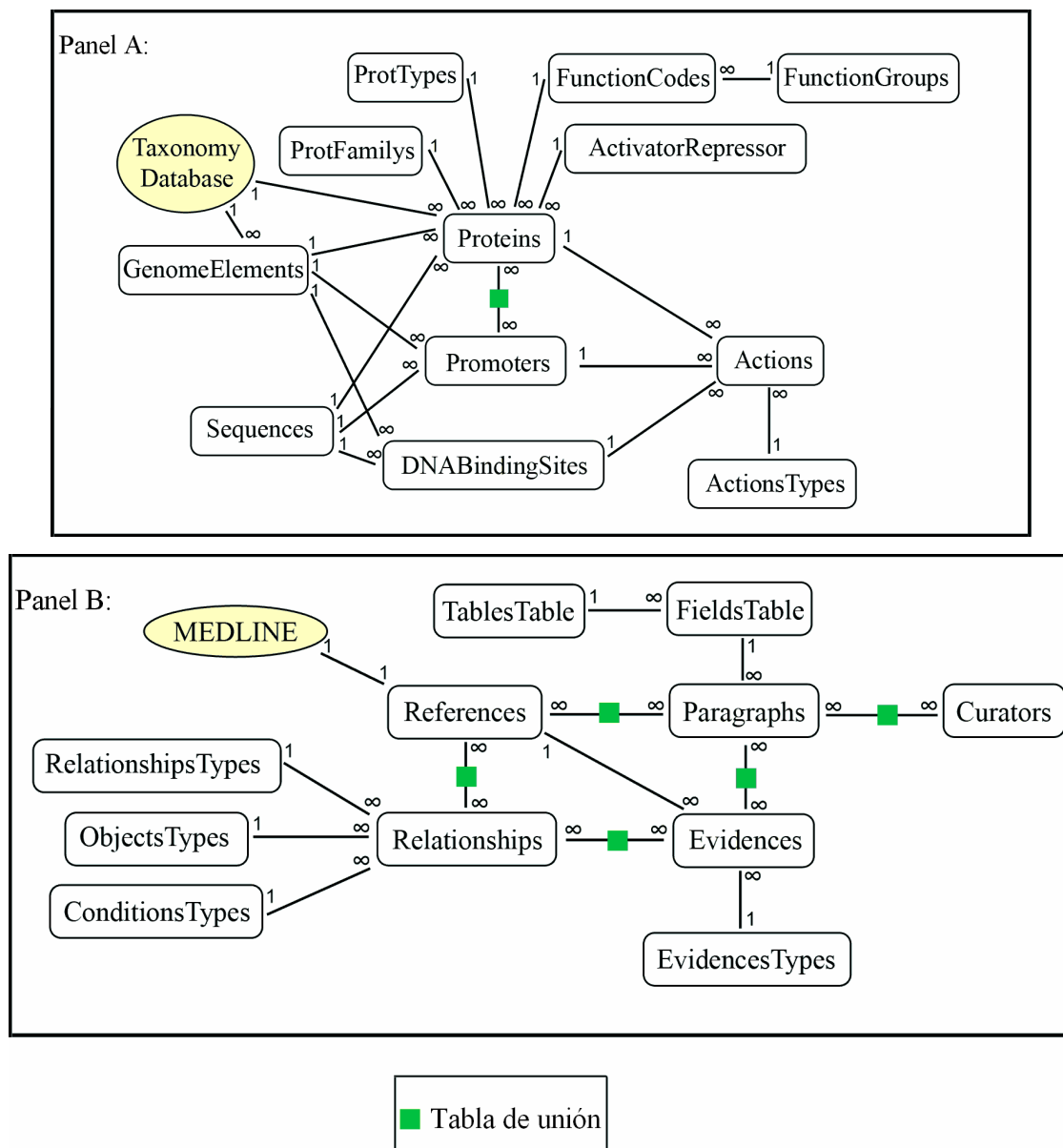


Figura 7: Estructura de la base de datos BacTregulators (Panel A: base de datos de proteínas. Panel B: base de datos de referencias bibliográficas).

La tabla de Proteínas se relaciona indirectamente con otras dos tablas: (i) Tabla de Tipos de Acciones, tabla maestra con una relación de 1 a muchos con la tabla de Acciones (un tipo de acción puede tener muchas acciones), y (ii) Tabla de Sitios de Unión al ADN, conectada también con la tabla de Acciones, por una relación de uno a muchos, con la tabla de Elementos de Genoma, por una relación de uno a muchos (un elemento de genoma contendrá multitud de sitios de unión al ADN) y con la Tabla de Secuencias por una relación de uno a muchos (una secuencia puede tener varios sitios de unión a ADN).

A continuación pasamos a describir cada una de las tablas mencionadas.

- Tabla de Proteínas (Tabla 1).

Constituye el eje de la base de datos y a partir de ella se construye la ficha de cada proteína que se le muestra al usuario. En su concepción inicial, muy ambiciosa, todas las proteínas citadas en BacTregulators tendrían el mismo tratamiento y por tanto, debería de contener los datos para construir la ficha no sólo de las proteínas reguladoras de la transcripción, objeto de esta base de datos, sino los de todas aquellas proteínas relacionadas con cada regulador, ya sean las proteínas cuya transcripción se regula, proteínas bajo el control de las reguladas, reguladores de la expresión de la proteína reguladora, etc. De todas formas en esta versión, hemos sido más realistas y nos centramos únicamente en los reguladores de la transcripción, incluyendo, eso sí, todos los elementos que nos permitan estudiar los circuitos y redes de regulación en cada caso. Podemos agrupar los datos en esta tabla en los siguientes tipos:

- (1) Datos de identificación de la proteína.
- (2) Datos de la secuencia del gen que codifica a la proteína y de la organización genética de la región en que se encuentra.
- (3) Características taxonómicas del organismo fuente de la proteína.
- (4) Datos de “conocimiento” acerca de la proteína.

Los datos cargados automáticamente difieren según la DB fuente, así por ejemplo en las proteínas procedentes de SWP TrEMBL, no se proporcionan datos acerca del elemento de genoma y posición absoluta del gen en el genoma, ni "startposition" ni "endposition", etc, tampoco se proporciona el COG.

(1) *Datos de identificación de la proteína* (fondo amarillo): Los campos del segundo al quinto y del noveno al undécimo guardan datos obtenidos automáticamente de las bases de datos “fuente”. El campo clave, el primero, es el “ProteinId” propio de BacTregulators, cada proteína incluida en ella queda definida por su identificador numé-

rico, de 1 a n. El segundo campo, "AccNumber", almacena el número de acceso de la proteína en su base de datos fuente. El noveno campo, "DBName", nos dice cuál es la base de datos fuente. Los campos "ShortName" y "LongName", guardan los nombres proporcionados por la ficha SWPROT, TrEMBL, o NCBI. El campo COG guarda el identificador COG, que nos proporciona la información referente a ortólogos, ("Clusters of Orthologous Groups of Proteins") (<http://www.ncbi.nlm.nih.gov/COG/>).

Tabla 1: Proteínas				
Campo	Tipo	Descripción	Ejemplo	
1	ProteinId	int uns.	Autonumeric	1
2	AccNumber	char(10) not null	Accession number	P34000
3	ShortName	char(15)		ACRR_ECOLI
4	GeneName	varchar(200)		
5	LongName	char(255)		Potential acrAB operon repressor
6	ProtTypeId	tinyint uns.	See ProteinsTypes master table	1 (Transcription factor)
7	FamilyId	tinyint uns.	See ProteinsFamyls master table	1 (TetR)
8	ActOrReprId	tinyint uns.	See ActivatorRepressor master table	3 (Activator and repressor)
9	DBName	Enum	Swiss-Prot, TrEMBL, NCBI, or Entry manually generated from literature data.	Swiss-Prot
10	COG	char(4)	COG (Code Ortholog Group) identifier	
11	FuncCodeId	tinyint uns.	See FunctionCodes master table	
12	StartPosition	int uns.	Absolute start position	
13	EndPosition	int uns.	Absolute end position	
14	Orientation	enum('+', '-')	Positive or negative orientation	
15	GeneUp	char(100)	Upstream gene	
16	OrientUp	enum('+', '-')	Upstream gene orientation	
17	GeneDown	char(100)	Downstream gene	
18	OrientDown	enum('+', '-')	Downstream gene orientation	
19	TaxId	Mediumint uns.	Use the NCBI taxonomic database	217992
20	GenomElemId	Mediumint uns.	see GenomeElements master table	
21	Sequenceld	int uns.	see Sequence master table	1
22	AutLastUpdate	Date		2003-02-20
23	ManLastUpdate	Date		2004-06-28
24	Function	Text	Overall view of the bacterial function in which the protein is involved	
25	PromotersText	Text	Promoters of the gene coding the protein (name, location and characteristics)	
26	RegulGenes	Text	Which genes regulate this protein and what relation with (activator or repressor)	
27	Sigma	Text		
28	RegulNet	Text	Regulation network in which the protein is involved	
29	Effectors	Text	Molecules and conditions (temperature, pH...) that trigger the protein	
30	Pathogen	Text	Involvement of the protein in pathogenicity	
31	Mutations	Text	Published data about mutations	
32	3DStructure	Text		
33	AssocFunc	Text	Functions associated to the different protein domains	
34	Oligomeriz	Text	Oligomerization	
35	Comments	Text		

Los campos sexto, séptimo y octavo incluyen datos almacenados en Tablas maestras. El sexto, “ProfTypeId”, de la tabla de Tipos de Proteínas (1 en caso de que la proteína sea un factor transcripcional), el séptimo, “FamilyId”, de la tabla de Familias de Proteínas (1 para TetR), y el octavo, “ActOrReprId”, de la tabla de “Activadores Represores” (3 si la proteína exhibe ambas acciones).

En el campo vigésimo primero guardamos el identificador de la secuencia que nos proporcionará la secuencia de aminoácidos de la proteína, almacenada en el campo de texto de la correspondiente tabla de Secuencias.

Si la Base de datos fuente es NCBI, en el campo undécimo guardamos el identificador del código de función del NCBI (Tabla 2) (el código funcional empleado por el NCBI para aquellas proteínas cuya función se conoce, consiste en una letra, existiendo un total de 25 letras para otras tantas funciones, que tratamos como carácter numérico. Además las 25 letras se agrupan en 4 categorías: “*Information storage and processing*”, “*Cellular processes and signaling*”, “*Metabolism*” y “*Poorly characterized*”).

Tabla 2: código de función del NCBI.

Código	Categoría	Descripción de la Función
J	1	Translation, ribosomal structure and biogenesis
A	1	RNA processing and modification
K	1	Transcription
L	1	Replication, recombination and repair
B	1	Chromatin structure and dynamics
D	2	Cell cycle control, cell division, chromosome partitioning
Y	2	Nuclear structure
V	2	Defense mechanisms
T	2	Signal transduction mechanisms
M	2	Cell wall/membrane/envelope biogenesis
N	2	Cell motility
Z	2	Cytoskeleton
W	2	Extracellular structures
U	2	Intracellular trafficking, secretion, and vesicular transport
O	2	Posttranslational modification, protein turnover, chaperones
C	3	Energy production and conversion
G	3	Carbohydrate transport and metabolism
E	3	Amino acid transport and metabolism
F	3	Nucleotide transport and metabolism
H	3	Coenzyme transport and metabolism
I	3	Lipid transport and metabolism
P	3	Inorganic ion transport and metabolism
Q	3	Secondary metabolites biosynthesis, transport and catabolism
R	4	General function prediction only
S	4	Function unknown

(2) *Datos de la secuencia del gen que codifica a la proteína y de la organización genética de la región en que se encuentra* (fondo anaranjado): El cuarto campo almacena el nombre del gen que codifica a la proteína de la ficha. Si la base de datos fuente es

el NCBI podemos almacenar en el campo vigésimo el Identificador de elemento de Genoma en que esté el gen, y en los campos duodécimo, decimotercero y decimocuarto, “StartPosition”, “EndPosition” y “Orientation” contienen las posiciones del primer y el último par de bases del gen, y su orientación, + o -, respectivamente. Los campos “GeneUp”, “OrientUp”, “GeneDown” y “OrientDown” proporcionan el nombre y orientación de los genes situados al extremo 5’ y 3’ del que nos ocupa, aún no están funcionando.

(3) *Características taxonómicas del organismo fuente de la proteína* (fondo azul): Si la Base de datos fuente es NCBI, el campo decimonoveno guarda el TaxId, usado por la base de datos de taxonomía del NCBI (“NCBI taxonomic database”).

(4) *Datos de “conocimiento” acerca de la proteína* (fondo verde): Se almacenan en los campos del vigésimo cuarto al trigésimo quinto. Se trata de campos virtuales que en realidad no existen físicamente en esta tabla (fondo verde en la Tabla 1), están definidos en la Tabla de campos, “FieldsTable”, y su contenido se extrae de la Tabla de Párrafos. En líneas generales, en este aspecto BacTregulators hereda a la anterior base de datos AraC/XylS. Se trata de campos de texto que almacenan el conocimiento extraído manualmente de las referencias bibliográficas. Como en AraC/XylS, y ya lo tratamos en el apartado correspondiente, en estos campos la información se organiza en párrafos referenciados y conectados a MEDLINE y como novedad, se indican las evidencias experimentales descritas en dichas referencias que apoyan lo dicho en cada párrafo.

En el campo “Function” incluimos los datos acerca de las funciones bacterianas en las que la proteína está implicada.

El campo “PromotorText” incluye la información referente a los promotores (nombre, posición y secuencia, características, etc).

En el campo “RegulGenes” guardamos la información sobre los genes cuya expresión se activa o reprime por la proteína reguladora, y las funciones en las que están implicadas las proteínas que codifican.

En el campo “Sigma” se almacena la información que se conozca sobre los factores sigma (cuál está implicado, si se establece contacto con el regulador, etc).

En el campo “RegulNet” esquematizamos la red de regulación en la que está implicada la proteína.

El campo “Pathogen” muestra los datos acerca de implicación de la proteína en la posible patogenicidad del microorganismo.

En el campo “Mutations” almacena la información sobre las mutaciones que encontramos en las referencias bibliográficas.

Los 3 siguientes campos almacenan información estructural. En “3DStructure” guardamos la descripción de la estructura de aquellas proteínas en que se haya resuelto. Un campo afín al anterior es el “AssocFunc” que indica las funciones asociadas a cada uno de los dominios de la proteína. El siguiente campo “Oligomeriz” nos indica si la proteína se presenta en dímeros, tetrámeros, dímeros de dímeros, etc.

Un campo a modo de cajón de sastre para la información que no se ajustara a los campos anteriores es “Comment”.

- Tabla de Secuencias (Tabla 3).

La secuencia de aminoácidos de la proteína se almacena en el segundo campo de la tabla, “Sequence”. Dicha secuencia se descarga de la base de datos fuente para cada caso. En el tercer campo almacenamos el “CRC64” (código de comprobación de errores). El cuarto campo, “Score”, nos indica la puntuación con la que esta secuencia se incluye en su familia.

La Tabla de Secuencias contiene en los tres últimos campos la información necesaria para reconstruir el "mejor" alineamiento secuencia/consensus. Así, en el quinto campo se guarda la posición en el consensus de inicio del alineamiento. En el sexto campo, la posición en la secuencia de inicio del alineamiento. En el séptimo campo se guarda la secuencia y posiciones de los “gaps” en dicho alineamiento.

- Tabla de Elementos de Genomas (Maestra) (Tabla 4).

Gracias a esta tabla, BacTregulators es también una base de datos genómica y no sólo de proteínas. Los elementos de genoma se entienden como cromosómicos y extra-cromosómicos.

Dicha tabla contiene 7 campos. El Identificador de Elemento de Genoma, “GenomeElemId”, propio de BacTregulators, constituye el campo clave. El segundo campo, “GenomeElement”, almacena el nombre de cada elemento. En el tercer campo guardamos el identificador del genoma en el NCBI. En el cuarto campo, el identificador de genoma en Entrez. En el último campo guardamos el identificador de la base de datos de taxonomía del NCBI.

- Tabla de Familias de Proteínas (Maestra) (Tabla 5).

Contiene los datos correspondientes a las familias de proteínas almacenadas en BacTregulators. En este momento BacTregulators contiene 3 familias de reguladores transcripcionales: AraC/XylS, TetR e IclR.

Esta tabla se organiza en 9 campos, los 3 primeros contienen los datos de denominación de la familia y los restantes los datos referentes a los umbrales empleados en la definición de la familia.

Como siempre, el primer campo, el Identificador numérico de la familia es el campo clave de la tabla, los dos siguientes almacenan el nombre de la familia (corto y largo, respectivamente). Los campos cuarto y quinto almacenan los valores de los umbrales 0 (superior) y 1 (inferior) dados por el programa "pfmake". El sexto campo guarda el valor del umbral que nosotros empleamos tras haber realizado el proceso de validación de la familia. En el séptimo campo, "InterProAc" guardamos el número de acceso de la familia en la base de datos Interpro. En el campo "MethodAc", se almacena el número de acceso de la herramienta empleada para definir a la familia (HMM de Pfam, perfil de Prosite, etc), en el ejemplo este campo se indica "NULL" porque la familia TetR se define con el perfil que hemos diseñado y por tanto carece aún de número de acceso.

- Tabla de Tipos de Proteínas (Maestra) (Tabla 6).

Aunque BacTregulators sea una base de datos de proteínas reguladoras de la transcripción, su filosofía, tal como ya anticipamos, es más amplia, pretendiendo ser una base de datos de proteínas en general (todas aquéllas implicadas en la función de los factores de transcripción), de ahí la necesidad de esta tabla, que con solo dos campos nos indica el tipo de proteína de que se trata ya sea factor de transcripción, proteína de membrana, toxina, etc. En el ejemplo de la estructura de la Tabla 6 se muestra el ejemplo del tipo 1, factor de transcripción.

5. Tabla de ActivadorRepresor (Maestra) (Tabla 7).

En esta tabla incluimos el dato de tipo de actividad de un regulador de la transcripción, que puede ser activador (id. 1), represor (id. 2) o ambos (id. 3). La Tabla 7 muestra la tabla para el registro tipo 1, activador.

- Tabla de Acciones (Tabla 8).

Se refiere a la acción sobre el promotor de una proteína mediante su acción sobre su sitio de unión en el ADN. El primer campo, "ActionId", es la clave. El segundo campo contiene el "ProteínId" de la proteína que actúa (Tabla de proteínas). El tercer campo, el identificador del sitio de unión al ADN al que se une dicha proteína para efectuar su acción (Tabla de sitios de unión al ADN). El cuarto campo contiene el identificador del promotor afectado por la unión de la proteína a su sitio (Tabla de promotores). Finalmente, el quinto campo contiene el identificador del tipo de acción, represor, activador o ambos de la Tabla maestra de Tipos de Acciones, con la que esta tabla está relacionada (Figura 7).

- Tabla de Promotores (Tabla 9) (no activada).

Esta tabla almacena los datos sobre las posiciones propias de un promotor. Para ello empleamos 11 campos.

Los tres primeros hacen referencia al identificador del promotor (campo clave), nombre del promotor, e identificador del elemento de genoma en el que se encuentra, respectivamente.

El resto de los campos guardan los datos de posición en el ADN. En el cuarto, la posición absoluta, en el elemento de genoma, del TSP del primer gen transcrito a partir de ese promotor. En el quinto, la posición del codón de inicio respecto del TSP. En el sexto, almacenamos la posición absoluta (respecto al elemento de genoma), del codón de inicio. Los 4 campos siguientes, del séptimo al décimo, guardan respectivamente: posición relativa de la base de la región -10 más próxima al TSP, posición relativa de la base de la región -10 más alejada del TSP, e igual para la región -35 con los dos siguientes campos.

Por último, el campo "SequenceId" señala a la Tabla de Secuencias (que almacenará todas las secuencias, de aminoácidos y de ADN).

- Tabla de Tipos de Acciones (Maestra) (Tabla 10).

En esta tabla almacenamos con un identificador numérico, en el primer campo (clave) el tipo de acción, que almacenamos en el segundo campo (represión, activación, "enhacer").

- Tabla de Sitios de Unión al ADN (Tabla 11).

En esta tabla almacenamos los datos referentes a las secuencias de ADN a las que se unen las proteínas reguladoras.

En ella se agrupan los datos tales como “identificador del sitio de unión” (primer campo), nombre del sitio de unión, por ejemplo *tetO* (segundo campo), con el identificador del elemento de genoma en que se encuentra dicho sitio de unión, cuando esté secuenciado (tercer campo), y los datos de la posición de inicio y de fin de cada sitio (campos cuarto y quinto respectivamente), referidas al TSP. Por último, el sexto campo guarda un identificador de secuencia del sitio de unión (recordemos que teóricamente la tabla de secuencias guarda tanto proteínas como ADN).

Tabla 3: Secuencias			
Campo	Tipo	Descripción	Ejemplo
1	SequenceId	Int unsigned not null	Automumeric
2	Sequence	text not null	
3	CRC64	char(16)	
4	Store	float(6,3)	Value obtained by the profile, signature or HMM that defines the family 737.00
5	ConsLowLimit	smallint unsigned	Pos. de inicio en Cons.
6	SeqLowLimit	smallint unsigned	Pos. de inicio en Sec.
7	Gaps	varchar(200)	Número y pos. de gaps

Tabla 4: Elementos de Genoma			
Campo	Tipo	Descripción	Ejemplo
1	GenomElemId	int uns.	Autonumeric 3
2	GenomeElement	varchar(100)	Aeropyrum pernix
3	NCBI_Id	char(9)	Identifier in NCBI NC_000854
4	Entrez_Id	smallint uns.	GI:47118298
5	Superkindom	enum('Bacteria','Archaea'),	
6	GenomaSize	int uns.	1669695
7	GenomTaxId	mediumint uns.	272557

Tabla 5: Familias de Proteínas			
Campo	Tipo	Descripción	Ejemplo
1	FamilyId	tinyint uns.	Autonumeric 1
2	FamShortName	char(20)	TetR
3	FamLongName	char(250)	Tetracycline resistant
4	CutOff0	float(7,4)	8.5
5	CutOff1	float(7,4)	6.5
6	CutOffAutoValid	float(7,4)	10.5
7	InterProAc	char(9)	IPR001647
8	MethodAc	char(9)	NULL
9	Consensus	varchar(100)	Consensus sequence REQILDAALELFAER...

Tabla 6: Tipos de Proteínas			
Campo	Tipo	Descripción	Ejemplo
1	ProtTypeId	tinyint uns.	Autonumeric 1
2	ProtType	char(30)	Transcription factor, Membrane prot., Enz.,...

Tabla 7: ActivadorRepresor			
Campo	Tipo	Descripción	Ejemplo
1	ActOrReprId	tinyint uns.	Autonumeric 1
2	ActOrRepr	char(30)	Activator, repressor or both Activator

Tabla 8: Acciones			
Campo	Tipo	Descripción	Ejemplo
1	🔑 ActionId	int uns.	Autonumeric
2	ProteinId	int uns.	
3	BindingSiteId	int uns.	
4	PromoterId	int uns.	
5	ActionTypeId	Smallint uns.	

Tabla 9: Promotores			
Campo	Tipo	Descripción	Ejemplo
1	🔑 PromoterId	mediumint uns.	Autonumeric
2	PromName	char(20)	
3	GenomElemId	mediumint uns.	Genome element identifier
4	TSP	int uns.	Transcription start point del primer gen que se transcribe a partir de ese gen (pos. absoluta en genoma)
5	CodonRelPos	mediumint.	TSP relative position of initiation codon
6	CodonAbsPos	int uns.	Absolute position of initiation codon
7	Near10	smallint	TSP relative position of promoter sigma -10 region nearest to TSP
8	Far10	smallint	TSP relative position of promoter sigma -10 region farrest to TSP
9	Near35	smallint	TSP relative position of promoter sigma -35 region nearest to TSP
10	Far35	smallint	TSP relative position of promoter sigma -35 region farrest to TSP
11	SequenceId	int uns.	

Tabla 10: Tipos de Acciones			
Campo	Tipo	Descripción	Ejemplo
1	🔑 ActionTypeId	tinyint uns.	Autonumeric
2	ActionType	char(20)	Repression, activation, enhance
			Activación

Tabla 11: Sitios de unión a ADN			
Campo	Tipo	Descripción	Ejemplo
1	🔑 BindingSiteId	int uns.	Autonumeric
2	BSName	char(30)	Binding site name
3	GenomElemId	mediumint uns.	Genome element identifier
4	StartSite	Smallint	TSP relative binding site start position nearest of first gene TSP that is transcribed
5	EndSite	Smallint	idem con end position
6	SequenceId	int uns.	

Tabla 12: ProteínaPromotor			
	Campo	Tipo	
1	PRI	IdProtein	int uns.
2	PMI	IdPromoter	int uns.
3	PRO	ProteinOrder	Tinyint

Finalmente, señalar que la relación entre las tablas de Proteínas y de Promotores, de muchos a muchos, se realiza mediante la Tabla de unión de “ProteínaPromotor” (Tabla 12). En esta tabla se señala qué proteínas se transcriben desde cada promotor. Contiene los campos clave de las dos tablas relacionadas, “IdProtein” e “IdPromoter” y el tercer campo “ProteinOrder” que indica la posición que ocupa el gen que expresa a la proteína desde ese promotor.

5.2. Descripción de la Base de Datos de Referencias Bibliográficas.

Las tablas y relaciones que conforman esta parte de la base de datos se esquematiza en la Figura 7 panel B. El núcleo lo constituye la Tabla de Párrafos (recordemos que el párrafo es la unidad de conocimiento), que se relaciona directamente con 4 tablas, las de Referencias, Evidencias, "Curators" y Campos. Con las 3 primeras tablas la relación es de muchos a muchos mediante las correspondientes tablas de unión (PárrafoReferencia, PárrafoEvidencia y CuratorPárrafo, respectivamente). La relación con la Tabla de Campos es de uno a muchos.

Para no extendernos excesivamente en la explicación escrita de la estructura plasmada en la Figura 7 panel B, pasamos a describir las diferentes tablas.

- Tabla de Párrafos (Tabla 13).

Su función es la de guardar el contenido de los campos de tipo P (ver tabla de campos), es decir los de tipo párrafo.

La Tabla de Párrafos contiene 5 campos por registro, cuatro numéricos y uno de texto. El campo clave como es natural consiste en un identificador de párrafos ("ParagraphId"). El segundo campo "RegisterID". El tercer campo "FieldId" (ver tabla maestra "FieldsTable"), nos indica el identificador del campo "virtual" al que se liga el párrafo de cada registro de la tabla de párrafos. El campo "ParagraphOrder" indica el orden en el que se coloca este párrafo en el campo correspondiente respecto a otros párrafos que también pudiera contener dicho campo. Por último en "ParagraphText" guardamos el texto del párrafo.

- Tabla de Referencias (Tabla 14).

Almacena de los datos correspondientes a cada referencia bibliográfica usada en BacTregulators. Podríamos considerarla en sí misma una base de datos de referencias.

Como es de esperar el campo clave es el identificador de referencia "ReferenceId" que se empleará para relacionarla con párrafos y evidencias. Su organización de campos no requiere explicación.

- Tabla de Evidencias (Tabla 15).

Esta tabla identifica cada evidencia ("EvidenceId"), la asigna a un tipo de evidencia ("EvidTypeId" de la tabla maestra de Tipos de Evidencias), guarda su texto ("EvidenceText") y la liga a la referencia de la que se extrajo ("ReferenceId" de la Tabla de Referencias).

- Tabla de Tipos de Evidencias (Maestra; no activada) (Tabla 16).

Es una tabla de solo dos campos que asigna cada evidencia a un tipo de evidencia, pero aún no se han definido los tipos de evidencias.

Los datos referentes a los mantenedores ("Curators") de la base de datos, se guardan como lista en lugar de tabla. La lista contiene un identificador, el nombre y correo electrónico.

Las relaciones de muchos a muchos que se establece entre las tablas de la Base de Datos de Referencias se explicitan mediante las tres tablas de unión, PárrafoReferencia, PárrafoEvidencia y CuratorPárrafo, cuya estructura mostramos en las Tablas 17, 18 y 19.

Tabla 13: Párrafos

Tabla 13: Párrafos				
Campo		Tipo	Descripción	Ejemplo
1	🔑 ParagraphId	int uns.	Autonumeric	
2	RegisterId	int uns.	Indicate the value of identifier register	
3	FieldId	smallint uns.	See FieldsTable master table	
4	ParagraphOrder	mediumint uns.		
5	ParagraphText	Text		

Tabla 14: Tabla de Referencias

Tabla 14: Tabla de Referencias				
Campo		Tipo	Descripción	Ejemplo
1	🔑 ReferenceId	int uns.	Autonumeric	
2	Title	char(250)		
3	Authors	char(250)		
4	Journal	char(30)		
5	Year	Year		
6	Volumen	Smallint uns.		
7	FirstPage	Smallint uns.		
8	LastPage	Smallint uns.		
9	MedlineId	int uns.		
10	InputDate	Date		

Tabla 15: Evidencias

Tabla 15: Evidencias				
Campo		Tipo	Descripción	Ejemplo
1	🔑 EvidenceId	int uns.	Autonumérico	
2	EvidTypeId	smallint uns.	Ver la tabla maestra de Tipos de Evidencias	
3	EvidenceText	Text	Texto de la evidencia	
4	ReferenceId	int uns.	Ver la Tabla de Referencias	

Tabla 16: Tipos de Evidencias

Tabla 16: Tipos de Evidencias				
Campo		Tipo	Descripción	Ejemplo
1	🔑 EvidTypeId	smallint uns.	Autonumérico	
2	EvidenceType	char		

Tabla 17: PárrafoEvidencia

Tabla 17: PárrafoEvidencia			
Campo		Tipo	
1	PGI	🔑 ParagraphId	int uns.
2	EVI	🔑 EvidenceId	int uns.

Tabla 18: PárrafoReferencia

Tabla 18: PárrafoReferencia			
Campo		Tipo	
1	PGI	🔑 ParagraphId	int uns.
2	RFI	🔑 ReferenceId	int uns.

Tabla 19: CuratorPárrafo

Tabla 19: CuratorPárrafo			
Campo		Tipo	
1	CUI	🔑 CuratorId	int uns.
2	PGI	🔑 ParagraphId	int uns.

3. Descripción de la Base de Datos de Relaciones.

Otro tipo de “datos” almacenados en BacTregulators es el de las “Relaciones”. En la versión que nos ocupa, esta área está aún desactivada, esperando a próximas actualizaciones, pero por su carácter innovador, así como por su utilidad práctica, consideramos adecuado describirla brevemente.

Está constituida por 4 tablas: Relaciones, Tipos de Relaciones, Tipos de Objetos, Tipos de Condiciones.

Por la naturaleza abstracta de los términos usados, “Objetos”, “Relaciones” y “Condiciones”, más propios de Computación, creemos interesante usar un ejemplo práctico concreto: el objeto A es la proteína represora TetR y el objeto B la proteína TetA (bomba de extrusión de tetraciclina) este caso la relación representaría la unión ambos objetos, es decir, la de TetR a TetA. La condición en la que la relación se verifica en este ejemplo sería ausencia del inductor tetraciclina. De esta forma el registro de este ejemplo almacenaría en la Tabla de Relaciones la siguiente relación: - “El *objeto A* (TetR) *reprime* la expresión (relación) del *objeto B* (TetA) mientras permanezca *la condición X* (ausencia de tetraciclina)”.

Existiría una interfaz para las relaciones en la entrada BacTregulators que se le muestra al usuario las relaciones en las que interviene la proteína objeto de esa entrada. Su formato aún está en fase de diseño. Para explicitar este tipo de datos, las Relaciones, hemos construido las cuatro tablas anteriormente mencionadas y que pasamos a describir.

- Tabla de Relaciones (Tabla 20, los Id y ejemplos son hipotéticos).

Explicitará las relaciones entre pares de objetos de BacTregulators. Esta aplicación permitirá al usuario interrogar a la base de datos sobre este tipo de cuestiones.

- Tabla de Tipos de Relaciones (Maestra) (Tabla 21).

Almacenaría las diferentes clases de relaciones empleadas en la tabla anterior.

- Tabla de Tipos de Objetos (Maestra) (Tabla 22).

Almacena las clases de objetos empleados en las relaciones en BacTregulators. El tipo de objeto se almacena en el campo “ObjectsTypes”. Dicho campo puede tomar los siguientes valores: factor activador de la transcripción, factor represor de la trans-

cripción, factor sigma, enzimas biosintéticas, enzimas degradativas, proteína de membrana, proteína de bomba.

- Tabla de Condiciones (Maestra) (Tabla 23).

Almacenaría las condiciones en las que se establecen las relaciones entre los objetos.

La conexión de la Base de Datos de Relaciones con BacTregulators se realizaría a través de la base de datos de Referencias, en concreto mediante las dos Tablas de Unión “RelaciónReferencia” y “RelaciónEvidencia”, (Tablas 24 y 25) que explicitan la conexión de muchos a muchos que se establece entre la Tabla de Relaciones y las de Referencias y Evidencias (una Relación puede aparecer en muchas Referencias y viceversa e idem con la Evidencias).

5.4. Conexión entre las tres Bases de Datos de BacTregulators.

El andamiaje que conecta a las distintas Bases de Datos de BacTregulators lo constituyen la Tabla de Campos y la Tabla de Tablas, que pasamos a describir.

- Tabla de Campos (Maestra) (Tabla 26).

En esta tabla organizamos como registros a todos los campos que conforman a BacTregulators. En cada registro los dos primeros campos son de identificador del registro (“campo”) que estemos tratando. Así, el primero, clave, es el identificador de campo "FieldId" (con el que se liga a la Tabla de párrafos), y el segundo, "FieldName" nos da el nombre del registro (“campo”). Los dos siguientes campos dan información acerca de la tabla a la que pertenece el registro (“campo”) y del orden que ocupa en la entrada BacTregulators, "TableId" y "FieldOrder". El quinto campo "FieldType" nos indica mediante un carácter el tipo de registro (campo) (N: Normal, E: Enum type, T: Text, P: Paragraph). El sexto campo "FieldDescript" guarda la denominación con la que este registro (campo) se muestra al usuario en la página html.

La utilidad de esta tabla, aparte de proporcionar la cohesión interna a BacTregulators, es que almacena las instrucciones para construir el Formulario o la Entrada de BacTregulators que se facilita al usuario (dice de qué tabla tomar el contenido de cada campo y en que orden éste se coloca, así como el tipo de campo y el título con que se muestra al usuario dicha información en la página html).

- Tabla de Tablas (Maestra) (Tabla 27).

Sirve para organizar las tablas que constituyen el cuerpo de BacTregulators. Se organiza en tres campos (Tabla 27), el primero, clave, "TableId" guarda el identificador de la tabla, el segundo "TableName", su nombre, y el tercero "TableDescript", una breve descripción de dicha tabla.

Como habíamos anticipado, la unión entre las diferentes áreas de BacTregulators, Base de Datos de Secuencias, de Referencias y de Relaciones, se logra mediante la Tabla de Campos y la Tabla de Tablas, ambas relacionadas por el campo "TableId".

Así pues, la clave (en sentido arquitectónico) de BacTregulators es la Tabla de Campos, que se relaciona con la Base de Datos de Referencias mediante la Tabla de Párrafos por el campo "FieldId" común a ambas tablas. Por otro lado, el tercer campo de la Tabla de Campos, el campo "TableId", la conecta con la Tabla de Tablas y con todas las tablas de BacTregulators, cada una definida por su primer campo "TableId".

Tabla 20: Relaciones			
Campo	Tipo	Descripción	Ejemplo
1	RelationId	int uns.	Autonumérico
2	ObjAId	int uns.	Ver la tabla maestra de Objetos
3	ObjATypeId	int uns.	Ver la tabla maestra de Tipos de Objetos
4	ConditionId	int uns.	Ver la tabla maestra de Condiciones
5	RelTypeId	int uns.	Ver la tabla maestra de Tipos de Relaciones
6	ObjBId	int uns.	Ver la tabla maestra de Objetos
7	ObjBTypeId	int uns.	Ver la tabla maestra de Tipos de Objetos

Tabla 21: Tipos de Relaciones			
Campo	Tipo	Descripción	Ejemplo
1	RelTypeId	smallint uns.	Autonumérico
2	RelType	Char	Unión

Tabla 22: Tipos de Objetos			
Campo	Tipo	Descripción	Ejemplo
1	ObjTypeId	smallint uns.	Autonumérico
2	ObjectType	char(30)	Activador transcripcional, represor transcripcional, factor sigma, proteína de bomba, ...

Tabla 23: Condiciones			
Campo	Tipo	Descripción	Ejemplo
1	ConditionId	smallint uns.	Autonumérico
2	Condition	char(30)	Presencia o de un inductor, temperatura, pH, aerobiosis o anaerobiosis, ...

Tabla 24: RelaciónReferencia		
Campo	Tipo	
1	RelationId	int uns.
2	ReferenceId	int uns.

Tabla 25: RelaciónEvidencia		
Campo	Tipo	
1	RelationId	int uns.
2	EvidenceId	int uns.

Tabla 26: Tabla de Campos



Campo	Tipo	Descripción	Ejemplo
1  FieldId	smallint uns.	Autonumeric	1
2 FieldName	char(20)		ProtTypeId
4 TableId	tinyint uns.		1 (Proteins)
5 FieldOrder	tinyint uns.	Ordinal con el que el campo aparece en la Entrada	5
5 FieldType	Enum	N: Normal, E: Enum, T: Texto, P: Párrafo	N
6 FieldDe-script	char(100)	Título del campo en la Entrada	“Protein type”
7 Extra	char(50)	Stores the master table name if exists (and suffix for ambiguous fields if necessary) or enum type values	ProtTypes

Tabla 27: Tabla de Tablas

Campo	Tipo	Descripción	Ejemplo
1  TableId	smallint uns.	Autonumérico	2
2 TableName	char(20)		DNABindingSites
3 TableDescript	char(100)		Tabla de los sitios de unión a ADN

Capítulo 2:

La familia TetR de reguladores transcripcionales.

6.1. Resumen

Se ha desarrollado un *perfil* para las proteínas de la familia TetR de represores. La secuencia peptídica que define al *perfil* está constituido por 47 aminoácidos que corresponden al motivo hélice-giro-hélice de unión a ADN y zonas adyacentes en la estructura tridimensional de TetR, QacR, CprB y EthR.

Se ha establecido un conjunto no redundante de 2353 proteínas pertenecientes a esta familia mediante el análisis de las bases de datos de genomas y proteínas con el *perfil* desarrollado para TetR. Hemos encontrado miembros de la familia TetR en 85 géneros de bacterias gram-positivas, α , β y γ -proteobacterias, cianobacterias y arqueas. El conjunto de genes que ellas regulan se conoce sólo para 85 de los 2353 miembros de la familia. Estas proteínas están implicadas en el control de la transcripción de bombas de eflujo de drogas con distintas estructuras, rutas de biosíntesis de antibióticos, respuesta a estrés osmótico y compuestos tóxicos, control de rutas catabólicas, procesos de diferenciación y patogenicidad. La red de regulación en la que cada miembro de la familia está implicado, puede ser tan simple como en TetR (TetR unido a su operador reprime su propia transcripción y la de *tetA* y se libera del operador en presencia de tetraciclina), o más compleja implicando cascadas reguladoras sucesivas, en las que un regulador miembro de la familia TetR está regulado por otro regulador o el miembro de la familia dispara una respuesta celular frente a estímulos ambientales.

Del análisis de los dos co-cristales de TetR y QacR con sus respectivos operadores, de sus estructuras tridimensionales con y sin ligandos, y de los análisis de los alineamientos múltiples del fragmento de 47 residuos de los 2353 miembros de la familia TetR, se han identificado dos grupos de residuos. El primero incluye posiciones muy conservadas implicadas en el mantenimiento de la arquitectura y la orientación adecuada del dominio hélice-giro-hélice, por lo que tendrían un papel estructural. En el segundo grupo se han identificado residuos menos conservados, implicados en el contacto con el esqueleto de fosfato y las bases diana del operador.

La información pertinente relacionada con la familia de reguladores TetR, se recoge en la base de datos "bacTregulators", accesible en "www.bactregulators.org".

6.2. Introducción

Las bacterias en el medio ambiente están expuestas a variaciones de temperatura, nutrientes, disponibilidad de agua y a la presencia de moléculas tóxicas originadas por los elementos bióticos o abióticos de su entorno (incluyendo moléculas deletéreas originadas por sus propios metabolismos). Estos cambios pueden hacerlas vivir en con-

diciones lejos de las óptimas. Sobrevivir en estos medios cambiantes requiere un amplio espectro de respuestas adaptativas rápidas, las cuales suelen estar controladas mediante proteínas reguladoras. Los reguladores transcripcionales pueden responder tanto a señales específicas ambientales como celulares y modular la transcripción, traducción o algunos otros eventos en la expresión genética, de manera que las respuestas fisiológicas se modifican apropiadamente (Beinlich *et al.*, 2001; Cases *et al.*, 2003; Cohen *et al.*, 1993; Fromknecht *et al.*, 2003; Fuqua *et al.*, 1996; Henikoff *et al.*, 1988; McKay *et al.*, 1981; Ramos *et al.*, 2002; Ramos *et al.*, 2001; Roessner *et al.*, 2002; Rojo, 1999; Surette *et al.*, 1998; Tropel y van der Meer, 2004; Ventre *et al.*, 2004; Winans y Bassler, 2002).

Los reguladores de la transcripción de procariotas se clasifican en familias atendiendo a similitudes de secuencia, estructurales y criterios funcionales (Busenlehner *et al.*, 2003; Egan, 2002; Fuqua *et al.*, 1994; Gallegos *et al.*, 1993; Grkovic *et al.*, 2001, A; Grkovic *et al.*, 2001, B; Haydon y Guest, 1991; Henikoff *et al.*, 1988; Henikoff *et al.*, 1990; Martínez-Hackert y Stock, 1997; North *et al.*, 1993; Poole *et al.*, 1996; Ramos *et al.*, 1990; Rigali *et al.*, 2002; Schell, 1993; Spiro, 1994; Summers, 1992). La Tabla 28 incluye el listado de las familias más importantes de reguladores microbianos de la transcripción, el tipo de motivo de unión a ADN que exhiben, si sus miembros son preferentemente activadores o represores o si muestran una acción alternativa.

En este capítulo de esta Tesis Doctoral se analiza la familia TetR, una de las familias de reguladores de la transcripción mejor representada y ampliamente distribuida en bacterias, con un dominio de unión a ADN de tipo hélice-giro-hélice (Levy, 1988; Levy *et al.*, 1999; Méndez *et al.*, 1980; Orth *et al.*, 2000).

Utilizando el *perfil* que se describe más adelante identificamos miembros de la familia de represores TetR en SWISS-PROT y TrEMBL y en todas las proteínas traducidas a partir de las secuencias de los genomas procariotas secuenciados. Tras la compilación de los resultados obtenidos de las bases de datos de proteínas y de ácidos nucleicos, se encontró que la familia de reguladores TetR incluía 2353 secuencias no redundantes (Diciembre, 2004). La función específica regulada por miembros de la familia se conoce solamente en unos 85 casos (Tabla 29). Estas proteínas controlan genes cuyos productos están implicados en resistencia a múltiples drogas, enzimas implicadas en diferentes rutas catabólicas, biosíntesis de antibióticos, estrés osmótico y patogenicidad de bacterias tanto gram negativas como gram positivas (Tabla 29). La información más relevante y referente a estas proteínas se ha recopilado en la base de datos disponible en <http://www.bactregulators.org> (Martínez-Bueno *et al.*, 2004; ver también el Capítulo 1).

Tabla 28. Familias de reguladores de la transcripción en procariotas.

Familia	Acción	Funciones reguladas	DBD "motif"	Posición	Referencias
LysR	Activador/Represor	Metabolismo del carbono y del nitrógeno	HTH	N-terminal	Henikoff et al., 1988; Schell, 1993.
AraC/XylS	Activador	Metabolismo nitrogenado, respuesta a estrés y patógenesis	HTH	C-terminal	Gallegos et al., 1997; Tobes and Ramos, 2002.
TetR	Represor	Biosíntesis de antibióticos, bombas de eflujo, estrés osmótico, etc.	HTH	N-terminal	Aramaki et al., 1993; Aramaki et al., 1995a; Aramaki et al., 1995b.
LuxR	Activador	"Quorum sensing", biosíntesis y metabolismo, etc	HTH	C-terminal	Fuqua et al., 1994; Pao et al., 1994; Redfield, 2002.
LacI	Represor	Utilización de fuente de carbono	HTH	N-terminal	Chandler, 1992; Weickert and Adhya, 1992.
ArsR	Represor	Resistencia a metales	HTH	Central	Busenlehner et al., 2003; Wu and Rosen, 1991.
IclR	Represor/Activador	Metabolismo del carbono, bombas de eflujo	HTH	N-terminal	Neal and Chater, 1987; Reizer et al., 1996; Reverchon et al., 1991; Sunnarborg et al., 1990; Molina-Henares et al., 2006.
MerR	Represor	Resistencia y detoxificación	HTH	N-terminal	Helmann et al., 1990; Summers, 1992.
AsnC	Activador/Represor	Biosíntesis de aminoácidos	HTH	N-terminal	Friedberg et al., 2001.
MarR	Activador/Represor	Resistencia múltiple a antibióticos	HTH	Central	Alekshun and Levy, 1999; Ariza et al., 1994; Seoane and Levy, 1995; Sulavik et al., 1995.
NtrC(EBP)	Activador	Asimilación de nitrógeno, síntesis de aminoácidos aromáticos, flagelos, rutas catabólicas, respuesta a fagos, etc.	HTH	C-terminal	Kustu et al., 1991; Morett and Segovia, 1993.
OmpR	Activador	Metales pesados y virulencia ("Response regulator of a two-component system")	Winged helix	C-terminal	Martinez-Hackert and Stock, 1997.
DeoR	Represor	Metabolismo glucídico	HTH	N-terminal	Ramos et al., 2002; van Rooijen and de Vos, 1990; Zeng et al., 1996.
Cold shock	Activador	Resistencia a bajas temperaturas	Dominio de unión a ARN	Variable	Brandt et al., 1994; La Teana et al., 1991; Schindelin et al., 1993.
GntR	Represor	Metabolismo general	HTH	N-terminal	Haydon and Guest, 1991; Reizer et al., 1991; Rigali et al., 2002.
Crp	Activador/Represor	Respuesta global, represión catabólica y anaerobiosis	HTH	C-terminal	Chandler, 1992; Gambino et al., 1993; Ma et al., 1995.

Tabla 29: Funciones específicas reguladas por los miembros de la familia TetR de represores.

N	ID SPTR	Nombre	Organismo	Función	G	REF.
1	P34000	AcrR	<i>Escherichia coli</i>	AcrR reprime la expresión del operón <i>acrAB</i> que confiere multiresistencia a drogas y probablemente también controla el gen <i>micF</i> .	1	Fralick, 1996; Jellen-Ritter and Kern, 2001; Ma <i>et al.</i> , 1996; Ma <i>et al.</i> , 1995; Rand, <i>et al.</i> , 2002; Rodionov <i>et al.</i> , 2001; White <i>et al.</i> , 1995.
2	Q53901	ActII	<i>Streptomyces coelicolor</i>	Situado en el agrupamiento <i>act</i> que contiene genes reguladores y para la exportación de antibióticos.	1	Caballero <i>et al.</i> , 1991; Fernández-Moreno <i>et al.</i> , 1991.
3	Q9F8V9	AmeR	<i>Agrobacterium tumefaciens</i>	Regula negativamente al operón <i>ameABC</i> que codifica proteínas similares a las de los sistemas de eflujo tipo RND (Nodulation-Cell Division).	1	Peng and Nester, 2001.
4	Q9RG61	AmrR	<i>Pseudomonas aeruginosa</i>	Probablemente regula los genes <i>amrAB</i> que codifican un sistema de eflujo implicado en el fenotipo de impermeabilidad a aminoglicósidos de <i>Pseudomonas aeruginosa</i> .	1	Westbrock-Wadman <i>et al.</i> , 1999.
5	Q9KJC4	ArpR	<i>Pseudomonas putida</i> S12	Parece ser un represor de la expresión del operón <i>arpABC</i> . ArpABC en <i>Pseudomonas putida</i> S12 está sólo implicada en la multiresistencia a drogas pero no en la tolerancia a disolventes orgánicos.	1	Kieboom and de Bont, 2001.
6	Q6VV70	BpeR	<i>Burkholderia pseudomallei</i>	Controla la expresión de la bomba de eflujo BpeAB-OprB que extruye gentamicina, estreptomicina eritromicina y acriflavina.	1	Chan <i>et al.</i> , 2004.
7	P31676	EnvR	<i>E. coli</i> K12	Regula el operón <i>acrEF</i> que codifica una bomba de eflujo responsable de la multiresistencia a drogas en <i>E. coli</i> . Su especificidad de sustrato es (antibióticos, pigmentos básicos y detergentes) es similar a la de AcrAB.	1	Klein <i>et al.</i> , 1991.
8	P96222	EthR	<i>Mycobacterium tuberculosis</i>	Resistencia a etionamida.	1	Baulard <i>et al.</i> , 2000; Engohang-Ndong <i>et al.</i> , 2004; Garnier <i>et al.</i> , 2003.
9	P72185	HemR	<i>Propionibacterium freudenreichii</i>	Probablemente regula a <i>hemX</i> , que parece estar implicado en transporte de hemo.	1	Hashimoto <i>et al.</i> , 1997.
10	Q93QZ7	HydR	Tn5398 <i>Clostridium difficile</i>	Implicado en la resistencia a eritromicina.	1	Farrow <i>et al.</i> , 2001.
11	O68442	IfeR	<i>Agrobacterium tumefaciens</i> 1D1609	Parece ser un represor que controla la expresión de un supuesto sistema de eflujo de isoflavonoides codificado por el operón <i>ifeABR</i> .	1	Palumbo <i>et al.</i> , 1998.
12	Q9ZGB7	LanK	<i>Streptomyces cyano-genus</i>	Probablemente es un regulador de resistencia a landomicina A.	1	Rebets <i>et al.</i> , 2003; Westrich <i>et al.</i> , 1999.
13		LfrR	<i>Mycobacterium smegmatis</i>	Controla la expresión del gen <i>lfrA</i> cuyo producto confiere resistencia a fluoroquinolonas, bromuro de etidio y acriflavina.	1	Li <i>et al.</i> , 2004.
14	O34619	LmrA	<i>Bacillus subtilis</i>	Probable represor del operón de resistencia a la lincomicina.	1	Kumano <i>et al.</i> , 1997; Kumano <i>et al.</i> , 2003; Murata <i>et al.</i> , 2003.

15	P39897	MtrR	<i>Neisseria gonorrhoeae</i>	Represor de la transcripción de los genes <i>mtrCDE</i> , que codifican un bomba de eflujo que confiere multiresistencia a drogas. MtrR actúa directa o indirectamente como regulador positivo de la expresión de los genes <i>farAB</i> .	1	Delahay <i>et al.</i> , 1997; Hagman <i>et al.</i> , 1995; Hagman and Shafer, 1995; Lucas <i>et al.</i> , 1995; Lucas <i>et al.</i> , 1997; Pan and Spratt, 1994; Rouquette <i>et al.</i> , 1999; Rouquette-Loughlin <i>et al.</i> , 2002; Salah-Bey and Thompson, 1995; Zaranonelli <i>et al.</i> , 1999; Zaranonelli <i>et al.</i> , 2001.
16	Q9F0Y2	Pip	<i>Streptomyces coelicolor</i>	Un regulador inducido por la pristinamicina I que controla la expresión de genes de multiresistencia a drogas.	1	Folcher <i>et al.</i> , 2001.
17	Q9F147	PqrA	<i>Streptomyces coelicolor</i>	Probablemente represor de <i>pqrB</i> , que codifica una bomba de eflujo que confiere resistencia al paraquat.	1	Cho <i>et al.</i> , 2003.
18	P23217	QacR	<i>Staphylococcus aureus</i>	Regula la bomba QacA de extrusión de múltiples drogas.	1	Grkovic <i>et al.</i> , 1998; Grkovic <i>et al.</i> , 2001; Mitchell <i>et al.</i> , 1999; Paulsen <i>et al.</i> , 1995; Paulsen <i>et al.</i> , 1996; Rouch <i>et al.</i> , 1990; Schumacher <i>et al.</i> , 2001; Schumacher <i>et al.</i> , 2002; Tennent <i>et al.</i> , 1985.
19	O52558	RifQ	<i>Amycolatopsis mediterranei</i>	Situado en la agrupación de genes de la biosíntesis de rifampicina y probablemente relacionado con el gen adyacente que codifica una proteína de eflujo de rifampicina.	1	August <i>et al.</i> , 1998.
20	Q9KIH5	RmrR	<i>Rhizobium etli</i> . Plasmid B	Probablemente regula al operón <i>rmrAB</i> relacionado con una bomba de eflujo multi-droga implicada en la sensibilidad a fitoalexinas, flavonoides y salicilatos.	1	González-Pasayo and Martínez-Romero, 2000.
21	Q9AMH9	SimReg 2	<i>Streptomyces antibioticus</i>	Incluido en el agrupamiento de genes de la biosíntesis de simociclinona de <i>Streptomyces antibioticus</i> ; probablemente regula a la proteína hipotética de exportación SimEX.	1	Trefzer <i>et al.</i> , 2002.
22	Q8KLP4	SmeT	<i>Stenotrophomonas maltophilia</i>	Un represor de la bomba de eflujo multi-droga SmeDEF de <i>Stenotrophomonas maltophilia</i> .	1	Sánchez <i>et al.</i> , 2002; Zhang <i>et al.</i> , 2001.
23	Q9R9T9	SrpR	<i>Pseudomonas putida</i>	Probable regulador de la bomba de resistencia a disolventes SrpABC de la cepa S12.	1	Jefferson <i>et al.</i> , 2003; Kieboom <i>et al.</i> , 1998a; Kieboom <i>et al.</i> , 1998b; Wery <i>et al.</i> , 2001.
24	P39885	TcmR	<i>Streptomyces glaucescens</i>	Un regulador de la resistencia a tetraceno-micina C reprimiendo la expresión del gen <i>tcmA</i> , que codifica una bomba de exportación.	1	Guilfoile and Hutchinson, 1992a; Guilfoile and Hutchinson, 1992b.
25	P09164	TetR	<i>Escherichia coli</i>	Controla la expresión de la resistencia a tetraciclina mediada por el gen <i>tetA</i> , que codifica una bomba de eflujo que actúa como un antiportador acoplado la exportación de $[Mg^{2+}$ -tetraciclina] con la entrada de protones.	1	Baumeister <i>et al.</i> , 1992a; Baumeister <i>et al.</i> , 1992b; Berens <i>et al.</i> , 1992; Berens <i>et al.</i> , 1997; Biburger <i>et al.</i> , 1998; Helbl and Hillen, 1998; Helbl <i>et al.</i> , 1998; Hillen and Berens, 1994; Hillen <i>et al.</i> , 1984; Hinrichs <i>et al.</i> , 1994; Kisker <i>et al.</i> , 1995; Klock <i>et al.</i> , 1985; Muller <i>et al.</i> , 1995; Orth <i>et al.</i> , 1998; Orth <i>et al.</i> , 1999a; Orth <i>et al.</i> , 1999b; Orth <i>et al.</i> , 2000; Tiebel <i>et al.</i> , 2000; Unger <i>et al.</i> , 1984; Wissmann <i>et al.</i> , 1991.
26	Q9AIU0	TtgR	<i>Pseudomonas putida</i>	Regula la bomba de eflujo TtgABC que media la tolerancia a disolventes orgánicos y la resistencia a ampicilina, tetraciclina, cloramfenicol y ácido nalidixico.	1	Duque <i>et al.</i> , 2001; Terán <i>et al.</i> , 2003.
27	Q93PU7	TtgW	<i>Pseudomonas putida</i>	<i>ttgW</i> es un pseudogen.	1	Rojas <i>et al.</i> , 2001; Rojas <i>et al.</i> , 2003.

28	Q9RP98	UrdK	<i>Streptomyces fradiae</i> Tu2717	Probablemente regula una bomba de eflujo de urdamicina A.	1	Faust <i>et al.</i> , 2000.
29	Q9AJL5	VarR	<i>Streptomyces virginiae</i>	Regula la transcripción de <i>varS</i> , el transportador específico de virginiamicina S de una forma dependiente de virginiamicina S.	1	Namwat <i>et al.</i> , 2001.
30	P96676	YdeS	<i>Bacillus subtilis</i>	Similar a un regulador de transporte de antibióticos complejos en <i>Streptomyces hygroscopicus</i> .	1	Beloin <i>et al.</i> , 1997.
31	Q54189	ArpA	<i>Streptomyces griseus</i>	Reprime la expresión de <i>adpA</i> . AdpA activa la expresión de <i>strR</i> , y la proteína StrR activa la expresión de los genes de biosíntesis de estreptomycin. ArpA también controla la morfogénesis.	2, 5, 8	Horinouchi <i>et al.</i> , 2001; Ohnishi <i>et al.</i> , 1999; Onaka <i>et al.</i> , 1995; Onaka and Horinouchi 1997; Onaka <i>et al.</i> , 1997; Sugiyama <i>et al.</i> , 1998; Vujaklija <i>et al.</i> , 1993; Yamazaki <i>et al.</i> , 2000.
32	Q93M20	Aur1B	<i>Streptomyces aureofaciens</i>	Incluido en el agrupamiento de genes de biosíntesis del poliketido auricina en <i>Streptomyces aureofaciens</i> auricin.	2	Novakova <i>et al.</i> , 2002.
33	Q9LBV6	BarA	<i>Streptomyces virginiae</i>	Probablemente implicado en la regulación de la biosíntesis de virginiamicina.	2	Kawachi <i>et al.</i> , 2000; Nakano <i>et al.</i> , 1998.
34	Q8KNI9	CalR1	<i>Micromonospora echinospora</i>	Incluido en el agrupamiento de genes de la calicheamicina.	2	Ahlert <i>et al.</i> , 2002.
35	O66129	CprB	<i>Streptomyces coelicolor</i>	CprB está implicado en el control de la biosíntesis de actinorhodina y undecilprodiosina y de la morfogénesis.	2	Natsume <i>et al.</i> , 2004; Onaka <i>et al.</i> , 1998.
36	O24741	FarA	<i>Streptomyces lavendulae</i> FRI-5	Receptor específico de IM-2; juega un papel importante en la regulación del metabolismo secundario y de la biosíntesis de los antibióticos showdomicina y minimicina en <i>Streptomyces lavendulae</i> . FarA actúa como un regulador negativo de la transcripción para la biosíntesis de antibióticos nucleósidos y del pigmento azul, "encendiendo" su expresión en presencia de IM-2. También actúa como un regulador de la transcripción positivo para la biosíntesis de D-cicloserina, "apagando" su expresión en presencia de IM-2.	2	Kitani <i>et al.</i> , 1999; Kitani <i>et al.</i> , 2001; Waki <i>et al.</i> , 1997.
37	Q939Q2	JadR*	<i>Streptomyces venezuelae</i>	Incluido en el agrupamiento de genes para la biosíntesis del componente dideoxiazúcar de la jadomicina B.	2	Wang <i>et al.</i> , 2002.
38	Q56153	JadR2	<i>Streptomyces venezuelae</i>	Reprime la biosíntesis de jadomicina B y parece controlar la pigmentación celular.	2	Yang <i>et al.</i> , 2001; Yang <i>et al.</i> , 1995.
39	Q9ZN97	MphB	<i>Escherichia coli</i> Plasmid pTZ3721	Represor de la biosíntesis de antibióticos.	2	Katayama and Noguchi, 1999; Noguchi <i>et al.</i> , 2000a.
40	Q9XDF0	NonG	<i>Streptomyces griseus</i> <i>sbsp. griseus</i>	Probablemente relacionado con la biosíntesis de la nonactina.	2	Walczak <i>et al.</i> , 2000.
41	Q9RF02	PhlF	<i>Pseudomonas fluorescens</i>	Un represor del operón <i>phlABCD</i> responsable de la biosíntesis del fungicida 2,4-diacetilfloroglucinol (PHL).	2	Schnider-Keel <i>et al.</i> , 2000.
42	Q9ZHP8	TylQ	<i>Streptomyces fradiae</i>	El receptor de butirolactona TylQ es un regulador potencial de la producción del antibiótico macrólido tilosina.	2, 8	Stratigopoulos and Cundliffe 2002.
43	Q8VQC6	VanT	<i>Vibrio anguillarum</i>	Regula positivamente la producción de serina metaloproteasa, pigmento y biofilm.	2, 5	Croxatto <i>et al.</i> , 2002.
44	Q9RPK9	TarA	<i>Streptomyces tendae</i>	Receptor hipotético de gamma-butirolactona, que regula la biosíntesis de nikkomicina.	2, 8	Engel <i>et al.</i> , 2001.
45	Q9XCC7	TylP	<i>Streptomyces fradiae</i>	Probable receptor de gamma-butirolactona que regula la producción de tilosina y la diferenciación morfológica.	2, 5, 8	Bate <i>et al.</i> , 1999; Stratigopoulos and Cundliffe 2002; Stratigopoulos <i>et al.</i> , 2002.
46	Q59213	Bm1P1	<i>Bacillus megaterium</i>	Probablemente actúa como un regulador positivo implicado en la expresión del gen P450BM-1, interfiriendo con la unión de la proteína represora Bm3R1 a las regiones reguladoras de P450BM-1	3	Shaw <i>et al.</i> , 2000; Shaw <i>et al.</i> , 1997; Shaw <i>et al.</i> , 1998.
47	O68276	Bm1P1	<i>Bacillus megaterium</i> ATCC14581	Afecta negativamente a la expresión del nivel basal de P450BM-1, una monooxigenasa de tipo P450 inducible por barbituratos. Los citocromos P450BM-3 y P450BM-1 catalizan la hidroxilación de ácidos	3	He <i>et al.</i> , 1995; Liang and Fulco, 1995; Liang <i>et al.</i> , 1995; Palmer <i>et al.</i> , 1998; Shaw and Fulco, 1992; Shaw <i>et al.</i> , 2000a; Shaw <i>et al.</i> , 2000b; Shaw <i>et al.</i> , 1998.

48	P43506	Bm3R1	<i>Bacillus megaterium</i>	grasos. Un represor de la transcripción implicado en la regulación de proteínas inducidas por barbituratos en <i>Bacillus megaterium</i> .	3	English <i>et al.</i> , 1994; English <i>et al.</i> , 1996; English <i>et al.</i> , 1997; He <i>et al.</i> , 1995; Liang <i>et al.</i> , 1998; Liang and Fulco, 1995; Palmer <i>et al.</i> , 1998; Shaw and Fulco, 1992; Shaw and Fulco, 1993; Shaw <i>et al.</i> , 2000a; Shaw <i>et al.</i> , 2000b; Shaw <i>et al.</i> , 1997; Shaw <i>et al.</i> , 1998.
49	Q9aj68	ButR	<i>Streptomyces cinnamonensis</i> .	Supuesto represor de la transcripción de la crotonil-CoA reductasa.	3	Liu <i>et al.</i> , 1999; Liu <i>et al.</i> , 2001.
50	Q93TU7	CampR	<i>Rhodococcus sp.</i> NCIMB 9784	Probablemente regula la 6-oxocamfor hidrolasa.	3	Grogan <i>et al.</i> , 2001.
51	Q51597	CamR	<i>Pseudomonas putida</i> . Plasmid CAM.	Un regulador negativo de la expresión del operón de la citocromo P-450cam hidroxilasa.	3	Aramaki <i>et al.</i> , 1993; Aramaki <i>et al.</i> , 1995a; Fujita <i>et al.</i> , 1993.
52	O33453	CymR	<i>Pseudomonas putida</i>	Un represor que controla la expresión de los operones <i>cym</i> y <i>cmt</i> y que es inducible por <i>p</i> -cumato pero no por <i>p</i> -cymeno.	3	Choi <i>et al.</i> , 2003; Eaton 1996; Eaton 1997; Ohta <i>et al.</i> , 2001.
53	Q9RAJ1	DhaR	<i>Mycobacterium sp</i> <i>GPI</i>	Reprime la expresión de <i>dhaA</i> . <i>dhaA</i> es el gen de una haloalcano dehalogenasa incluido en el agrupamiento genético del catabolismo del 1-clorobutano.	3	Poelarends <i>et al.</i> , 2000.
54	Q9RA03	KstR	<i>Rhodococcus erythropolis</i> Strain <i>SQ1</i>	Reprime la expresión de <i>kstD</i> que codifica 3-cetosteroide-dehidrogenasa, implicada en la degradación de intermediarios de esteroides de la ruta de degradación del fitoesterol.	3	van Der Geize <i>et al.</i> , 2000.
55	Q8VV87	LexA - like	<i>Terrabacter sp.</i> strain DBF63	Probablemete implicado en la degradación del dibenzofurano.	3	Kasuga <i>et al.</i> , 1997.
56		AcnR	<i>Corynebacterium glutamicum</i>	Represor del gen <i>acn</i> que codifica el enzima aconitasa que controla el ciclo de Krebs.	3	Krug <i>et al.</i> , 2005.
57	Q9fa56	PaaR	<i>Azoarcus evanssi</i>	Probablemente regula los genes <i>paa</i> , que son responsables de la ruta del catabolismo aerobio del ácido fenilacético.	3	Mohamed <i>et al.</i> , 2002.
58	Q9XDW2	PsbI	<i>Rhodopseudomonas palustris</i>	Incluido en el agrupamiento genético que participa en la biodegradación aeróbica del <i>p</i> -cumato.	3	Puskas <i>et al.</i> , 2000.
59	O85706	ThlR	<i>Clostridium acetobutylicum</i> DSM 792	Posiblemete actúa como un represor de la transcripción del operón <i>thlRBC</i> , que está implicado en la biosíntesis de tiolasa.	3	Winzer <i>et al.</i> , 2000.
60	Q59431	UidR	<i>E.coli</i>	Un represor del operón <i>uidRABC(gusRABC)</i> que comprende una beta-D-glucuronidasa (<i>uidA</i>), una glucurónido permeasa (<i>uidB</i>) y una proteína asociada a membrana (<i>uidC</i>).	3	Blanco <i>et al.</i> , 1986.
61	P22645	YDH1	<i>Xanthobacter autotrophicus</i>	Probablemente regula al gen <i>dhlA</i> implicado en la degradación del 1,2-dicloroetano.	3	Janssen <i>et al.</i> , 1989.
62	P17446	BetI	<i>Escherichia coli</i>	Represor dependiente de colina del regulón <i>bet</i> implicado en el estrés osmótico.	4	Andresen <i>et al.</i> , 1988; Lamark <i>et al.</i> , 1995; Lamark <i>et al.</i> , 1996; Røkenes <i>et al.</i> , 1996.
63	Q8NLK1	McbR	<i>Corynebacterium glutamicum</i>	En ausencia de L-metionina, reprime la expresión de seis enzimas clave en la biosíntesis de aminoácidos azufrados, L-cisteína y L-metionina, incluida la utilización de sulfonato y la reducción de sulfito.	4	Rey <i>et al.</i> , 2003.

64	Q9EVJ6	MphR	<i>Escherichia coli</i>	Reprime al operón <i>mph(A)-mrx-mphR(A)</i> en ausencia de eritromicina. La eritromicina induce la síntesis de macróido 2'-fosfotransferasa [Mph(A)], que inactiva a la eritromicina.	4	Noguchi <i>et al.</i> , 2000b.
65	Q9F9Z7	PhaD	<i>Pseudomonas oleovorans</i>	Biosíntesis de poli-3-hidroxicanoatos (PHAs) de cadena de longitud media (MCL) como material de reserva intracelular.	4	Klinke <i>et al.</i> , 2000; Zhang <i>et al.</i> , 2001a.
66	Q9ZF45		<i>Lactococcus lactis</i>	Regula al operón <i>purDEK</i> , que codifica enzimas de la ruta de novo de nucleótidos con purinas.	4	Nilsson and Kilstrup, 1998.
67	P06969	TtK	<i>Escherichia coli</i>	Cotranscrito con el gen <i>dut</i> (deoxy-uridine-triphosphatase).	4	El-Hajj <i>et al.</i> , 1988; Wang and Weiss, 1992.
68	P32398	Yhgd o YixD	<i>Bacillus subtilis</i>	Probablemente relacionado con la biosíntesis de protohemo IX.	4	Hansson and Hederstedt, 1992.
69	Q9F6W0	CasR	<i>Rhizobium etli</i>	Represor del gen <i>casA</i> , que codifica a la calsimina, una proteína tipo camodulina implicada en el desarrollo del bacteroide durante la simbiosis y la fijación simbiótica de nitrógeno.	5	Xi <i>et al.</i> , 2000.
70	Q9RQQ0	IcaR	<i>Staphylococcus aureus</i>	Represor del operón <i>ica</i> que interviene en la formación del componente polisacárido extracelular que actúa como aglutinante en la formación de biofilms.	5	Cramton <i>et al.</i> , 1999; Jefferson <i>et al.</i> , 2003.
71	Q8GLC6	IcaR	<i>Staphylococcus epidermidis</i>	Represor del operón <i>ica</i> que interviene en la formación del componente polisacárido extracelular que actúa como aglutinante en la formación de biofilms.	5	Cho <i>et al.</i> , 2002; Conlon <i>et al.</i> , 2002a; Conlon <i>et al.</i> , 2002b; Knobloch <i>et al.</i> , 2001; Ziebuhr <i>et al.</i> , 1999.
72	Q8KX64	LitR	<i>Vibrio fischeri</i>	Interviene en inducción de luminiscencia, además desempeña un papel positivo en la modulación la capacidad para colonizar a las formas juveniles de calamares y puede controlar el fenotipo de colonia opaco/translucido	5, 8	Fidopiastis <i>et al.</i> , 2002.
73	P21308	LuxR	<i>Vibrio harveyi</i>	Requerido para la expresión del operón <i>luxCDABEGH</i> (luciferasa), responsable bioluminiscencia de la bacteria.	5	Bassler <i>et al.</i> , 1994a; Bassler <i>et al.</i> , 1994b; Cao and Meighen, 1989; Chatterjee <i>et al.</i> , 1996; Johnston <i>et al.</i> , 1986; Martin <i>et al.</i> , 1989; McCarter, 1998; Miyamoto <i>et al.</i> , 1988; Miyamoto <i>et al.</i> , 1996; Miyamoto <i>et al.</i> , 1990; Miyamoto <i>et al.</i> , 1994; Miyamoto <i>et al.</i> , 1998; Shao and Hor, 2001; Silva <i>et al.</i> , 2003; Swartzman <i>et al.</i> , 1990; Swartzman and Meighen, 1993; Swartzman <i>et al.</i> , 1992.
74	Q9ANS7	LuxT	<i>Vibrio harveyi</i>	Activa la expresión de LuxO, la proteína fosfotransferasa que regula la luminiscencia en <i>Vibrio harveyi</i> .	5	Lin <i>et al.</i> , 2000.
75	O50285	OpaR	<i>Vibrio parahaemolyticus</i>	Un regulador de la transcripción que controla la morfología opaca de las colonias de <i>Vibrio parahaemolyticus</i> .	5	McCarter, 1998; Shao and Hor, 2001.
76	Q9XDV7	Orf2	<i>Streptomyces griseus</i>	Probablemente relacionado con la diferenciación dependiente de la fuente de carbono en <i>Streptomyces griseus</i> .	5	Ueda <i>et al.</i> , 1999.

77	Q9L8G8	SmcR	<i>Vibrio vulnificus</i>	Parece jugar un papel importante en la adaptación a condiciones de escasez de nutrientes y en la regulación de algunos genes dependientes de la fase de crecimiento, incluidos algunos factores de virulencia (proteasa, hemolisina). SmcR reprime la movilidad y la producción de fimbrias y de biofilm.	5, 6	Choi <i>et al.</i> , 2002; McDougald <i>et al.</i> , 2000; McDougald <i>et al.</i> , 2001; Shao and Hor, 2001.
78	O30343	HapR	<i>Vibrio cholerae</i>	Un regulador de la transcripción con un papel central en el control de la virulencia de <i>Vibrio cholerae</i> , en una vía dependiente de la densidad celular.	6	Jobling and Holmes, 1997; Kovacicova and Skorupski, 2002; Krukoniš <i>et al.</i> , 2000; McCarter, 1998; Zhu <i>et al.</i> , 2002.
79	Q8KU49	Ef0113	<i>Enterococcus faecalis</i>	Localizado en una isla de patogenicidad en <i>Enterococcus faecalis</i> resistente a vancomicina.	6	Shankar <i>et al.</i> , 2002.
80	O24739	BarB	<i>Streptomyces virginiae</i>	Regula la biosíntesis de virginiamicina.	7	Kinoshita <i>et al.</i> , 1997.
81	O86852	ScbR	<i>Streptomyces coelicolor</i>	Actúa como el receptor citoplasmático que une específicamente a la gamma-butirolactona SCB1 y que negativamente regula la transcripción del <i>gensecA</i> , responsable de la síntesis de SCB1.	2, 7, 8	Aigle <i>et al.</i> , 2000; Takano <i>et al.</i> , 2001; Takano <i>et al.</i> , 2000.
82	Q9JN89	MmfR	<i>Streptomyces coelicolor</i> plasmid SCP1	Supuesto regulador dependiente de lactona.	8	Yamazaki <i>et al.</i> , 2003a.
83	Q9S3L4	AmtR	<i>Corynebacterium glutamicum</i>	AmtR es un regulador del control de nitrógeno.	9	Burkovski, 2003; Jakoby <i>et al.</i> , 2000.
84	Q9EX90	PsrA	<i>Pseudomonas putida</i>	PsrA está implicado en la cascada de regulación que controla la regulación del gen <i>rpoS</i> en respuesta a la densidad celular.	9	Kojic and Venturi, 2001; Rousseau <i>et al.</i> , 2004.
85	P36656	YjdC	<i>Escherichia coli</i>	Probablemente implicado en la tolerancia a cobre.	1 0	Fong <i>et al.</i> , 1995.

Código de funciones.

1. Regulación de bombas de eflujo y transportadores implicados en resistencia a antibióticos y tolerancia a compuestos tóxicos.
2. Regulación de biosíntesis de antibióticos.
3. Regulación de rutas catabólicas.
4. Biosíntesis de productos importantes para bacterias (por ejemplo osmoprotectores, nucleótidos, aminoácidos, PHAs, protohemo).
5. Regulación de diferenciación (esporulación, formación de micelio), fenotipo de colonia, formación de biofilm.
6. Regulación de genes implicados en virulencia.
7. Regulación de síntesis de butirolactonas.
8. Receptores de butirolactonas o de autoinductores.
9. Regulación global.
10. Otros

6.3. Definiendo la familia TetR

6.3.1. El perfil de la familia TetR.

La familia TetR toma su nombre de la proteína TetR que es el miembro mejor caracterizado de la familia, tanto a nivel genético como bioquímico (Helbl *et al.*, 1995; Hillen y Berens, 1994; Hinrichs *et al.*, 1994; Kamionka *et al.*, 2004; Orth *et al.*, 2000; Tovar *et al.*, 1988). Esta proteína controla la expresión de los genes *tet*, cuyos productos confieren resistencia a la tetraciclina (Hinrichs *et al.*, 1994; Kisker *et al.*, 1995; Levy, 1984; Levy, 1988; Saenger *et al.*, 2000; Yamaguchi *et al.*, 1999a; Yamaguchi *et al.*, 1999b). Los miembros de la familia TetR muestran un alto grado de similitud de secuencia en el dominio de unión a ADN (ver más abajo). Interpro (Mulder *et al.*, 2003) asigna proteínas a la familia TetR mediante la *signature* PS01081 de PROSITE (Sigrist *et al.*, 2002), el motivo PRINTS PR00455 (Attwood *et al.*, 2003; Attwood *et al.*, 2000) y el perfil tipo HMM (del inglés *Hidden Markov Model*) PF00440 de Pfam (Bateman *et al.*, 2002; Bateman *et al.*, 2000).

Con el fin de establecer un criterio único para definir a la familia TetR, se decidió desarrollar un perfil convencional, pues este tipo de perfil se ha demostrado que es matemáticamente equivalente a los perfiles HMM, pero es de un uso sencillo y logra una máxima sensibilidad (Gattiker *et al.*, 2003).

El desarrollo del perfil comenzó con la selección de un conjunto de 120 secuencias pertenecientes a la familia TetR atendiendo a dos criterios: (1) una puntuación alta para la *signature* PS01081 de PROSITE, y (2) una puntuación alta con el HMM PF00440.

Mediante BLAST las 120 secuencias se agruparon en 42 conjuntos y se seleccionó una secuencia representativa de cada grupo, las cuales se alinearon con CLUSTAL (<http://clustalw.genome.ad.jp/>).

Este procedimiento reveló que la región más conservada correspondía al dominio hélice-giro-hélice descrito en los cristales TetR y QacR (Grkovic *et al.*, 2001, A; Hinrichs *et al.*, 1994; Orth *et al.*, 1999a; Orth *et al.*, 2000; Orth *et al.*, 1999b; Schumacher *et al.*, 2002a; Schumacher *et al.*, 2001; Schumacher *et al.*, 2002b). La secuencia inicial correspondiente al dominio hélice-giro-hélice se fue extendiendo progresivamente hasta que la puntuación del alineamiento múltiple disminuía.

La Figura 8 muestra el alineamiento final de las 42 secuencias seleccionadas. Este fragmento conservado, de 47 residuos, correspondía en los cristales TetR y QacR con la casi totalidad de la alfa-hélice 1, el dominio hélice-giro-hélice formado por las alfa-hélices 2 y 3, y cinco residuos de la alfa-hélice 4 que conecta la región de interacción con el ADN con el dominio de reconocimiento de efector de la proteína (ver en la Figura 9 en la que se muestra para la estructura tridimensional de TetR).

El alineamiento final (Figura 8) se empleó para la construcción del perfil convencional que se utiliza para identificar a los miembros de la familia TetR. El perfil TetR se construyó usando el programa "pfmake" disponible: http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_pfmake.html (Bucher y Bairoch, 1994; Bucher *et al.*, 1996).

Usando el programa "pfsearch", disponible en: <http://bioweb.pasteur.fr/seqanal/interfaces/pftools.html#pfsearch> (Bucher *et al.*, 1996), el perfil construido se enfrentó contra las 660992 proteínas de bacterias y arqueas depositadas en las bases de datos SWISS-PROT y TrEMBL (versión de Diciembre, 2004). El programa "pfsearch" propuso un valor provisional de Z-score umbral de 8,5 para considerar a una proteína como perteneciente a la familia TetR y ello permitió seleccionar un conjunto de (Bennik *et al.*, 2000) proteínas como miembros hipotéticos de esta familia.

Para comprobar la calidad del nuevo perfil de TetR respecto a su especificidad (falsos positivos), y sensibilidad (falsos negativos), se implementó una nueva herramienta que se ha llamado "Provalidator", que usa los recursos Interpro, Swiss-Prot, Prodom, TIGRfam, CoGnitor, NCBI-RPS-BLAST y PSI-BLAST (Corpet *et al.*, 1999; Haft *et al.*, 2001; Hofmann *et al.*, 1999; Rice *et al.*, 2000; Schultz *et al.*, 2000; Tatusov *et al.*, 2001; Zdobnov y Apweiler, 2001).

En el primer paso se buscaron los posibles falsos positivos entre las 2357 proteínas que se habían asignado a la familia TetR. Interpro asignó 2319 que se consideraron como verdaderos. Las 38 proteínas restantes se analizaron con otros recursos tales como TIGRfam, Prodom, NCBI-RPS-BLAST and PSI-BLAST (Haft *et al.*, 2001; Zdobnov and Apweiler, 2001). Esto nos permitió asignar 34 proteínas a la familia TetR. Tres de los cuatro falsos positivos (Q89RN6, Q988I6 y Q6N8G8) eran miembros de la familia AraC/XylS de activadores de la transcripción (Gallegos *et al.*, 1997; Tobes and Ramos, 2002). Estas proteínas tienen dos dominios hélice-giro-hélice en su extremo C-terminal, característica típica de los miembros de la familia AraC/XylS (Gallegos *et al.*, 1997; Manzanera *et al.*, 2000). Estas 3 proteínas se identificaron como potenciales miembros de la familia TetR porque uno de sus dominios hélice-giro-hélice es similar al dominio de unión a ADN de TetR. El cuarto falso positivo es una transposasa (Q981E7).

"Provalidator" detectó quince falsos negativos (Q742Y2, Q8CJK3, Q73ZY1, Q6D1J7, Q8KU64, Q9A917, Q880T2, Q6D2Z4, Q885G7, Q8PC90, Q9A466, Q9S6C0, Q9ZH26, Q6A626, Q8G822) que son proteínas asignadas a la familia TetR por INTERPRO, pero cuyo Z-score está entre 6,407 y 8,487.

En resumen, el perfil de TetR con un Z-score umbral de 8,5 identificó proteínas que no eran detectadas por INTERPRO y entre las 660992 proteínas analizadas, sólo se encontraron 4 falsos positivos. Estos resultados indican que el nuevo perfil es muy eficaz en la detección de miembros de la familia TetR.

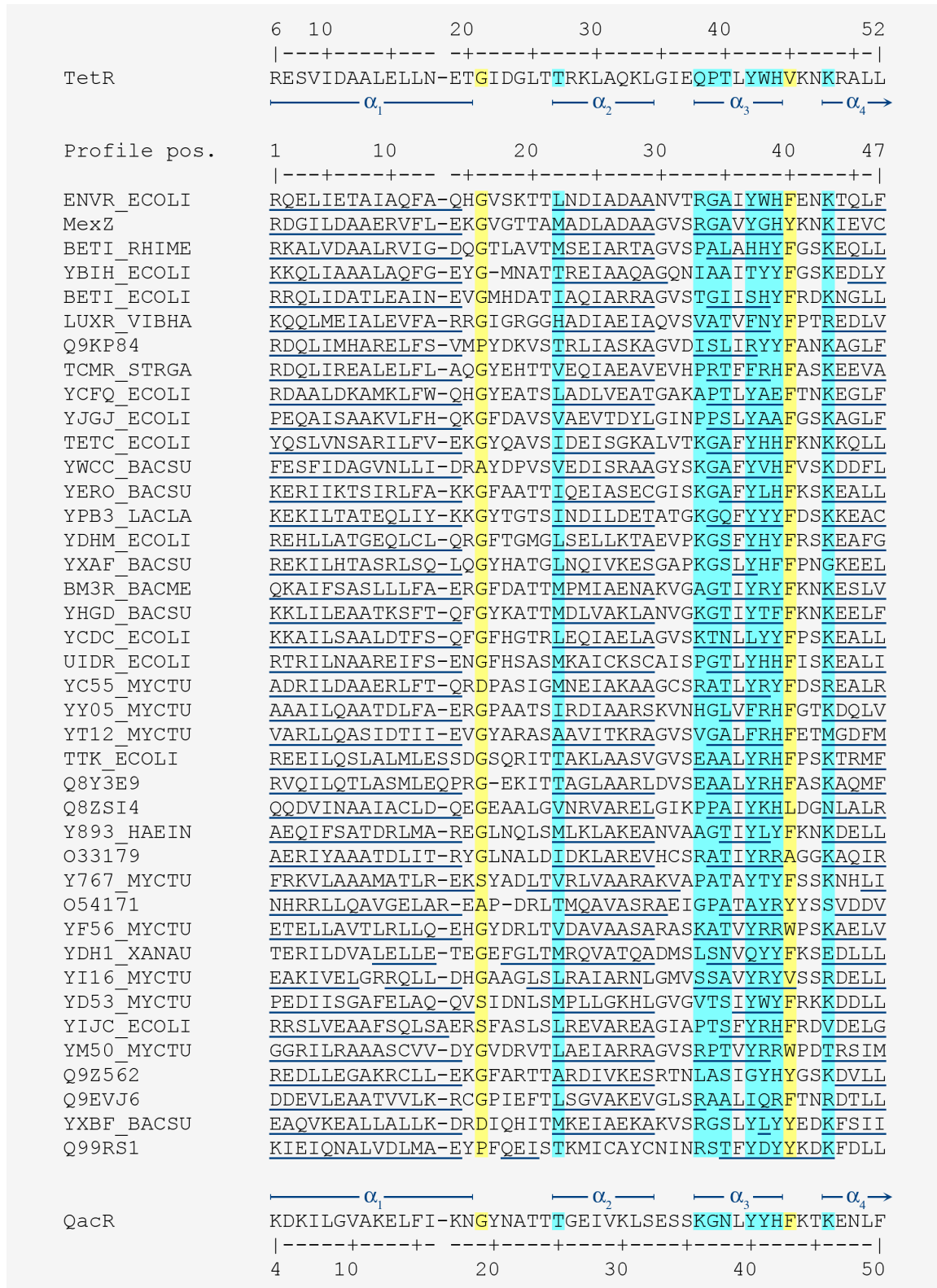


Figura 8. Alineamiento de 42 miembros de la familia TetR que ilustra el perfil de la familia TetR. La columna azul indica los residuos de las α -hélices implicados en contactos con el ADN en la estructura de los cristales de TetR y QacR; la columna amarilla indica los de los giros. Los residuos más conservados están sombreados. Las líneas azules indican la predicción de α -hélices. Las abreviaturas empleadas son: BACME, *Bacillus megaterium*; BACSU, *Bacillus subtilis*; ECOLI, *Escherichia coli*; HAEIN, *Haemophilus influenzae*; LACLA, *Lactobacillus lactis*; MYCTU, *Mycobacterium tuberculosis*; RHIME, *Rhizobium meliloti*; STRGA, *Streptomyces* sp.; VIBHA, *Vibrio haemophilus*; XANAU, *Xantomonas* sp.

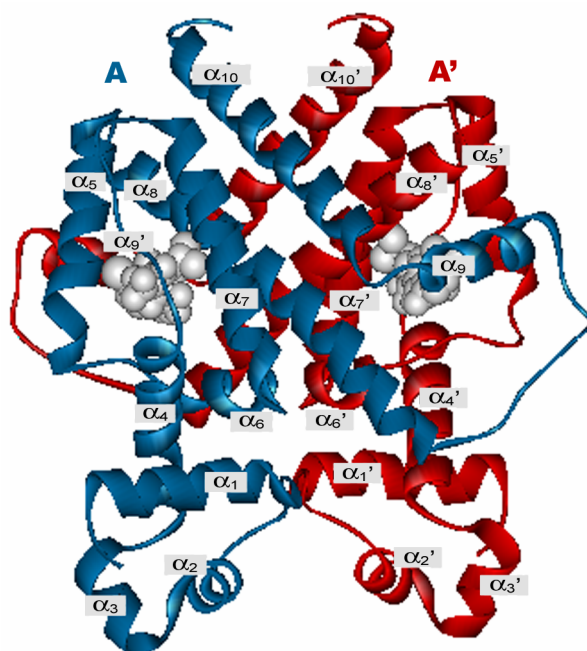


Figura 9. Diagrama "ribbon" (cinta) del homodímero de TetR. Los monómeros se muestran en azul y rojo. Dos moléculas de tetraciclina, una unida a cada monómero, se muestran en gris. Las α -hélices 2 y 3 en el monómero azul y α -2' y α -3' en el monómero rojo constituyen el dominio hélice-giro-hélice (HTH) de unión a ADN. La α -hélice 1 y parte de la hélice α -4 junto con α -hélices 2 y 3 comprenden a la secuencia que mejor define el perfil de la familia TetR. Esta figura está basada en el trabajo de Hinrichs *et al.* (Hinrichs *et al.*, 1994).

6.3.2. Identificación de miembros de la familia TetR en las diferentes bases de datos de ADN y proteínas.

Usando el perfil que se acaba de describir se rastrearon las bases de datos Swiss-Prot y TrEMBL y los 196 genomas microbianos, completos e incompletos, disponibles en NCBI en diciembre de 2004.

Se detectaron 73 proteínas en Swiss-Prot, 2277 en TrEMBL y 2410 en las ORF traducidas correspondientes a los 196 genomas microbianos. Para eliminar la redundancia, el conjunto de las 4758 proteínas TetR se analizó usando el programa SEQUNIQ desarrollado en nuestro laboratorio (Molina-Henares *et al.*, resultados no publicados), que integra el conjunto de secuencias disponibles en las bases de datos de proteínas y ácidos nucleicos. Encontramos 2353 secuencias que sobrepasaron el *Z-score* umbral de

8,5 y por tanto pertenecen a la familia TetR. El dominio hélice-giro-hélice en 2348 miembros de la familia estaba localizado en el extremo N-terminal de las proteínas.

La Tabla 30 muestra que los miembros de la familia TetR se detectaron en 144 genomas microbianos pertenecientes a 80 géneros y 133 especies de bacterias gram-positivas, y α , β y γ -proteobacterias, cianobacterias y arqueas, indicando una amplia distribución taxonómica.

Comprobamos que las proteínas de la familia TetR estaban codificadas por genes tanto cromosómicos como plasmídicos, y que la movilidad de estos últimos elementos podría ser una fuente de dispersión genética de los genes de la familia mediante transferencia horizontal (Heuer *et al.*, 2004; Szczepanowski *et al.*, 2004), como también ocurre con ciertos genes catabólicos (Dogra *et al.*, 2004; Itoh *et al.*, 2004; Martínez-Bueno *et al.*, 2002; Vimr *et al.*, 2004; Wilson *et al.*, 2003) y determinantes de resistencia frente a agentes antimicrobianos (Balsalobre *et al.*, 2003; Fomenko *et al.*, 2003; Grohmann *et al.*, 2003; Schouls *et al.*, 2003).

Se encontró que los miembros de la familia TetR son particularmente abundantes en microorganismos expuestos a condiciones ambientales variables, tales como microorganismos del suelo (por ejemplo, *Nocardia* sp., *Streptomyces* sp., *Bradyrhizobium* sp., *Mesorhizobium* sp., *Pseudomonas* sp., *Bacillus* sp., *Ralstonia* sp., etc.); patógenos de animales y vegetales (por ejemplo, *Agrobacterium* sp., *Brucella* sp., *Escherichia coli*, *Mycobacterium* sp., *Salmonella* sp., etc.), extremófilos (por ejemplo, *Deinococcus*) y bacterias metanogénicas tales como *Methanosarcina acetivorans*. Por el contrario, los miembros de la familia TetR no aparecen en parásitos intracelulares tales como *Chlamydia*, *Mycoplasmas* y endosimbiontes como *Buchera*, en consonancia con sus estilos de vida en ambientes más o menos invariables (Cases *et al.*, 2003). Sin embargo hay que señalar que Dugan *et al.* (2004) recientemente han encontrado que *Chlamydia suis* puede adquirir resistencia frente a tetraciclina mediante transferencia horizontal de islas genómicas que portan los genes *tet*.

Como corolario podemos decir que parece que las proteínas de la familia TetR están implicadas en la adaptación a ambientes variables y complejos. Esto correlaciona con el hecho de que microorganismos con abundancia de factores sigma ECF poseen además muchos miembros de la familia TetR (Cases *et al.*, 2003, Mahren y Braun, 2003; Martínez-Bueno *et al.*, 2002; Ohki *et al.*, 2003; Yoshimura *et al.*, 2004).

Tabla 30. Distribución de proteínas TetR en procariotas.

Microorganismo	Tño.genoma (Mb)	Nº
<i>Nocardia farcinical</i> FM 10152	6.21	151
<i>Streptomyces coelicolor</i> A3(2)	9.05	150
<i>Streptomyces avermitilis</i> MA-4680	9.12	116
<i>Mycobacterium avium subsp. Paratuberculosis</i> str. k10	4.83	108
<i>Agrobacterium tumefaciens</i> str. C58	11.35	61
<i>Bradyrhizobium japonicum</i> USDA 110	9.11	59
<i>Mycobacterium bovis</i> AF2122/97	4.35	51
<i>Mycobacterium tuberculosis</i> CDC1551	4.40	51
<i>Mycobacterium tuberculosis</i> H37Rv	4.41	51
<i>Bacillus licheniformis</i> ATCC 14580	8.44	48
<i>Mesorhizobium loti</i> MAFF303099	7.60	47
<i>Rhodopseudomonas palustres</i> CGA009	5.46	40
<i>Pseudomonas aeruginosa</i> PAO1	6.26	38
<i>Bacillus cereus</i> ATCC 10987	5.22	36
<i>Bacillus anthracis</i> str. 'Ames Ancestor'	5.50	32
<i>Bacillus anthracis</i> str. A2012	5.37	32
<i>Bacillus anthracis</i> str. Sterne	5.23	31
<i>Bacillus anthracis</i> str. Ames	5.23	30
<i>Bacillus cereus</i> ZK	5.30	30
<i>Bacillus cereus</i> ATCC 14579	5.43	29
<i>Bordetella bronchiseptica</i> RB50	5.34	28
<i>Pseudomonas syringae</i> pv. tomato str. DC3000	6.40	28
<i>Sinorhizobium meliloti</i> 1021	6.69	28
<i>Bacillus thuringiensis</i> serovar konkukian str. 97-27	5.24	27
<i>Clostridium acetobutylicum</i> ATCC 824	4.13	27
<i>Caulobacter crescentus</i> CB15	4.02	26
<i>Lactobacillus plantarum</i> WCFS1	3.31	26
<i>Pseudomonas putida</i> KT2440	6.18	25
<i>Burkholderia pseudomallei</i> K96243	7.25	24
<i>Ralstonia solanacearum</i> GMI1000	5.81	24
<i>Photobacterium profundum</i> SS9	6.40	23
<i>Oceanobacillus iheyensis</i> HTE831	3.63	22
<i>Bordetella parapertussis</i> 12822	4.77	21
<i>Burkholderia mallei</i> ATCC 23344	5.84	21
<i>Bacillus halodurans</i> C-125	4.20	20
<i>Bacillus subtilis</i> subsp. subtilis str. 168	4.21	20
<i>Acinetobacter</i> sp. ADP1	3.60	19
<i>Bordetella pertussis</i> Tohama I	4.09	17
<i>Chromobacterium violaceum</i> ATCC 12472	4.75	17
<i>Shewanella oneidensis</i> MR-1	5.13	17
<i>Vibrio vulnificus</i> CMCP6	5.13	17
<i>Escherichia coli</i> CFT073	5.23	16
<i>Gloeobacter violaceus</i> PCC 7421	4.66	16
<i>Methanosarcina acetivorans</i> C2A	5.75	16
<i>Streptococcus mutans</i> UA159	2.03	16
<i>Vibrio parahaemolyticus</i> RIMD 2210633	5.17	16
<i>Vibrio vulnificus</i> YJ016	5.26	16
<i>Xanthomonas axonopodis</i> pv. citri str. 306	5.18	16
<i>Corynebacterium glutamicum</i> ATCC 13032	3.31	15
<i>Deinococcus radiodurans</i> R1	3.28	15
<i>Erwinia carotovoras</i> subsp. atroseptica SCRI1043	5.06	15
<i>Xanthomonas campestris</i> pv. campestris str. ATCC 33913	5.08	15
<i>Escherichia coli</i> O157:H7	5.59	14
<i>Corynebacterium efficiens</i> YS-314	3.15	13
<i>Escherichia coli</i> O157:H7 EDL933	5.53	13
<i>Lactococcus lactis</i> subsp. lactis I11403	2.37	13

<i>Leifsonia xyli</i> subsp. <i>xyli</i> str. CTCB07	2.58	13
<i>Listeria monocytogenes</i> str. 4b F2365	2.91	13
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi str. CT18	5.13	13
<i>Salmonella typhimurium</i> LT2	4.95	13
<i>Treponema denticola</i> ATCC 35405	2.84	13
<i>Corynebacterium diphtheriae</i> NCTC 13129	2.49	12
<i>Escherichia coli</i> K12	4.64	12
<i>Listeria innocua</i> Clip11262	3.01	12
<i>Listeria monocytogenes</i> EGD-e	2.94	12
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi Ty2	4.79	12
<i>Shigella flexneri</i> 2a str. 301	4.61	12
<i>Vibrio cholerae</i> O1 biovar eltor str. N16961	4.03	12
<i>Nostoc</i> sp. PCC 7120	7.21	11
<i>Enterococcus faecalis</i> V583	3.22	10
<i>Mycobacterium leprae</i> TN	3.27	10
<i>Shigella flexneri</i> 2a str. 2457T	4.60	10
<i>Geobacter sulfurreducens</i> PCA	3.81	9
<i>Leptospira interrogans</i> serovar Copenhageni str. Fiocruz L1-130	4.63	9
<i>Leptospira interrogans</i> serovar Lai str. 56601	4.69	9
<i>Propionibacterium acnes</i> KPA171202	2.56	9
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> MRSA252	2.90	9
<i>Bifidobacterium longum</i> NCC2705	2.26	8
<i>Brucella melitensis</i> 16M	3.29	8
<i>Brucella suis</i> 1330	3.32	8
<i>Photorhabdus luminescens</i> subsp. <i>laumondii</i> TTO1	5.69	8
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50	2.90	8
<i>Symbiobacterium thermophilum</i> IAM 14863	3.57	8
<i>Yersinia pestis</i> biovar Medievalis str. 91001	4.80	8
<i>Yersinia pseudotuberculosis</i> IP 32953	4.84	8
<i>Bacteroides fragilis</i> YCH46	5.31	7
<i>Bacteroides thetaiotaomicron</i> VPI-5482	6.26	7
<i>Bdellovibrio bacteriovorus</i> HD100	3.78	7
<i>Desulfovibrio vulgaris</i> subsp. <i>vulgaris</i> str. Hildenborough	3.77	7
<i>Fusobacterium nucleatum</i> subsp. <i>nucleatum</i> ATCC 25586	2.17	7
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> MSSA476	2.80	7
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2	2.82	7
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315	2.84	7
<i>Yersinia pestis</i> CO92	4.83	7
<i>Desulfotalea psychrophila</i> LSv54	3.66	6
<i>Lactobacillus johnsonii</i> NCC 533	1.99	6
<i>Mannheimia succiniciproducens</i> MBEL55E	2.31	6
<i>Rhodopirellula baltica</i> SH 1	7.15	6
<i>Streptococcus agalactiae</i> NEM316	2.21	6
<i>Yersinia pestis</i> KIM	4.60	6
<i>Aquifex aeolicus</i> VF5	1.59	5
<i>Methylococcus capsulatus</i> str. Bath	3.30	5
<i>Streptococcus agalactiae</i> 2603V/R	2.16	5
<i>Streptococcus pyogenes</i> MGAS10394	1.90	5
<i>Streptococcus pyogenes</i> MGAS8232	1.90	5
<i>Clostridium perfringens</i> str. 13	3.09	4
<i>Methanococcus maripaludis</i> S2	1.66	4
<i>Staphylococcus epidermidis</i> ATCC 12228	2.50	4
<i>Streptococcus pyogenes</i> M1 GAS	1.85	4
<i>Streptococcus pyogenes</i> MGAS315	1.90	4
<i>Streptococcus pyogenes</i> SSI-1	1.89	4
<i>Thermus thermophilus</i> HB27	2.13	4
<i>Clostridium tetani</i> E88	2.80	3
<i>Haemophilus influenzae</i> Rd KW20	1.83	3
<i>Methanosarcina mazei</i> Go1	4.10	3
<i>Nitrosomonas europaea</i> ATCC 19718	2.81	3
<i>Pasteurella multocida</i> subsp. <i>multocida</i> str. Pm70	2.26	3
<i>Porphyromonas gingivalis</i> W83	2.34	3

<i>Streptococcus pneumoniae</i> R6	2.04	3
<i>Streptococcus pneumoniae</i> TIGR4	2.16	3
<i>Synechocystis</i> sp. PCC 6803	3.57	3
<i>Thermoanaerobacter tengcongensis</i>	2.69	3
<i>Wolinella succinogenes</i> DSM 1740	2.11	3
<i>Haemophilus ducreyi</i> 35000HP	1.70	2
<i>Halobacterium salinarum</i> NRC-1	2.57	2
<i>Legionella pneumophila</i> str. Lens	3.41	2
<i>Legionella pneumophila</i> str. Paris	3.64	2
<i>Legionella pneumophila</i> subsp. pneumophila str. Philadelphia 1	3.40	2
<i>Methanothermobacter thermautotrophicus</i> str. Delta H	1.75	2
<i>Neisseria meningitidis</i> MC58	2.27	2
<i>Neisseria meningitidis</i> Z2491	2.18	2
<i>Thermotoga maritima</i> MSB8	1.86	2
<i>Xylella fastidiosa</i> 9a5c	2.73	2
<i>Archaeoglobus fulgidus</i> DSM 4304	2.18	1
<i>Campylobacter jejuni</i> subsp. jejuni NCTC 11168	1.64	1
<i>Coxiella burnetii</i> RSA 493	2.00	1
<i>Helicobacter hepaticus</i> ATCC 51449	1.80	1
<i>Mycoplasma penetrans</i> HF-2	1.36	1
<i>Picrophilus torridus</i> DSM 9790	1.55	1
<i>Pyrococcus abyssi</i> GE5	1.77	1
<i>Sulfolobus solfataricus</i> P2	2.99	1
<i>Sulfolobus tokodaii</i> str. 7	2.69	1
<i>Ureaplasma parvum</i> serovar 3 str. ATCC 700970	0.75	1

6.4. Proteínas de la familia TetR de estructura tridimensional conocida

El elevado grado de identidad en la secuencia primaria del fragmento que define la región hélice-giro-hélice del perfil de TetR, probablemente se refleje en una estructura tridimensional común para este dominio en los miembros de la familia.

Esta hipótesis se apoya en el hecho de la alta identidad en la estructura tridimensional del hélice-giro-hélice de TetR, QacR, CprB y EthR como se deduce de la superposición de estas regiones, y del elevado grado de conservación del alineamiento (Dover *et al.*, 2004; Natsume *et al.*, 2004; Schumacher *et al.*, 2002a; y Figura 10).

Como en otras familias de reguladores de la transcripción, no se apreció conservación de la secuencia primaria fuera del dominio hélice-giro-hélice, lo que probablemente sea reflejo de las diferencias en el tipo de señal detectada por los reguladores de la familia, por ejemplo, antibióticos con estructuras distintas, barbitúricos, homoserin-lactonas, disolventes orgánicos, colina, etc (ver Tabla 29). Sin embargo y a pesar de esto, sí observamos una conservación global de la estructura tridimensional (Figuras 17 y 18).

Dado que todos los miembros de la familia cuya función se conoce, actúan como represores, probablemente funcionen de un modo similar: la unión de la molécula inductora al dominio no conservado, causaría cambios conformacionales que se transmitirían a la región conservada, de unión a ADN, lo que resultaría en la liberación del represor del operador, permitiendo la transcripción desde los promotores correspondientes.

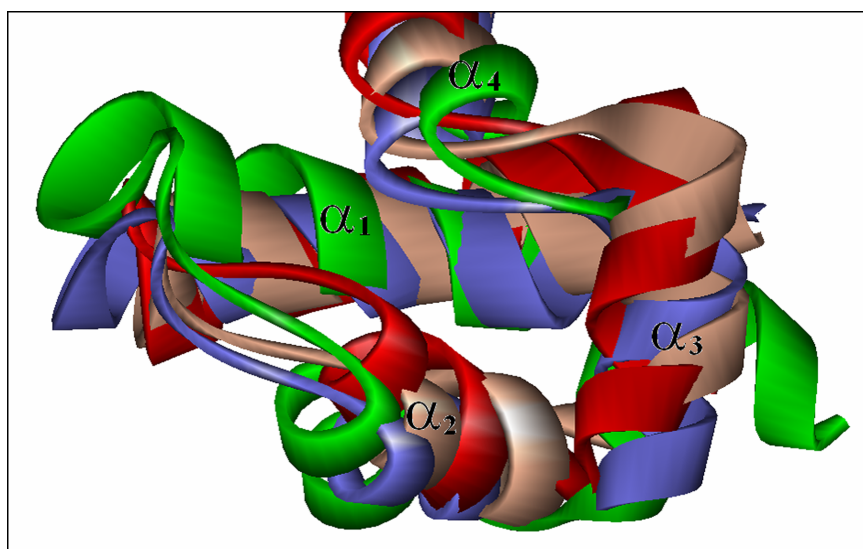


Figura 10: Superposición de la estructura tridimensional de la secuencia peptídica que define al perfil en las moléculas cristalizadas de función conocida (TetR rojo, QacR salmón, CprB verde y EthR azul). Figura realizada con el Método CE (*Combinatorial Extension*) (Shindyalov y ourne 1998), disponible en <http://cl.sdsc.edu/ce.html>.

Para entender mejor el mecanismo de acción de los miembros de la familia TetR se analizó en detalle la estructura tridimensional de los cuatro miembros de la familia cuyas estructuras cristalinas se habían dilucidado antes o durante la elaboración de este trabajo de Tesis Doctoral, a saber, TetR, QacR, CprB y EthR (Hinrichs *et al.*, 1994; Natsume *et al.*, 2004; Orth *et al.*, 1998; Orth *et al.*, 1999a; Orth *et al.*, 2000; Orth *et al.*, 1999b; Schumacher *et al.*, 2002a; Schumacher *et al.*, 2001; Schumacher *et al.*, 2002b). De esta manera buscábamos identificar características comunes y diferenciadoras entre los miembros de la familia TetR.

6.4.1. El regulador TetR.

Resistencia a la tetraciclina y el papel del regulador de la transcripción TetR. Las tetraciclinas están entre los antibióticos de amplio espectro más comúnmente empleados (Kustu *et al.*, 1991; Lamark *et al.*, 1995). Actúan mediante la unión a la subunidad menor del ribosoma, interrumpiendo la elongación de la cadena polipeptídica mediante un mecanismo desconocido. Muchas bacterias gram-negativas han desarrollado mecanismos de resistencia contra este antibiótico. El más frecuente implica a una proteína asociada a membrana (TetA) que expulsa el antibiótico de la célula bacteriana antes de que ocurra la inhibición de la elongación del polipéptido (Kaneko *et al.*, 1985; Levy *et al.*, 1999; Tauch *et al.*, 2000; Yamaguchi *et al.*, 1999a; Yamaguchi *et al.*, 1999b; Zhao y Aoki, 1992).

Contiguo a *tetA* y con orientación divergente se sitúa *tetR* (George y Levy, 1983) cuyo producto génico controla estrechamente la expresión de los dos genes, *tetA* y *tetR* (Hillen and Berens, 1994; Hinrichs *et al.*, 1994). La región intergénica contiene dos operadores idénticos separados por 11 pares de bases. TetR se une a dichos operadores evitando la transcripción desde ambos promotores (Figura 11; Orth *et al.*, 2000). En todas las estructuras de cristales de TetR resueltas hasta la fecha (PDB identifiers: 2TCT; 2TRT; 1A6I; 1BJO; 1BJY; 1BJZ; 1ORK; 1RP1), este represor aparece como un homodímero (Baumeister *et al.*, 1992a; Baumeister *et al.*, 1992b; Isackson y Bertrand, 1985; Kisker *et al.*, 1995; Orth *et al.*, 1999a; Orth *et al.*, 2000; Orth *et al.*, 1999b; Sizemore *et al.*, 1990). Este homodímero se une al operador (Figura 11).

Cada operador, de 15 pares de bases, muestra una secuencia palindrómica simétrica con un par de bases central (Figura 11A). Las secuencias de los operadores solapan con los promotores de los genes *tetA* y *tetR*, bloqueando su expresión cuando TetR está unido a sus operadores. Cuando la tetraciclina, acomplejada con Mg^{2+} se une a TetR (Jogun y Stezowski, 1976; Takahashi *et al.*, 1986) ocurre un cambio conformacional

que incapacita a la proteína TetR para permanecer unida al ADN. Como consecuencia, TetR y TetA se expresan (Orth *et al.*, 1998; ver figura 11).

El homodímero TetR está constituido por dos monómeros idénticos cada uno con diez α -hélices conectadas mediante giros y lazos (Figura 9). La estructura tridimensional del monómero de TetR se estabiliza principalmente mediante contactos hidrofóbicos hélice-hélice. La estructura global del homodímero TetR puede dividirse en 2 dominios de unión a ADN en el extremo N-terminal de cada monómero y dos núcleos reguladores, implicados en la dimerización y en la unión a ligando (Hinrichs *et al.*, 1994, Orth *et al.*, 1998; Orth *et al.*, 1999a; Orth *et al.*, 1999b). El dominio de unión a ADN está formado por las hélices $\alpha 1$, $\alpha 2$ y $\alpha 3$ y sus simétricas $\alpha 1'$, $\alpha 2'$ y $\alpha 3'$ (el carácter "prima" indica al segundo monómero). Las hélices $\alpha 4$ y $\alpha 4'$ conectan estos dominios con el dominio regulador constituido por las hélices $\alpha 5$ a $\alpha 10$ y sus simétricas $\alpha 5'$ a $\alpha 10'$ (Hinrichs *et al.*, 1994; Orth *et al.*, 1999a; Orth *et al.*, 1999b). El dominio regulador, responsable de la dimerización, contiene un bolsillo, por monómero, que acomoda a una molécula de tetraciclina en presencia de un catión divalente. Las hélices $\alpha 5$, $\alpha 8$ y $\alpha 10$ y sus simétricas $\alpha 5'$, $\alpha 8'$ y $\alpha 10'$ constituyen el andamio del dominio regulador, y constituyen la estructura más conservada en ambas conformaciones de TetR (Hinrichs *et al.*, 1994; Orth *et al.*, 1999a; Orth *et al.*, 2000; Orth *et al.*, 1999b).

El bolsillo de unión de tetraciclina es idéntico en ambos monómeros. La cavidad a la que se une el complejo $[\text{Tc-Mg}^{2+}]$ se describe en la Figura 12 (Orth *et al.*, 1998; Orth *et al.*, 1999a; Orth *et al.*, 1999b). La entrada a esta cavidad está limitada por la $\alpha 9'$ y el extremo C-terminal de $\alpha 8'$ y el lazo que las conecta, mientras que su fondo se cierra mediante el lazo $\alpha 4$ - $\alpha 5$ (Orth *et al.*, 1999a; Orth *et al.*, 2000; Orth *et al.*, 1999b).

Cuando el complejo $[\text{Tc-Mg}^{2+}]$ entra en el bolsillo, su anillo A contacta con el lazo $\alpha 4$ - $\alpha 5$ y se dispara una cascada de cambios conformacionales. La His100 y la Thr103, ambas de la $\alpha 6$, contactan con el Mg^{2+} lo que desplaza a la $\alpha 6$ que sufre el desenrollamiento de su extremo C-terminal que se transforma en un giro- β (Figura 12). El lazo $\alpha 6$ - $\alpha 7$, también resulta desplazado hacia el inductor, estableciéndose contacto entre la tetraciclina y la Arg104 y la Pro106. El desplazamiento de $\alpha 6$ obliga a $\alpha 4$ a moverse en la misma dirección a causa de los contactos de van der Waals. La His64 de $\alpha 4$, anclada a $\alpha 5$ y a la tetraciclina, actúa como un eje sobre el que $\alpha 4$ sufre un movimiento pendular, que hace rotar dicha hélice respecto al corazón de cada monómero.

Como consecuencia, los dominios de unión a ADN de ambos monómeros se desplazan uno respecto al otro reflejándose en el incremento de la distancia entre sus hélices $\alpha 3$ y $\alpha 3'$ de manera que se rompen los contactos con el ADN (Figura 13; Orth *et al.*, 1998; Orth *et al.*, 1999a; Orth *et al.*, 1999b).

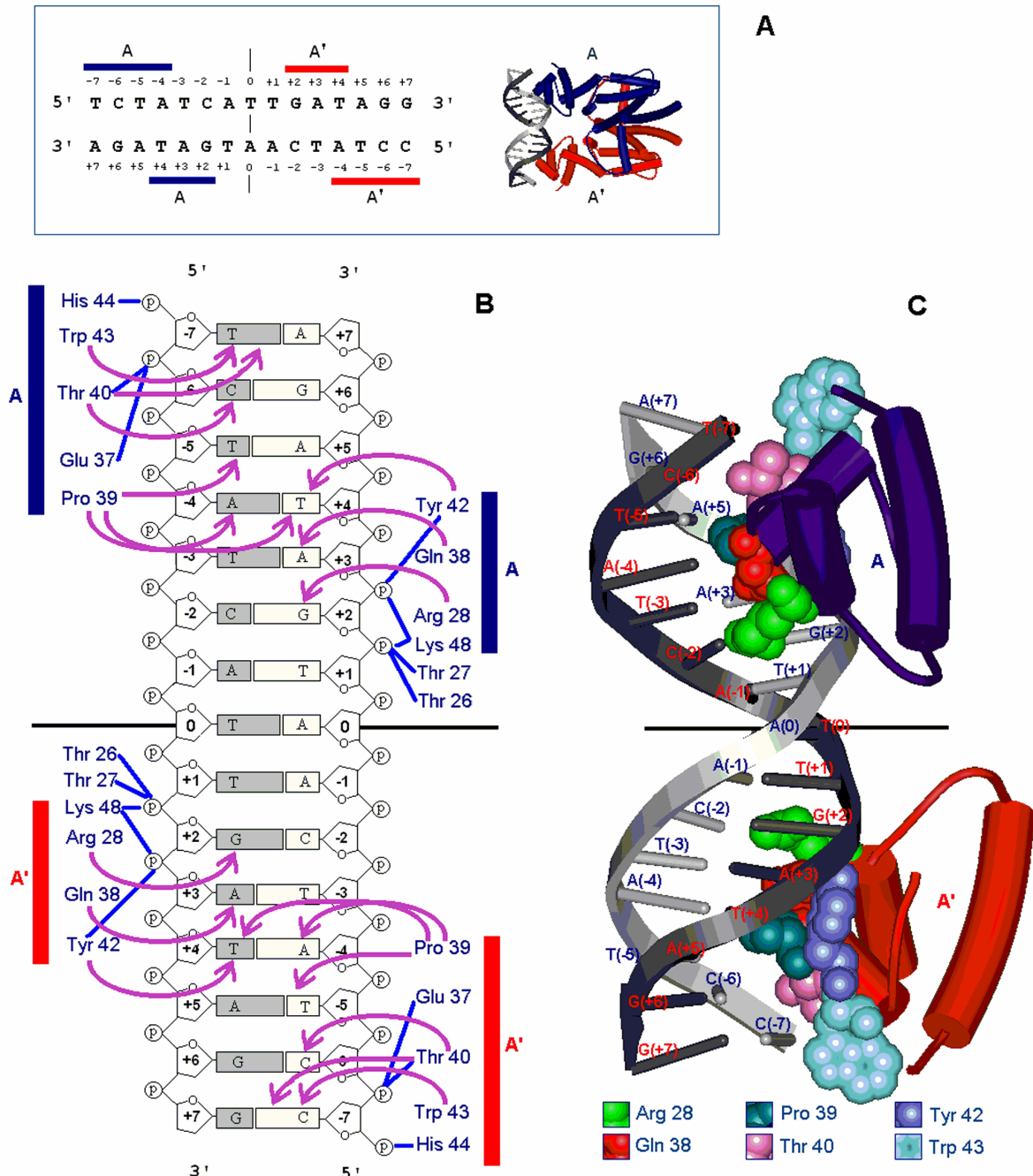


Figura 11. Unión de TetR a su operador. **A)** operador *tetR* y regiones de contacto. El operador *tetR* es una secuencia palindrómica. Las barras horizontales muestran los nucleótidos contactados por cada monómero del dímero TetR. **B)** Interacción de los residuos de TetR con los nucleótidos específicos (flechas) y con los fosfatos del esqueleto (líneas azules) en la región del operador. Los aminoácidos implicados en la unión al ADN se extienden desde el residuo 27 al 48. Los contactos con el operador se confirmaron mediante: ensayos de improntas *in vitro*, análisis de mutantes TetR, y estudios cristalográficos (Baumeister *et al.*, 1992a; Baumeister *et al.*, 1992b; Isackson and Bertrand, 1985; Orth *et al.*, 1998; Sizemore *et al.*, 1990). **C)** Representación de cada homodímero unido al operador *tet* en una representación del ADN en doble hélice. Figura basada en los trabajos de Hinrichs *et al.* (Hinrichs *et al.*, 1994), Orth *et al.* (Orth *et al.*, 1998) y Baumeister *et al.* (Baumeister *et al.*, 1992a; Baumeister *et al.*, 1992b)

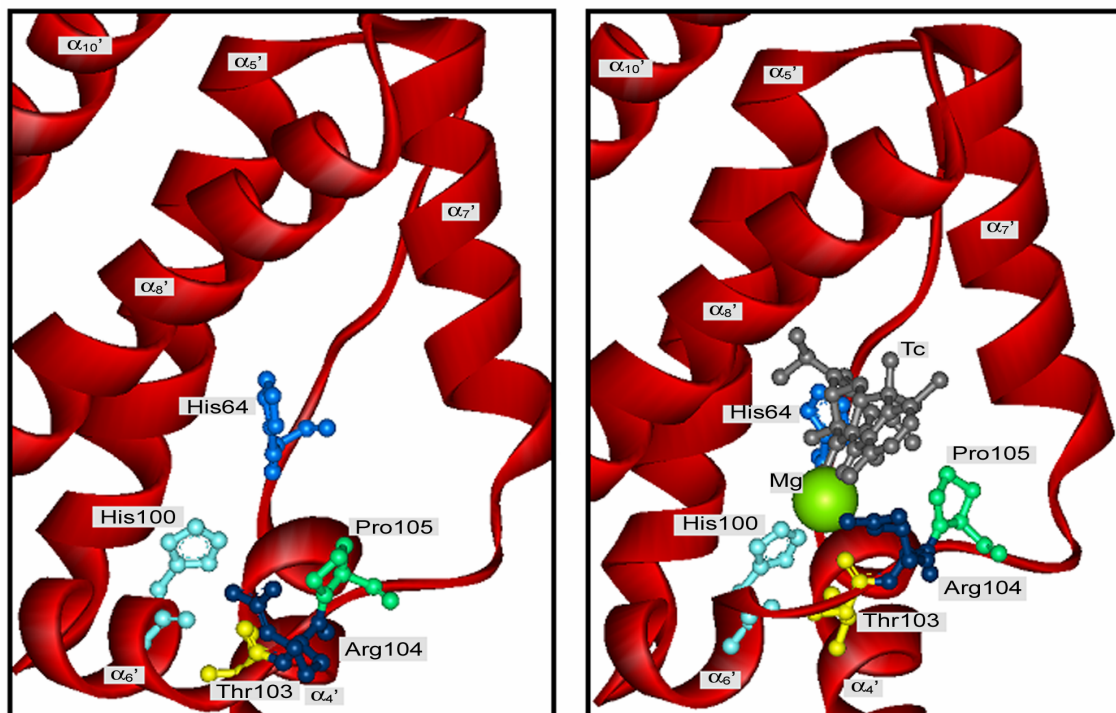


Figura 12. Representación de la cavidad de TetR implicada en la unión a tetraciclina. Panel izquierdo) En ausencia de Tc. Panel derecho) En presencia de Tc. La esfera verde representa al catión Mg^{2+} . Las interacciones específicas no se muestran para no complicar la figura pero se describen en el texto. Figura basada en los trabajos de Orth *et al* y Kisker *et al*.

Con la cascada de cambios conformacionales la tetraciclina queda retenida en la cavidad de unión de cada monómero y TetR pierde su capacidad de unión a su ADN diana de manera irreversible.

Se debe hacer notar que algunos residuos fuera del bolsillo influyen en la afinidad por la tetraciclina, tal como demostraron Kamionka *et al.* (1985), que aislaron un doble mutante (G96E, L205S) que mostraba disminuida su afinidad por el antibiótico. Es posible que el mecanismo de tipo "on/off" que funciona en TetR para responder a señales específicas podría ser usado de manera similar por otros miembros de la familia.

El dominio de unión a ADN de TetR: un dímero, con dos monómeros simétricos, que se une a un operador palindrómico. La cocrystalización de TetR con su operador estableció que el homodímero TetR se dispone perpendicularmente al eje longitudinal del ADN (Figura 11A). Dos regiones adyacentes del surco mayor del ADN, que cubren un área de 6 pares de bases sobre ambas cadenas están implicadas en el acoplamiento casi perfecto con los dos dominios de interacción con ADN (Figuras 11A y 11B; Orth *et al.*, 2000). No se encontraron moléculas de agua en la interfase ADN-TetR, donde la interacciones cruciales son hidrofóbicas (Orth *et al.*, 2000). Las interacciones de cada dominio hélice-giro-hélice con su ADN operador se resumen en las figuras 11A y 11B.

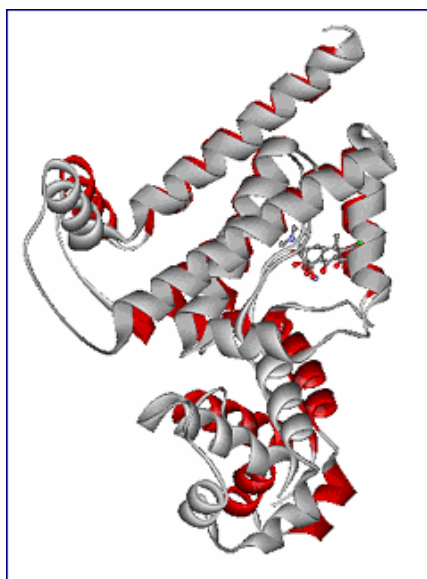


Figura 13. Superposición del monómero de TetR sin (gris) y con tetraciclina (rojo). Destacan las alteraciones inducidas en la región $\alpha 1$ - $\alpha 3$ implicada en la unión con el operador. El incremento de la distancia en $\alpha 3$ y $\alpha 3'$ con tetraciclina inhabilita a TetR para mantener las interacciones específicas mostradas en la Figura 11, por lo tanto el represor se libera. Figura basada en el trabajo de Orth *et al.* (Orth *et al.*, 2000).

El monómero A de TetR, se une a la cadena principal desde la posición -4 a la -7, mientras que con la cadena complementaria los contactos se dan entre la posición +4 a la +2, y el monómero simétrico, A', se une a la cadena principal desde la posición +2 a la +4 y a la complementaria desde la -4 a la -7 (Figuras 11A y 11B).

Los análisis cristalográficos revelaron que la hélice α -3 (que se extiende desde la Gln38 a la His44) es el elemento responsable principal en el reconocimiento específico, puesto que todos sus residuos contribuyen al mismo, excepto la Leu41 que forma parte del núcleo hidrofóbico que estabiliza la disposición y orientación correcta de las hélices α -1, α -2 y α -3. La Thr40 en el monómero A establece contacto directo con las bases T(-7) y C(-6) de la cadena principal del operador (Figuras 11A y 11B). El Trp43 también interacciona con la T(-7). La Pro39 interacciona con ambas cadenas, en concreto, con la T(-5) y la A(-4) de la cadena principal y con la T(-4) de la complementaria. En el resto del medio operador, la hélice α -3 del monómero A interacciona con la cadena complementaria, en concreto, la Tyr42 contacta con la T(+4) y la Gln38 con la A(+3).

La hélice α -2 proporciona un contacto específico adicional con la cadena complementaria y así la Arg28 contacta con la G(+2).

Aunque el dominio de unión a ADN de TetR mantiene su estructura gracias al núcleo hidrofóbico constituido por residuos de cada una de las tres hélices, α -1, α -2 y α -3, la interacción con el ADN causa cambios en dicho dominio. Uno de estos cambios

consiste en que la α -3 forma una vuelta de hélice 3_{10} en su extremo N-terminal como resultado de los contactos con el ADN.

Los puentes de hidrógeno entre Arg28 y G(+2), y Gln38 y A(+3) incrementan la separación entre los pares de bases 1 y 2, desde 3,4 Å a 3,9 Å (Orth *et al.*, 2000). Los dos grupos fosfato que flanquean a la G de la posición +2, establecen puentes de hidrógeno con las cadenas laterales de la Thr26, Thr27, Tyr42 y Lys48, y con el grupo amino de la cadena principal de la Thr27 y la Lys48 (Figura 11B). Estos contactos aproximan el ADN hacia TetR, cerca de la G(+2), generando una torsión del ADN que provoca una deformación en el área correspondiente a las posiciones +3 a +6 que compensa la desviación del ADN inducida en +2. Los estudios cristalográficos revelaron que la Lys48, localizada en α -4, también establecía contactos con el ADN (Figura 11B). La lisina en esta posición está muy conservada entre los miembros de la familia TetR, y se hipotetiza que este residuo juega un papel equivalente en otras proteínas de la familia.

6.4.2. El regulador QacR.

Dos dímeros de QacR se unen al operador para reprimir la expresión del gen qacA. QacA confiere resistencia frente a antisépticos catiónicos monovalentes y bivalentes y desinfectantes tales como los compuestos cuaternarios del amoníaco (el nombre *qac* deriva del inglés *quaternary ammonium compounds*) (Aramaki *et al.*, 1995a; Aramaki *et al.*, 1995b; Brown y Skurray, 2001).

El locus *qac* está constituido por los genes *qacA* y *qacR*, de transcripción divergente, y localizados en un plásmido (Grkovic *et al.*, 1998). En ausencia de drogas, la proteína de 188 residuos, QacR, mantiene reprimida la expresión del gen *qacA* que codifica al exportador multidroga, mediante la unión a dos palíndromos anidados, situados en el extremo 3' del promotor de *qacA* y solapantes con su punto de inicio de la transcripción (Grkovic *et al.*, 1998; Paulsen *et al.*, 1996). Por lo tanto, QacR parece reprimir la transcripción impidiendo la transición del complejo promotor/ARN-polimerasa hacia su estado transcripcional productivo, más que mediante bloqueo físico de la unión de la ARN-polimerasa.

La estructura tridimensional de QacR (identificadores PDB 1JTX, 1JTG, 1JTY, 1JUM, 1JUP, 1JUS, 1JTO) reveló que se trata de una proteína de carácter helical en su conjunto y que presenta un motivo de unión a ADN de tipo hélice-giro-hélice, situado en su región N-terminal, y un segundo dominio implicado en la unión de efectores y dimerización (Schumacher *et al.*, 2001; Schumacher *et al.*, 2002b). Hay que destacar que a diferencia de TetR, dos dímeros de QacR, en lugar de uno, se unen al sitio operador (Salah-Bey y Thompson, 1995; Sánchez *et al.*, 2002; Figura 14).

Los monómeros de cada dímero se han llamado proximal y distal para referirnos a su posición respecto al centro de simetría del operador palindrómico (Figura 15A y 15B). Se observó que la parte del operador a la que se unía un dímero, era simétrica y parcialmente solapante con la unida por el otro dímero (Schumacher *et al.*, 2002b; y Figuras 14 y 15). La existencia dentro del mismo fragmento de ADN de dos palíndromos parcialmente solapantes con idéntica simetría de bases, es llamativa (Figura 15A y 15B).

En este sentido, las secuencias palindrómicas reconocidas por cada dímero QacR son equivalentes a las descritas para la interfaz de TetR excepto por la longitud de los espaciadores, consistente en 3 pares de bases para TetR y 4 para QacR, apoyando la hipótesis de que las interacciones de otros miembros de la familia con sus secuencias operadoras puedan ser similares, independientemente del número de dímeros implicados, dado el tamaño de los dominios hélice-giro-hélice que se extienden una amplitud que abarca 7 pares de bases en el ADN.

La hélice α -3 de los monómeros distales de QacR_A y de QacR_B establece la mayoría de los contactos específicos con el operador (Schumacher *et al.*, 2002b). El residuo Tyr41 del monómero distal A (Figura 15) establece contactos hidrofóbicos con la base T(-10) de la cadena principal del ADN además de con el fosfato en la posición -11 de la misma cadena, mientras que la Tyr40 contacta con T(+7) (Figura 15B). Además, el acoplamiento estrecho con el ADN se facilita mediante puentes de hidrógeno específicos entre la Lys36 y la base G(+6) en la cadena complementaria, y entre la Gly37 en la base G(-8) en la principal. La Gly37 es importante en la represión pues el nucleótido con el que contacta, la G(-8), es el sitio de inicio de la transcripción del gen *qacA*.

Los monómeros proximales A y B también establecen una serie de interacciones críticas. Por ejemplo, la Tyr41 del monómero B proximal contacta con la C(-6) de la cadena principal, mientras que la Tyr40 contacta con la T(+3) y el fosfato (+2) de la complementaria (Schumacher *et al.*, 2002b). Hay que señalar al respecto la aparente intercambiabilidad T/C en el contacto específico, hidrofóbico, con un residuo de Tyr que también se observa en el co-cristal de TetR/operador. La Gly37 en el monómero A proximal contacta con la G(-4) en la cadena complementaria, mientras que la Lys36 contacta con la G(+1) en la cadena principal. Un hecho curioso es la diferencia de espaciado en las bases contactadas por Tyr40 y Lys36 en los monómeros proximal y distal, Figura 16. Un número de residuos de α 2, lazo α 2- α 3, α 3, y el dipolo positivo de la α 1 (N-terminal) también interaccionan con el esqueleto de fosfatos en ambas cadenas de ADN (Schumacher *et al.*, 2002b). La Figura 15B y 15C muestran cómo cada dímero se dispone sobre el surco mayor del ADN, de manera que cada uno queda casi en la cara opuesta al otro, formando un ángulo entre los ejes de ambos dímeros de menos de 180° (Figura 14).

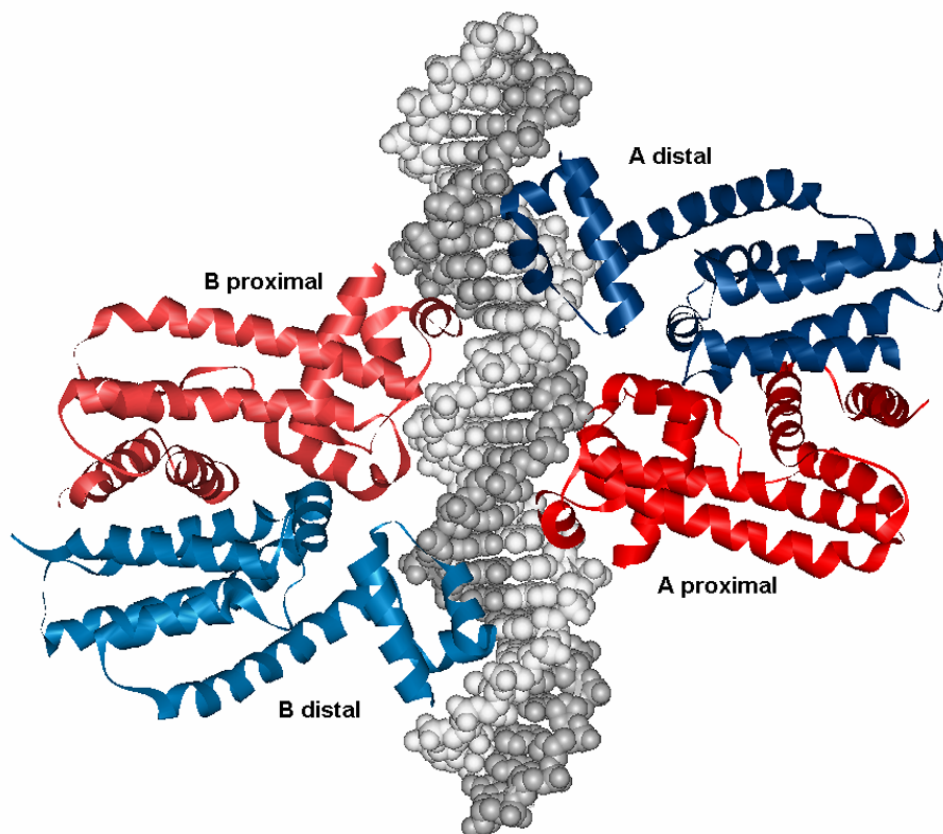


Figura 14. Diagrama de cintas de los dos monómeros QacR unidos a su secuencia de ADN diana representada en doble hélice (A) y detalles de los contactos establecidos por la α -hélice 3 de los monómeros de homodímeros diferentes cuando reconocen los sitios solapantes (B).

Los estudios de la unión de QacR al ADN han indicado que los dos dímeros se unen al ADN de manera cooperativa (Grkovic *et al.*, 2001a; Grkovic *et al.*, 2001b; Schumacher *et al.*, 2002b). Los análisis tridimensionales sugerían que tal cooperatividad no surge de interacciones proteína-proteína, pues el punto más próximo entre ambos dímeros dista 5.0 Å. Más bien la cooperatividad parece estar mediada por un cambio en la conformación del ADN desde la forma B a una configuración de bajo enrollamiento pero de alta afinidad, lo cual es patente en el análisis del cristal. Esta modificación de la conformación del ADN es necesaria porque la distancia óptima entre cada motivo hélice-giro-hélice del dímero QacR es de 37 Å. Esto requiere la expansión de los 34 Å que en ADN-B separan a las dos regiones de contacto sucesivas en el surco mayor. Este hecho ha sugerido que la unión del primer dímero QacR fuerza este cambio conformacional energéticamente desfavorable, que facilita la unión del segundo dímero (Schumacher *et al.*, 2002b). Los datos experimentales presentados por Grkovic *et al.* (2001b y 2003) sugerían que los dos dímeros deben unirse simultánea y cooperativamente al operador para ser capaces de mantener la deformación del ADN observada en el cristal.

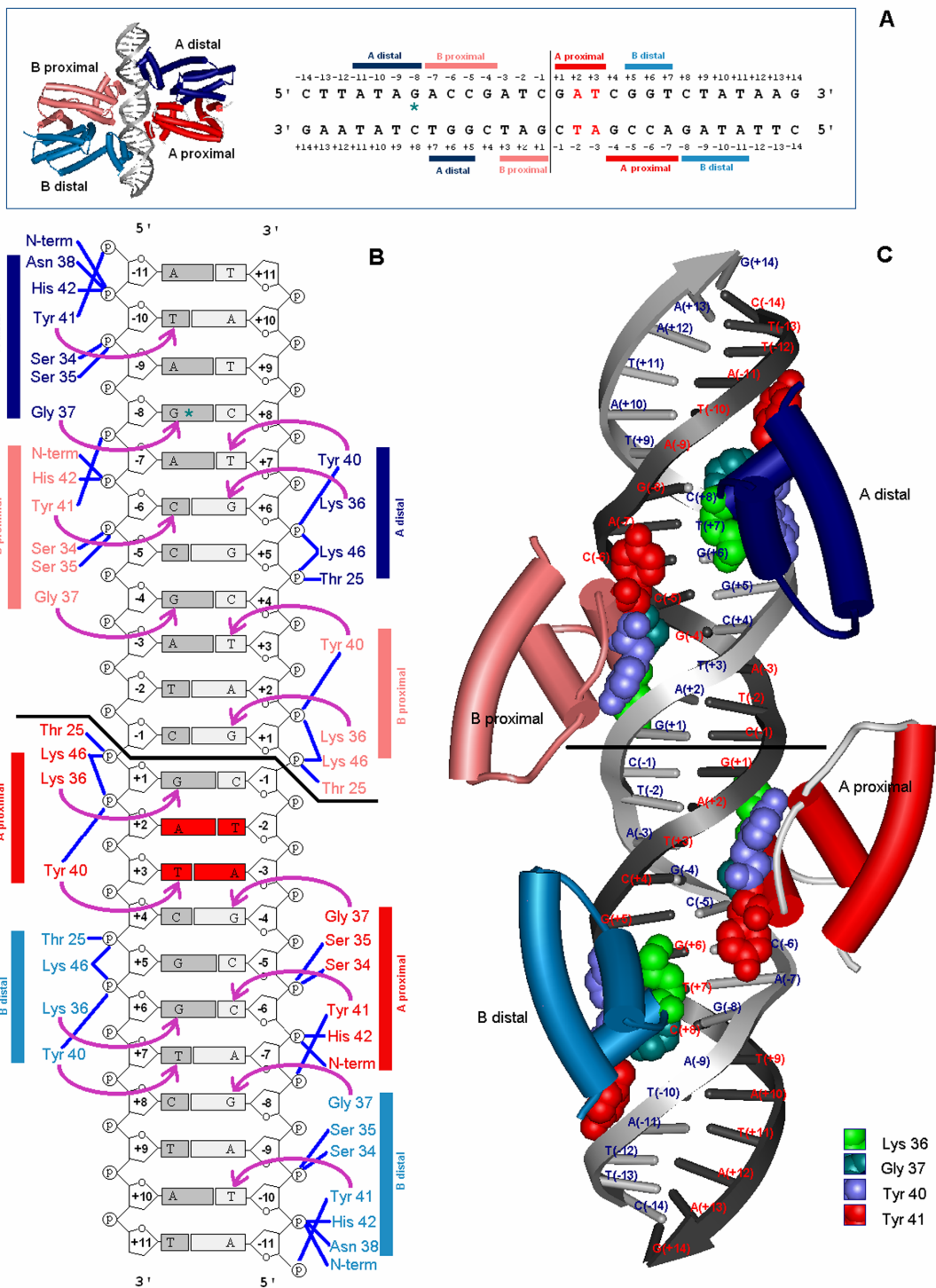


Figura 15. Unión de QacR a su operador. **A)** Interacción de QacR con el operador qac. **B)** Contactos de los residuos de la α -hélice 3 de los homodímeros QacR A y B con los nucleótidos específicos (flechas) y con los fosfatos del esqueleto del ADN (líneas azules) en el operador sintético usado para el co-cristal QacR_ADN (Schumacher *et al.*, 2002a; Schumacher *et al.*, 2001). **C)** Representación de los dos homodímeros QacR unidos al operador qac en una representación en doble hélice. Figura basada en la del trabajo de Schumacher *et al.*, 2001 y 2002b.

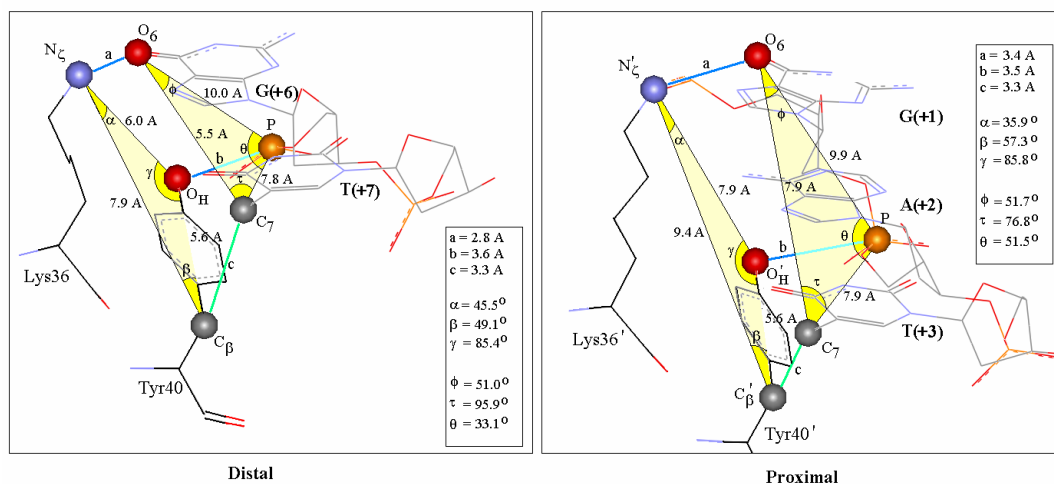


Figura 16. El par de residuos Lys36/Tyr40 de los monómeros distales y proximales aunque establecen contactos idénticos con las bases nitrogenadas se diferencian en el espaciado entre las bases contactadas. En los monómeros distales ambos residuos contactan con bases consecutivas, G(+6) y T(+7), mientras que en los monómeros proximales las bases contactadas, G(+1) y T(+3), están separadas por la A(+2). Esto obliga al par Tyr40/Lys36 a adoptar una disposición espacial algo diferente entre los monómeros distales y proximales, tal como se resume en la figura.

Shummacher y Brennan (2002a) observaron que TetR y QacR mostraban el mismo grado de especificidad de unión al ADN mediante mecanismos de unión ligeramente diferentes.

Ellos notaron que TetR emplea la Arg28, situada fuera de su hélice de reconocimiento, para el establecimiento de un contacto específico con una base (Orth *et al.*, 2000), mientras que QacR no emplea residuos fuera de la hélice $\alpha 3$ para establecer contactos específicos con el ADN. Estos autores también observaron que TetR deforma su sitio de unión induciéndole una curvatura de 17° hacia la proteína para optimizar la posición de sus motivos hélice-giro-hélice para las interacciones específicas con las bases en cada medio operador. Por su parte QacR ensancha ligeramente el surco mayor en todo el sitio de unión IR1 y lo curva sólo unos 3°. Estas particularidades se reflejan en las diferentes distancias entre los centros de los dominios hélice-giro-hélice observados en QacR (37 Å).

Así, una importante conclusión que se deriva de las comparaciones entre las estructuras de los complejos QacR-ADN y TetR-ADN es que incluso proteínas de la misma familia homólogas estructuralmente y que comparten una función similar, la represión, pueden utilizar mecanismos de acción ligeramente diferentes.

QacR como un modelo para el reconocimiento de múltiples efectores por un regulador. QacR se libera del operador de *qacA* mediante su interacción con varios efectores de naturaleza catiónica y lipofílica tales como la rodamida 6G, el cristal violeta y el bromuro de etidio (Grkovic *et al.*, 1998). Más recientemente, Grkovic *et al.*,

(2003) demostraron que otras sustancias tales como varios colorantes catiónicos bivalentes y alcaloides de origen vegetal, actuaban como efectores de QacR.

Pese a la existencia de dos bolsillos de unión, sólo se une una molécula de efector por homodímero, lo cual se demostró mediante estudios de equilibrio de diálisis y calorimetría utilizando el complejo QacR-R6G (Schumacher *et al.*, 2001). El cristal de QacR unido a distintos efectores reveló otro hecho destacable respecto al bolsillo de unión del efector, se trata de una cavidad con un volumen de unos 1100 Å³, capaz de adaptarse y alojar a moléculas diferentes en distintas partes del bolsillo (Schumacher *et al.*, 2002a y 2002b). Yu *et al.* (2003a y 2003b) han descrito una cavidad similar para el transportador multidroga AcrB, capaz de unir distintas moléculas orgánicas de estructuras muy diferentes.

Los estudios cristalográficos de Schumacher *et al.* (2001) y Murray *et al.* (2004) demostraron que el reconocimiento "multidroga" mediado por el dímero de QacR es un proceso más bien simple que, al contrario de lo esperado, no requiere de mecanismos moleculares sofisticados. De hecho, el dominio de unión a efector de QacR consta de 6 α -hélices (identificadores PDB: 1JTX, 1JT6, 1JTY, 1JUP, 1JUS, 1JTO, 1RKW, 1RPW). La entrada al bolsillo de unión de efector se realiza a través de una pequeña abertura formada por la divergencia de las hélices α_6 , α_7 , α_8 y α_8' . La estequiometría de una molécula de efector por dímero de QacR conduce a un proceso de inducción asimétrico en el que el monómero unido a la molécula de efector sufre el cambio estructural mayor.

La comparación de la estructura ligada a efector con la ligada a ADN revela que la unión de la droga desencadena un desenrollamiento de los residuos 89-93, lo que hace que se extienda la hélice α_5 en una vuelta. Esta transición libera los residuos Tyr92 y Tyr93 que actúan como sustitutos de la molécula de efector, del núcleo hidrofóbico de la proteína. La expulsión de ambos residuos de Tyr conlleva la recolocación de la hélice α_6 adyacente y del dominio de unión de ADN, ligado a ella. El resultado final de esta transición estructural es una translación de 9 Å y una rotación de 37° del dominio de unión a ADN, lo que incapacita al dímero QacR para unirse a su ADN diana.

6.4.3. La estructura tridimensional de CprB.

Las bacterias gram positivas pertenecientes al género *Streptomyces* utilizan las γ -butirolactonas como autorreguladores u hormonas bacterianas, junto con sus receptores específicos (receptores de γ -butirolactonas), para controlar sus procesos de diferenciación morfológica, producción de antibióticos, o ambos (Hinrichs *et al.*, 1994; Hirata *et*

al., 2004). La γ -butirolactona más representativa es la 2-isocaprilol-3R-hidroximetil- γ -butirolactona producida por *Streptomyces griseus*.

Esta butirolactona se conoce como factor-A y es esencial para la formación del micelio aéreo, la producción de estreptomycin y la producción de un pigmento amarillo, procesos en los que juega un papel esencial la proteína ArpA (Hara y Beppu, 1982a; Hara and Beppu, 1982b; Horinouchi y Beppu, 1990).

ArpA, no se ha podido purificar debido a una cierta insolubilidad intrínseca. Al contrario, la proteína CprB de *S. coelicolor* A3(2), con una identidad de secuencia del 30% con ArpA (Onaka *et al.*, 1998), se ha purificado y cristalizado (Natsume *et al.*, 2004), aunque el ligando específico de CprB es aún desconocido. No obstante, CprB se une a la misma secuencia de nucleótidos que ArpA (Sugiyama *et al.*, 1998) y lo que es más, CprB también actúa como un regulador negativo tanto del metabolismo secundario como de la morfogénesis en *S. coelicolor* al igual que ArpA en *S. griseus* (Natsume *et al.*, 2004; Onaka *et al.*, 1998).

CprB se presenta en dímeros de morfología semejante a una letra omega. Las dos subunidades del dímero se relacionan entre sí mediante un eje pseudo-binario (Figura 17). Cada monómero de CprB se compone de 10 α -hélices y tiene dos dominios, uno de unión a ADN (residuos 1-52) y otro regulador (residuos 77-215) (Figuras 17 y 18).

La estructura tridimensional de CprB es, en esencia, similar a la de QacR unido al ADN, excepto por la hélice α 10 (Schumacher *et al.*, 2001; Schumacher *et al.*, 2002b). Además, los dominios de unión a ADN en las dos proteínas son muy similares, tanto que ambos pueden superponerse con una desviación rms de 1.48 Å para 71 átomos de $C\alpha$ (Natsume *et al.*, 2004).

Aunque se carece de información específica sobre las interacciones de CprB con su operador, el grado de conservación de secuencia tan elevado con TetR y QacR permitió a los autores predecir que el núcleo del dominio de unión a ADN estaba compuesto por Ile14, Ile15, Ala18, Phe22, Leu32, Leu46 y Phe50 (Bennik *et al.*, 2000). Se ha sugerido que dos dímeros CprB se unen a la secuencia de ADN como lo observado para el complejo QacR-ADN (Grkovic *et al.*, 2001a; Grkovic *et al.*, 2001b; Natsume *et al.*, 2004; Schumacher *et al.*, 2001; Schumacher *et al.*, 2002b). La base para esta hipótesis parte del buen alineamiento de las secuencias primarias de los dominios de unión a ADN de CprB y QacR. Esto indicaría una relación evolutiva entre los dominios de unión a ADN de ambas proteínas.

El dominio regulador de CprB está constituido por 6 α -hélices (α 5- α 10) (Natsume *et al.*, 2004), y también se puede superponer sobre el dominio equivalente de QacR (PDB: 1JT0).

6.4.4. La estructura tridimensional de EthR.

La etionamida se ha empleado durante más de 30 años como tratamiento quimioterápico de segunda línea en pacientes con tuberculosis que han desarrollado resistencia frente a antibióticos de primera línea tales como la isoniazida y la rifampicina. La activación del profármaco etionamida está regulada mediante la monooxigenasa Bayer-Villinger EthA y el represor EthR de la familia TetR, cuyos marcos de lectura abierta están separados por 75 pares de bases en el genoma de *Mycobacterium tuberculosis*.

La expresión de *ethA* en *M. tuberculosis* está regulada por EthR. EthR reprime la transcripción del gen activador *ethA*. La sobreexpresión de *ethR* provoca resistencia a la etionamida, mientras que la inactivación cromosómica de *ethR* promueve la hipersensibilidad a la etionamida (Baulard *et al.*, 2000). EthR se une directa y específicamente a secuencias de ADN que corresponden a la región intergénica *ethRA* (Baulard *et al.*, 2000; Engohang-Ndong *et al.*, 2004). El operador de EthR abarca 55 pares de bases, siendo mucho mayor que los operadores de 15 pares de bases reconocidos por otros miembros de la familia. El operador se organiza como un palíndromo degenerado que contiene pares de secuencias invertidas solapadas y repetidas en tándem (Engohang-Ndong *et al.*, 2004).

En ausencia de ADN, EthR forma homodímeros en solución, y las medidas de SPR (*surface plasmon resonance*) sugieren que EthR octameriza cuando se une al ADN (Engohang-Ndong *et al.*, 2004).

El monómero de EthR está formado por 9 α -hélices (Figuras 17 y 18), organizadas en dos dominios (Dover *et al.*, 2004). El dominio N terminal comprende las 3 primeras α -hélices, con la segunda y tercera formando el motivo hélice-giro-hélice de unión a ADN. El dominio C terminal, supuestamente implicado en la unión a ligando, incluye las 6 α -hélices restantes, desde la quinta a la novena, y su función en EthR aún no se conoce. La cuarta α -hélice actúa de conector entre ambos dominios. La estructura del cristal revela que la interfase de dimerización, una característica estructural conservada entre los miembros de la familia TetR (Orth *et al.*, 2000; Schumacher *et al.*, 2002b), está constituida primariamente por las hélices 8 y 9.

Una de las características más llamativas de la estructura de EthR es la cavidad estrecha en forma de túnel (Figura 17) a la que dan lugar las α -hélices 4, 5, 7, y 8 abriéndose el túnel en la parte inferior del mismo (Dover *et al.*, 2004). El túnel tiene unos 20 Å de longitud y está revestido predominantemente, aunque no exclusivamente, por residuos aromáticos, en su mayoría pertenecientes a las hélices 5 y 7. El lazo que conecta las hélices 4 y 5 restringe la apertura de este túnel hidrofóbico. Dicho lazo, tal como indica la baja definición de su densidad electrónica, posee cierto grado de flexibilidad estructural, tal como sería de esperar. Dicha cavidad sería el sitio de unión para un ligando, o ligandos, aún desconocidos.

6.4.5. Estructura de los miembros de la familia TetR cristalizados y de funciones desconocidas.

Los nuevos métodos en genómica/proteómica están permitiendo la cristalización de proteínas, muchas de las cuales aún no tienen una función asignada. Las proteínas de la familia TetR pertenecientes a este grupo son: Cgl2612 de *Corynebacterium glutamicum* (pdb 1V7B); YbiH de *S. typhimurium* (pdb 1T33); YcdC de *Escherichia coli* (pdb 1PB6) y YfiR de YsiA de *B. subtilis* (pdb 1RKT y 1VIO, respectivamente) (Figuras 17 y 18).

6.5. Características estructurales de las proteínas de la familia TetR.

Del estudio conjunto de la estructura de los miembros de la familia TetR cristalizados en el momento de la realización de este estudio, se deduce una serie de características generales.

Todos son proteínas de tipo α -helical (constituidas sólo por α -hélices y los lazos y giros que las unen) que se presentan en homodímeros y en todos los monómeros distinguimos una región de unión a ADN con tres α -hélices ($\alpha 2$ y $\alpha 3$ formando parte del dominio hélice-giro-hélice) y un núcleo regulador, con un bolsillo de unión a inductor y una superficie de dimerización, constituida por dos α -hélices; la hélice α -4 es el conector entre ambos (Figura 17).

Podemos diferenciarlos en dos grupos según posean monómeros constituidos por 10 o por 9 α -hélices.

Con 10 α -hélices tenemos a TetR y a CprB. En realidad CprB tiene una estructura básica de 9 α -hélices, pero posee una décima α -hélice adicional, la última, tras las dos de dimerización, $\alpha 8$ y $\alpha 9$, mientras que en TetR, la hélice “adicional” se dispone entre las dos de dimerización $\alpha 8$ y $\alpha 10$ contribuyendo a la delimitación del bolsillo de unión a efector en el monómero compañero y que se desplaza cerrando dicha cavidad tras la entrada del efector (Figura 18).

El resto de los reguladores que se han cristalizado, QacR, EthR, YcdC, YbiH, Cgl2612, YsiA e YfiR, poseen monómeros con una organización típica en 9 α -hélices con la superficie de dimerización constituida por las 2 últimas $\alpha 8$ y $\alpha 9$. Si bien esto es cierto para QacR, YsiA, YfiR y Cgl2612, no ocurre lo mismo con las tres restantes, en las que un análisis más detallado muestra ciertas particularidades (Figura 18).

Podríamos decir que YcdC posee once α -hélices. Una hélice adicional está formada por la división de la $\alpha 7$ en dos más pequeñas, $\alpha 7$ y $\alpha 8$. La superficie de dimerización está formada por las 2 últimas $\alpha 9$ y $\alpha 10$, sin embargo $\alpha 9$ en su región C terminal, en los residuos Asp174 y Phe175, sufre una inflexión que dispone casi perpendicularmente a sus dos últimas vueltas respecto del eje del fragmento mayor de la hélice. De esta manera aparece una hélice “extra”, corta, $\alpha 10$, situada entre las 2 hélices de dimerización que recuerda a la α -hélice 9 de TetR (Figura 18).

La organización del monómero de YbiH es idéntica a la anterior. En este caso la inflexión en la primera hélice de dimerización se debe a un solo residuo la Gly182 que permite el cambio brusco de dirección del las dos últimas vueltas C-terminales de dicha hélice (Figura 18).

EthR tendría en realidad 8 α -hélices, carecería de la $\alpha 6$ propiamente dicha que constituye el lateral inferior externo del bolsillo de unión a efector en los otros monómeros. La $\alpha 5$ sufre una inflexión en su His117 que da a su mitad C-terminal la orientación adecuada para suplir a la hélice $\alpha 6$ (Figura 18).

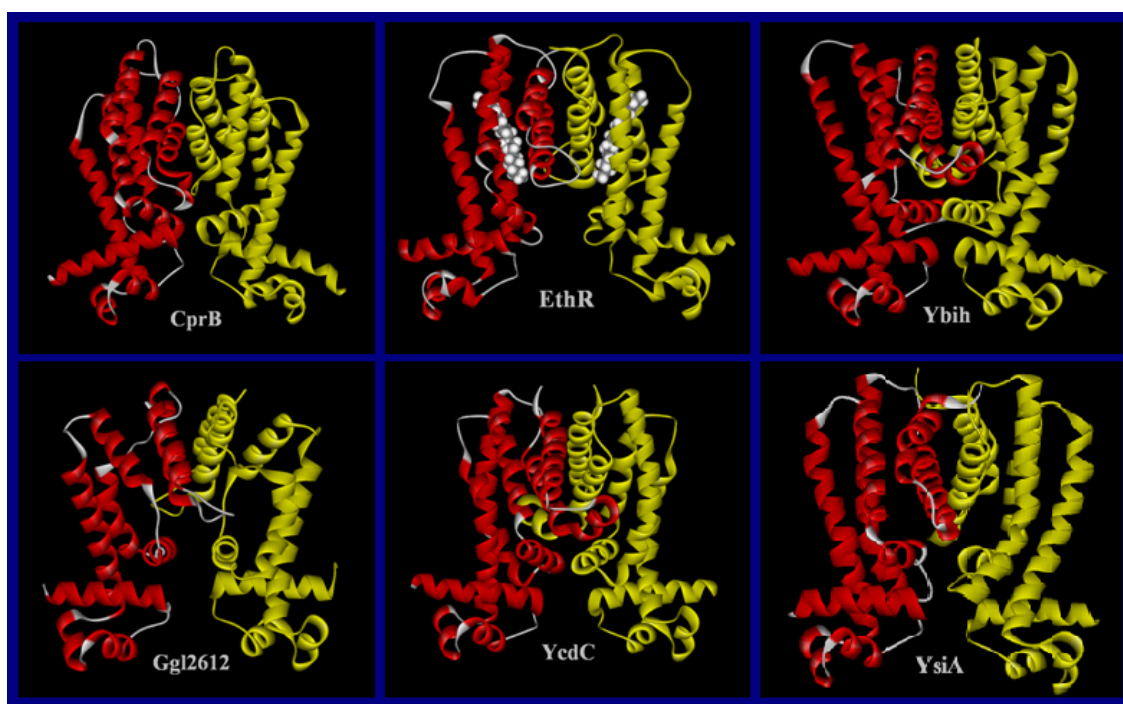


Figura 17. Estructura de otros miembros de la familia TetR cristalizados. Los miembros de la familia TetR son proteínas tipo α -helical que se presentan como homodímeros, cada uno con dos dominios de unión a ADN y dos regiones reguladoras, ambos relacionados por la α -hélice 4. Los dominios de unión a ADN están formados por las tres primeras α -hélices y los dominios reguladores por el resto de las α -hélices a partir de la quinta.

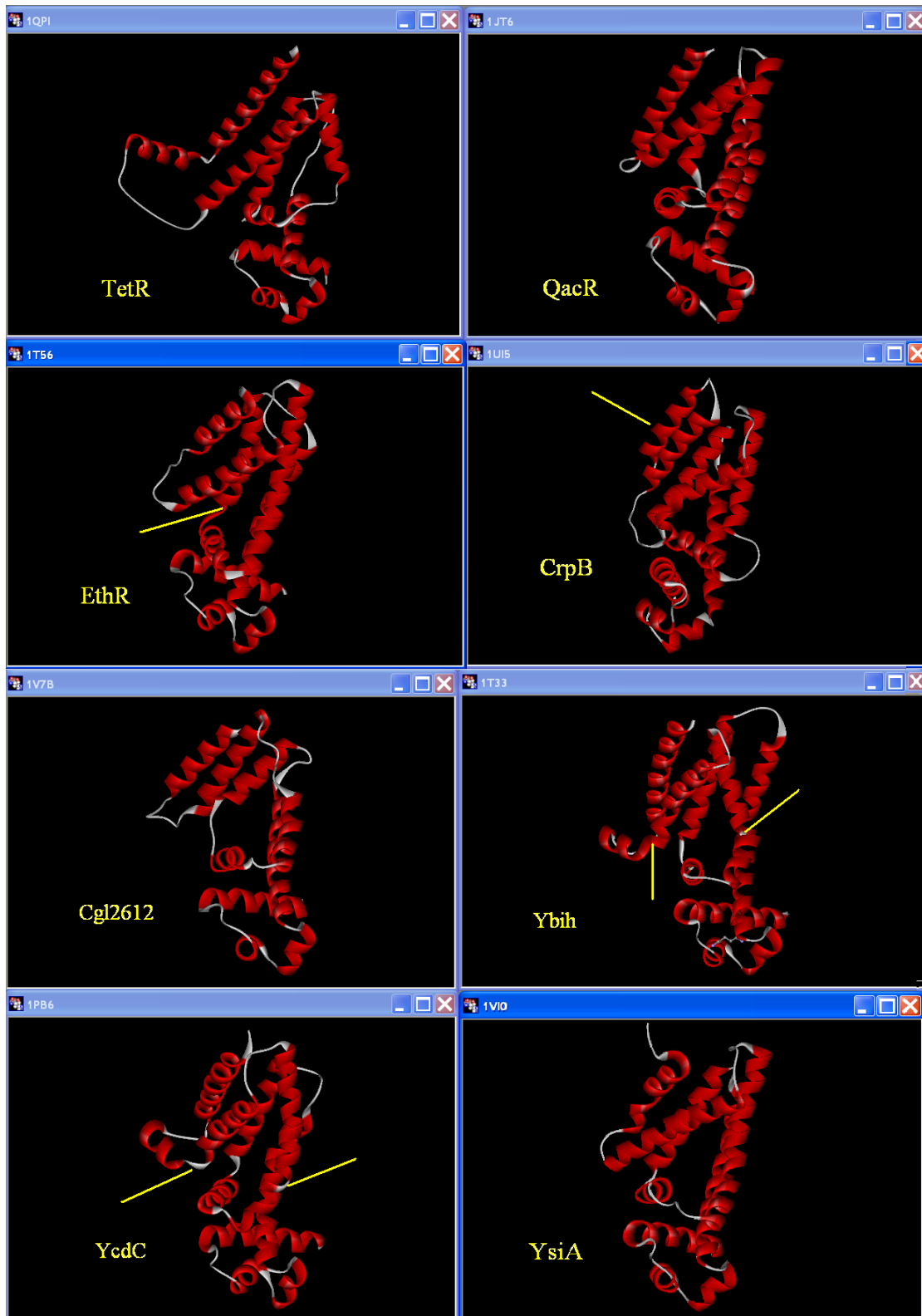


Figura 18. Estructura de los monómeros de los miembros de la familia TetR cristalizados. Se observa el dominio de unión a ADN en la zona inferior de cada monómero y la región reguladora en la que además de la cavidad de unión a efector, en posición central, se aprecian las dos α -hélices, paralelas o subparalelas que constituyen la superficie de dimerización. Con trazos amarillos se marcan las particularidades de algunos monómeros que se discuten en el texto.

6.6. Predicción de unión a ADN para las proteínas de la familia TetR basadas en el estudio de las estructuras de los cristales de TetR y QacR

Existe un solapamiento perfecto de los dominios de unión a ADN de QacR, TetR, CrpB y EthR, y no fue necesario introducir gaps en las alfa hélices implicadas en el contacto con el ADN en el alineamiento múltiple de los miembros de la familia TetR en este dominio. Esto se usó de base para hipotetizar que los residuos en la misma posición en el alineamiento múltiple de todos los miembros de la familia podrían jugar papeles equivalentes. Esto nos llevó a analizar cada aminoácido en el alineamiento múltiple del dominio de unión a ADN.

6.6.1. Estableciendo relación entre la posiciones en el *perfil* y la posición estructural.

El análisis comparativo de los co-cristales de QacR y TetR con sus correspondientes operadores revelaron que los residuos en las posiciones 22, 33, 34, 35, 37, 38, 39 y 43 estaban implicados en interacciones con el ADN operador (Figura 8). Se analizó la presencia de cada aminoácido en estas posiciones en el alineamiento múltiple de los miembros de la familia TetR (Tabla 31).

Encontramos dos tipos de posiciones, una en la que el residuo estaba muy conservado y otra en la que la conservación, en caso de existir, era baja. Las posiciones 37, 39 y 43 fueron las más conservadas, mientras que las posiciones 22, 33, 34, 35 y 38 alineaban diferentes residuos.

En la *posición 37* del perfil, correspondiente a la Tyr42 de TetR y la Tyr40 de QacR (Figura 19), el residuo Tyr aparece en el 74.16% de las proteínas alineadas (Tabla 31). El siguiente residuo en conservación en esta posición corresponde al otro aminoácido aromático, la fenilalanina (8%) y a continuación la histidina (4%) (Tabla 31). La Tyr-42 de TetR y la Tyr-40 de QacR se localizan en el centro de la α -hélice 3 y contactan con una timina localizada en el centro del palíndromo que constituye el operador, y además establece un contacto con el fosfato situado una posición más hacia el centro del palíndromo (Figuras 11B y 15B).

El residuo en la *posición 39* del perfil correspondía con la His44 en TetR y la His42 en QacR (Figura 19). Ambos co-cristales mostraban que estas His establecían contactos con el esqueleto de fosfatos (Figura 11B y 15B). El alineamiento múltiple de los miembros de la familia muestra que en la posición 39 aparece His o Tyr en el 78.33% de las secuencias alineadas. Resulta pues tentador proponer que este residuo sea crítico en las interacciones con la cadena de fosfato.

Pos 22	Pos 33	Pos 34	Pos 35	Pos 37	Pos 38	Pos 39	Pos 43
V 21.95	K 29.26	G 37.79	T 34.3	Y 74.16	R 22.75	H 44.5	K 77.25
L 20.6	R 20.07	A 18.93	S 20.87	F 8.05	Y 16.44	Y 33.83	R 10.07
M 18.66	P 10.27	P 12.55	A 19.6	H 4.63	H 13.29	R 4.9	L 2.82
I 17.65	Q 6.04	S 8.39	L 5.84	L 3.09	N 8.39	F 3.56	I 1.88
T 13.22	V 5.3	R 6.71	N 5.3	T 2.48	K 6.11	A 2.15	M 1.88
F 2.15	A 4.56	T 5.1	G 4.3	S 1.95	W 5.84	N 1.95	V 1.81
H 1.74	T 4.3	Q 3.56	V 2.28	N 1.88	A 5.64	Q 1.88	T 0.94
A 1.61	L 4.23	M 2.35	M 2.08	R 1.21	L 3.56	E 1.48	G 0.94
Y 1.28	E 3.36	N 1.07	Q 1.34	Q 0.6	S 3.56	L 1.34	Q 0.74
S 0.54	I 2.89	K 1.01	Y 1.28	A 0.47	F 2.89	W 1.21	A 0.47
P 0.4	H 2.89	V 0.81	I 1.14	I 0.47	Q 2.55	S 0.94	H 0.34
N 0.13	S 2.01	D 0.47	P 0.54	G 0.27	T 2.35	T 0.74	F 0.27
E 0.07	N 1.61	L 0.47	R 0.4	M 0.27	E 1.68	V 0.54	P 0.2
	G 1.07	E 0.4	E 0.2	V 0.2	V 1.61	C 0.4	S 0.13
	D 1.01	F 0.27	H 0.2	K 0.13	G 1.41	K 0.27	Y 0.07
	Y 0.67	I 0.07	K 0.13	C 0.07	D 1.07	I 0.13	E 0.07
	M 0.27	H 0.07	D 0.13	P 0.07	I 0.54	D 0.07	N 0.07
	C 0.13		C 0.07		C 0.2	G 0.07	
	F 0.07				M 0.07		
					P 0.07		

Tabla 31. Frecuencia de aminoácidos en cada una de las posiciones críticas para el reconocimiento del operador en los miembros de la familia TetR. (Las frecuencias de aminoácidos se expresan en porcentajes y están calculadas con los 2353 miembros de la familia TetR).

Se ha establecido que existe interacción entre la Lys-48 de TetR y la Lys-46 en QacR, y el esqueleto de fosfato del ADN. Estos residuos corresponden a la *posición 43* del alineamiento múltiple, y la lisina se localiza en el extremo amino de la α -hélice 4 de cada proteína (Figura 19). En el 77% de las secuencias aparece un residuo de lisina, y su interacción con el esqueleto de fosfato podría ser crucial para ajustar el dominio hélice-giro-hélice en su posición correcta sobre el ADN (Figura 11B y 15B).

En la *posición 22* del alineamiento múltiple (Thr27 en TetR y Thr25 en QacR) los cinco residuos más abundantes son Val, Leu, Met, Ile y Thr. La Thr27 en TetR y la Thr25 en QacR (Figura 19) están implicadas en interacciones con el esqueleto de fosfato.

En la familia TetR, los contactos de los residuos alineados en la posición 37, en el centro de la α -hélice 3 (Tyr en el 74% de los casos), 39, también en la α -hélice 3 (His o Tyr en el 98% de los casos) y el residuo 43, en la α -hélice 4 (Lys en 77% de los casos) (Figura 19), probablemente actúen orientando al motivo hélice-giro-hélice para interaccionar de forma adecuada con el surco mayor del ADN, anclando a la proteína al esqueleto de fosfato.

El papel de la Gly en la posición 16 del alineamiento múltiple, situada al fin de la α -hélice 1 y muy conservada, podría ser el de orientar la cadena polipeptídica en los cristales de TetR y QacR, permitiendo que el dominio hélice-giro-hélice quede dispuesto en la orientación apropiada para interaccionar con el ADN (Figura 20).

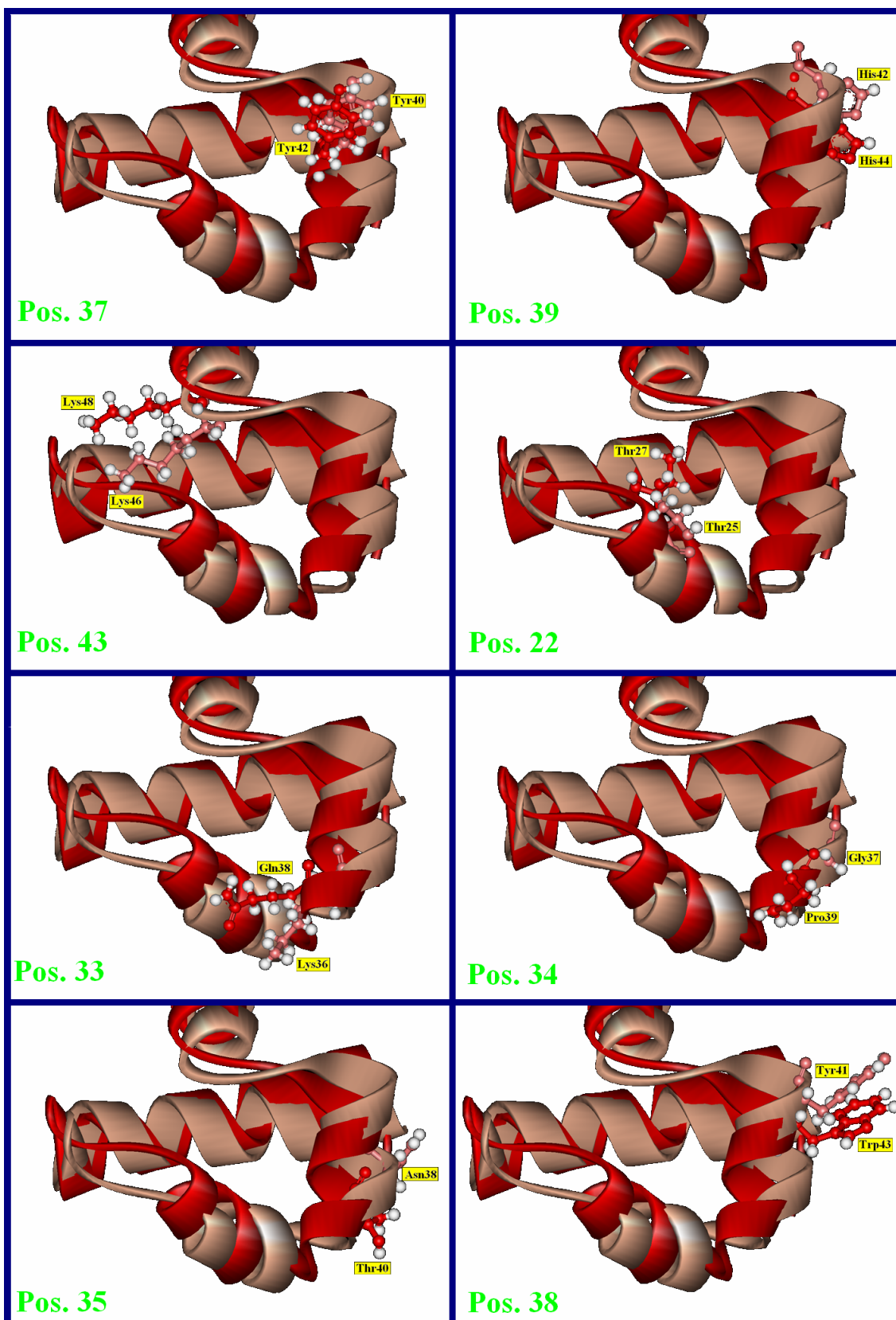


Figura 19. Superposición de los dominios de unión al operador de TetR(rojo) y QacR(salmón), mostrando los residuos de interacción con el ADN implicados en cada posición descrita en el texto.

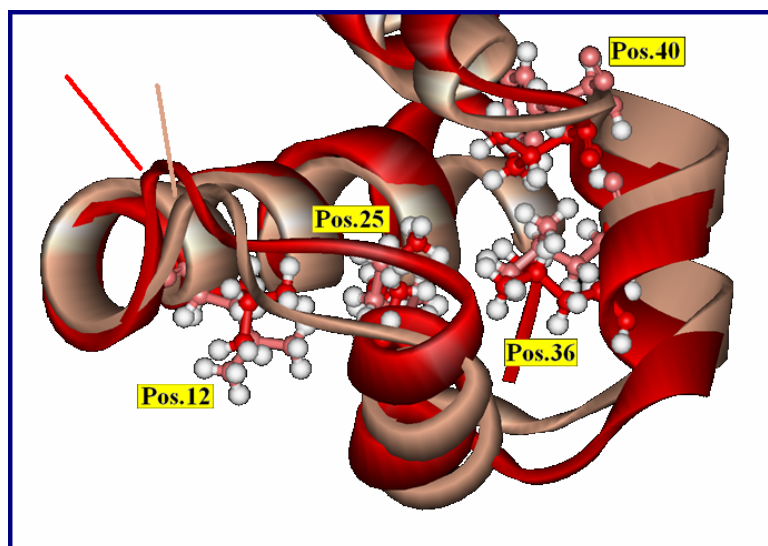


Figura 20. Superposición de los dominios de unión al operador de TetR (rojo) y QacR (salmón), mostrando los residuos conservados en las posiciones del perfil no implicadas en contactos con el ADN (no aparecen en la Tabla 31). Se trata de residuos hidrofóbicos orientados hacia el interior del núcleo del dominio que contribuyen al mantenimiento de la disposición espacial de las α -hélices, necesaria para la interacción con el surco mayor del ADN. Las Gly en posición,16 citadas en texto, están marcadas por sendas líneas al final de la α 1.

En las *posiciones 33, 34, 35 y 38* del alineamiento múltiple se encontraron distintos residuos (Tabla 31). En TetR y QacR los residuos correspondientes a esas posiciones establecían contactos específicos con diferentes bases del ADN, excepto la Asn38 de QacR (posición 35 en alineamiento múltiple) que contacta con el esqueleto de fosfato. Basándonos en la alta variabilidad de residuos en estas posiciones, se propone que estas posiciones dotan de especificidad a cada proteína de manera que puedan reconocer su propio operador mediante interacciones con bases nitrogenadas específicas.

6.6.2. Predicción de los residuos de unión a ADN en los otros miembros de la familia con estructura resuelta.

Si representamos los residuos de las ocho posiciones de contacto con el ADN consecutivamente, éstos forman "palabras" de 8 letras. El análisis de esas palabras en el alineamiento múltiple de la región del perfil de todas las secuencias, nos mostraría lo que podríamos considerar como "PALABRAS DE UNIÓN AL ADN" propias de los miembros de la familia TetR. Cuanto más similares son dos secuencias mayor es la probabilidad de que compartan la misma palabra.

La Tabla 32 (Anexo II) recoge las palabras de unión a ADN de los miembros de la familia TetR. Podemos diferenciar en ella dos tipos, las representadas por una sola

secuencia y las que son comunes a un conjunto de secuencias con una similitud global elevada. En los casos en los que se conozca la estructura tridimensional de un miembro de uno de estos conjuntos, dado el grado de conservación de la disposición espacial que exhiben los dominios de unión a ADN que conocemos, podríamos hipotetizar que los miembros del grupo no solo comparten la palabra (estructura primaria), sino también su disposición tridimensional que se muestra en el cristal.

Así por ejemplo en el co-cristal de QacR con su secuencia operadora tenemos la palabra TKGNY^YHK, que en la Tabla 32 (Anexo II) aparece con el número 56 y que está presente en un conjunto de secuencias anotadas como QacR de diversos *Staphylococcus* (*S. aureus*, *S. epidermidis*, *S. haemolyticus*), de manera que el conocimiento acerca de los residuos que contactan con el ADN que se deduce del co-cristal podría extenderse a las secuencias que comparten palabra.

El razonamiento equivalente sería aplicable a la palabra TQP^TYWHK, la número 379 de la Tabla 32 (Anexo II), que corresponde a la secuencia TetR clase D, cuyo co-cristal hemos estudiado, además de compartirla las clases B, E, H (*E. coli*, *S. enterica*, *S. typhi*, *Shigella flexneri*, *Proteus mirabilis*, *Pasteurella multocida*, *P. aerogenes*, *Neisseria meningitidis*, *Vibrio salmonicida*,...).

En el grupo de secuencias de TetR de la Tabla 32 (Anexo II) encontramos otras 3 palabras de interacción con ADN: TQP^AYWHK (nº 377) que corresponde a las variantes TetR clases B, C y G; TQP^SYWHK (nº 378) TetR(31) y TQP^GYWHK (nº 380) TetR de *Corynebacterium glutamicum*. Igual que el caso anterior, los datos sobre los residuos que contactan con el ADN que se deducen del co-cristal podrían extenderse a las secuencias que comparten palabra, puesto que las sustituciones del residuo en la posición 35 (T, A, S, G), probablemente implicado en las diferencias de especificidad de unión a su operador en cada clase de TetR, no sea tan drástica como para suponer una variación en la estructura 3D.

En la Figura 21 se representan los dominios de unión a ADN de CprB, EthR, Cgl2612, YbiH, YcdC e YsiA. En ellos se marcan los residuos correspondientes a las posiciones hipotéticas de interacción con sus operadores, deducidas de los co-cristales TetR y QacR.

La Tabla 32 (Anexo II) muestra que la palabra LKGAYFHK de CprB de *Streptomyces coelicolor* (nº 11), la comparten CprL y CprA del mismo organismo y BarZ de *S. virginiae* y que se agrupa junto a otras 2 palabras casi idénticas, IKGAYFHK (nº 12) de TylP de *S. fradiae*, ArpA de *S. griseus*, TarA de *S. tendae*, FarA de *Streptomyces* sp., y VKGAYFHK (nº 13) del receptor de butanolida de *S. virginiae* entre otros sin anotación funcional. La palabra de CprB constituiría un modelo de la interacción con el ADN para

los receptores de γ -butirolactonas, pues al igual que discutíamos para TetR, las sustituciones del residuo en la posición 22 (I, L y V), no supongan una variación en la estructura tridimensional del dominio de unión a ADN. Recordemos que CprB une la misma secuencia de nucleótidos que ArpA, con I en lugar de L en la posición 22.

Un grupo de secuencias de *Mycobacterium leprae* y *M. tuberculosis*, entre las que se encuentra EthR comparten la palabra VRPTYFYK, nº 679 de la Tabla 32.

Razonamientos paralelos podemos hacer con el resto de los miembros de la familia cuya estructura se ha resuelto, pero con función no conocida, Cgl2612 YKSGIYHR (nº 771), YbiH TIAATYYK (nº 831), YcdC LKTNLYYK (nº 669), YsiA VDGTYIYK.

6.6.3. Predicción de los residuos de unión a ADN en miembros de la familia sin estructura resuelta.

Hemos de destacar algunas palabras compartidas por grupos amplios de secuencias que constituyen ejemplos tipo dentro de la familia TetR.

Por ejemplo LRGAYWHK, la palabra nº 1, compartida por el grupo más amplio de secuencias de la tabla, anotadas como: AcrR (*E. coli*, *Shigella flexneri*, *Klebsiella pneumoniae*, *Enterobacter aerogenes*, *S. typhimurium*, *S. typhi*, *Y. pestis*); EnvR (*E. coli*, *S. typhi*); MtrR (*Neisseria meningitidis* y *N. gonorrhoeae*); TtgR, TtgW, ArpR, SrpR (*P. putida*); además de otras muchas secuencias sin anotación funcional.

Otro ejemplo remarcable lo constituye la palabra nº 822 HVATFNRYR, compartida por: LuxR (*V. harveyi*), LitR (*V. fischeri*), VanT (*V. anguillarum*), OpaR (*V. parahaemolyticus*), ScmR y VvpR (*V. vulníficus*) y HapR (*V. cholerae*).

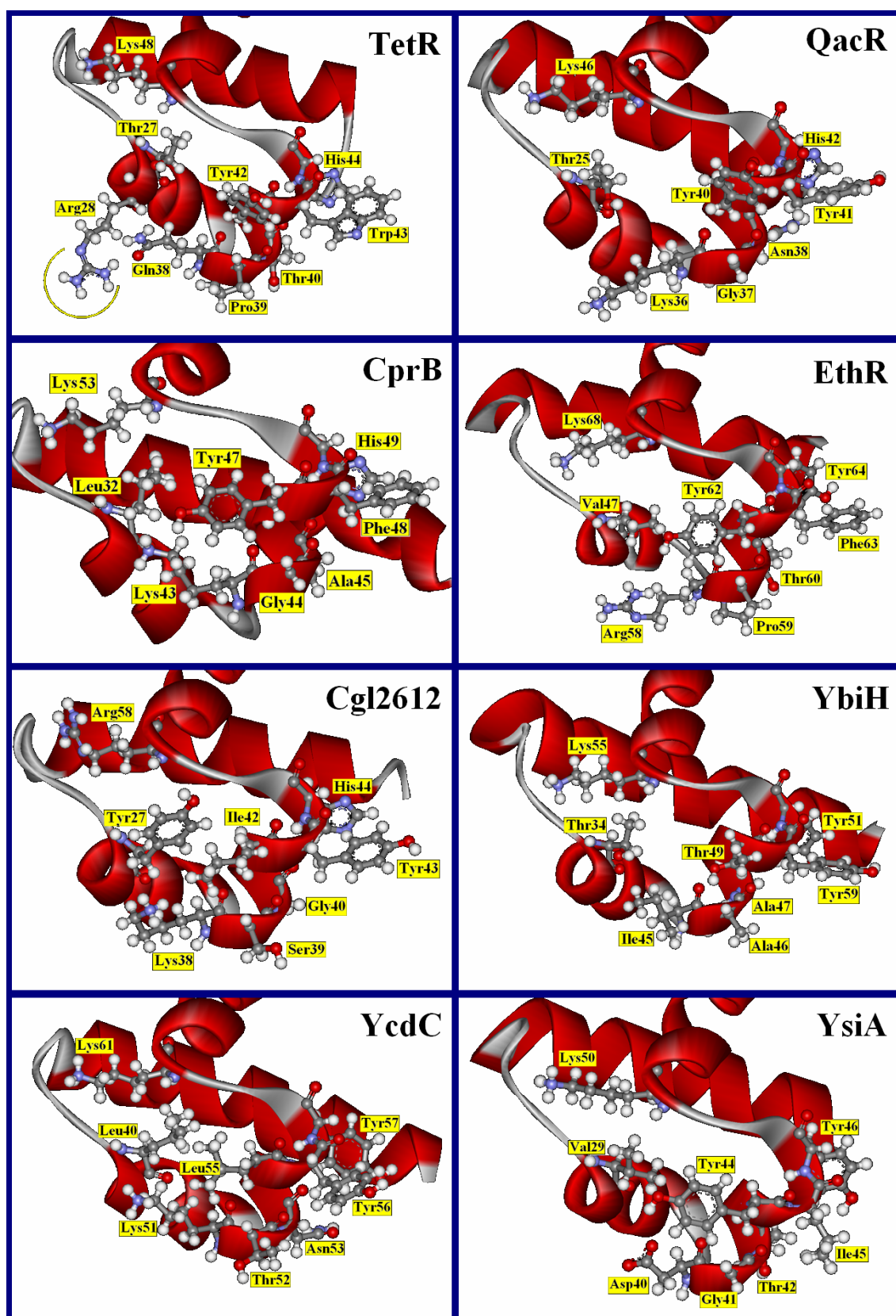


Figura 21. Dominios de unión a ADN con los residuos correspondientes a las posiciones de interacción con el ADN operador deducidas del análisis comparativo de los co-cristales de QacR y TetR con sus correspondientes operadores. (TetR: TQPAYWHK; QacR: TKGNYHYK; CprB: LKGAYFHK; EthR: VRPTYFYK; Cgl2612: YKSGIYHR; YbiH: TIAATYYK; YcdC: LKTNLYYK; YsiA: VDGTYIYK).

6.7. Algunos reguladores de la familia TetR forman parte de circuitos reguladores complejos

Los datos publicados indican que se conocen las funciones específicas de 85 miembros de la familia TetR (Tabla 29). Una información amplia acerca de cada una de estas proteínas está disponible en <http://www.bactregulators.org> (Martínez-Bueno *et al.*, 2004). Hemos agrupado las funciones reguladas por los miembros de la familia TetR en 10 conjuntos tipo (Tabla 29). La función más frecuente es la regulación de la expresión de bombas de extrusión implicadas en la resistencia frente a antibióticos y tolerancia a compuestos tóxicos.

Hemos observado que los miembros de la familia TetR a menudo autorregulan su expresión. Este control por retroalimentación asegura que la concentración del represor se encuentre dentro de unos límites óptimos (Beinlich *et al.*, 2001; Delany *et al.*, 2003; Marqués *et al.*, 1998; Sala *et al.*, 2003; Thieffry *et al.*, 1998). En este esquema de regulación sencillo, la síntesis del represor y de la(s) proteína(s) regulada(s) se desreprime en presencia de la molécula de inductor.

Sin embargo, las proteínas de la familia TetR también participan en otros tipos de redes de regulación que controlan procesos complejos tales como homeostasis metabólica (biosíntesis de amino ácidos, nucleótidos, protohemo y material de reserva); la síntesis de protectores osmóticos, *quorum sensing*, virulencia y procesos relacionados con la diferenciación (esporulación y biosíntesis de antibióticos en estreptomicetos) (Tabla 29 y www.bactregulators.org) (Martínez-Bueno *et al.*, 2004).

La Figura 22 muestra una serie de esquemas de regulación, desde los más simples a los más complejos, en los que interviene al menos un miembro de la familia TetR.

Pensando en la brevedad, a continuación sólo analizamos algunos conjuntos representativos de redes de regulación, incluidas proteínas implicadas en resistencia frente a drogas (AcrR de *E. coli* y MtrR de *Neisseria gonorrhoeae*), biosíntesis de un protector osmótico (BetI), una proteína clave implicada en la producción de antibióticos y diferenciación en *Streptomyces* (ArpA), una proteína implicada en la patogénesis de *Vibrio cholerae* (HapR) y algunas otras proteínas implicadas en *quorum sensing*.

6.7.1. El regulador AcrR es el regulador específico local de la bomba de extrusión AcrAB.

La resistencia múltiple a antibióticos en *E. coli* es un área reciente de atención, promoviendo la explicación de varios mecanismos implicados en este fenómeno. Uno de ellos es el transporte de diversos sustratos fuera de la célula mediante el transportador de eflujo AcrAB-TolC, responsable del fenotipo "Mar", multirresistencia a antibióticos (Nikaido, 1996). El conjunto de antibióticos frente al que AcrAB confiere resistencia incluye ampicilina, cloramfenicol, eritromicina, fluoroquinolonas, beta-lactámicos, novomicina, tetraciclina, tigeciclina y rifampicina (Hirata *et al.*, 2004; Klein and Plapp, 1992; Ma *et al.*, 1996; Nikaido, 1996; Nikaido, 1998; Oethinger *et al.*, 1998). AcrB es una proteína de membrana de gran tamaño (Ma *et al.*, 1995; Ma *et al.*, 1994; Yu *et al.*, 2003a; Yu *et al.*, 2003b) que se asocia a AcrB, una proteína de fusión de membrana (Okusu *et al.*, 1996) y a TolC, la proteína que forma el canal de extrusión de los sustratos al medio (Fralick, 1996; Koronakis *et al.*, 2000).

Los genes *acrA* y *acrB* forman un operón (Ma *et al.*, 1995) cuya transcripción está bajo el control de producto del gen *acrR*. El gen *acrR* tiene una orientación divergente respecto del operón *acrAB*. Esta observación es concordante con la función de AcrR como represor de la transcripción de *acrAB*. Los ensayos de retardo de movilidad en gel proporcionan evidencia directa de la unión de AcrR al promotor de *acrAB*, lo que se adoptó como una evidencia añadida a su papel como represor (Ma *et al.*, 1996). La secuenciación de ADN de ciertos aislados que sobreexpresan el gen *acrB* reveló que las cepas mutantes contenían inserciones que interrumpían el gen *acrR*, o mutaciones puntuales que hacían a AcrR un regulador no funcional, por ejemplo, la sustitución de un residuo de cisteína por arginina en la posición 45 de AcrR. Estas evidencias bioquímicas y genéticas apoyaban el papel regulador de AcrR.

MarA, SoxS y Rob son activadores de la transcripción relacionados, pertenecientes a la familia AraC/XylS (Amabile-Cuevas y Demple, 1991; George and Levy, 1983; Skarstad *et al.*, 1993) que activan la expresión de *acrAB* en respuesta a condiciones de estrés general (Ariza *et al.*, 1994; Ariza *et al.*, 1995; Barbosa y Levy, 2000; Benik *et al.*, 2000; Gambino *et al.*, 1993; Ma *et al.*, 1996; Ma *et al.*, 1993). Se ha descrito también que las condiciones de estrés general aumentaban la transcripción de *acrAB* incluso en un fondo genético deficiente en la síntesis de AcrR, y sorprendentemente, en estas condiciones se incrementaba también la expresión de *acrR*. Estos resultados sugirieron que AcrR actúa como un modulador secundario que afina el grado de transcripción de *acrAB* y previene la sobreexpresión, indeseada, de la bomba de extrusión.

6.7.2. El circuito Mtr de *Neisseria*.

La bomba de extrusión MtrCDE de *Neisseria gonorrhoeae* proporciona al gonococo resistencia frente a un conjunto de agentes antimicrobianos hidrofóbicos estructuralmente diversos y frente a péptidos antibióticos con estructura en lámina beta (proteína 1) o dos hélices (PC-8 y LC37), genéricamente conocidos como HAS (Guilfoile y Hutchinson, 1992a; Liu y Reynolds, 1999; Mahren y Braun, 2003; Schell, 1993).

Las mutaciones que inactivan *mtrR* o que conducen a que no se exprese, provocaron un incremento de la expresión del operón *mtrCDE* e incrementaron paralelamente la resistencia a HAS (Conlon *et al.*, 2002b; Guilfoile and Hutchinson, 1992a; Levy, 1988; Levy *et al.*, 1999; Orth *et al.*, 1998; Schell, 1993; Wissmann *et al.*, 1993). También se observó que cepas de *N. gonorrhoeae* hipersusceptibles a los HAS a menudo contenían mutaciones en los genes que *mtrCDE* que codifican la bomba de extrusión (Thieffry *et al.*, 1998).

El gen *mtrR* se transcribe divergentemente respecto al operón *mtrCDE* (Figura 22F). Los promotores de *mtrR* y *mtrC* se superponen en sus cajas -35 y los análisis de improntas *in vitro* mostraron que MtrR se unía a una región de 40 pares de bases dentro de la región -10 y -35 del promotor de *mtrR*, que contiene una secuencia palindrómica (Levy *et al.*, 1999). MtrR unido a su secuencia operadora impide su propia expresión y la del operón *mtrCDE* (Figura 22F).

La expresión de los genes *mtr* se incrementa mediante MtrA, un miembro de la familia AraC/XylS, aunque el mecanismo por el que esta proteína estimula la transcripción de *mtrCDE* es desconocido.

Por otro lado, Veal y Shafer (2003) han identificado recientemente un gen designado *mtrF* y situado en el extremo 3' de *mtrR*, que codifica una proteína de membrana citoplásmica de 56.1 kDa que presenta 12 dominios transmembrana. La expresión de *mtrF* se incrementaba en una cepa deficiente en MtrR, indicando que este gen, junto al operón *mtrCDE*, están sujetos a un control transcripcional dependiente de MtrR. Evidencias genéticas sugieren que MtrF es también importante en la expresión del fenotipo resistencia de alto nivel a detergentes del gonococo, y se ha propuesto que MtrF actuaría conjuntamente con la bomba MtrC-MtrD-MtrE para conferir al gonococo resistencia de alto nivel frente a ciertos HAS.

MtrR también controla al operón *farAB* que codifica una bomba de extrusión implicada en la resistencia frente a ácidos grasos de cadena corta (Figura 22F). La bomba FarAB usa a MtrE como componente de membrana externa (Kunin *et al.*, 1992).

6.7.3. BetI controla la ruta del osmorregulador colina-betaína en *Escherichia coli*.

En *Escherichia coli* la glicinbetaína actúa como osmoprotector en células sometidas a estrés hiperosmótico. Esta sustancia se acumula en grandes cantidades en el citoplasma, lo que permite a la célula ajustar su presión osmótica a la del medio evitando la deshidratación. La glicinbetaína es sólo uno de los varios compuestos usados como osmolitos por *E. coli*, pero con su acumulación, este microbio alcanza su grado más elevado de tolerancia osmótica (Knobloch *et al.*, 2001; Shaw *et al.*, 1998). Para acumular glicinbetaína *E. coli* necesita de un aporte externo de este compuesto o de sus precursores, colina o el aldehído de betaína.

La ruta de síntesis del osmorregulador glicinbetaína está codificada por los genes *bet*. El gen *betA* codifica una colina deshidrogenasa; *betB* codifica una deshidrogenasa del aldehído de la betaína; *betT* codifica un sistema transportador de colina; y *betI* codifica un represor de 21,8 kDa (Figura 22C). Los genes *bet* están ligados, aunque el gen *betT* se transcribe divergentemente respecto del operón *betIBA* (Kovacicikova y Skorupski, 2002; Showalter *et al.*, 1990). Análisis de extensión del cebador permitieron identificar dos promotores, parcialmente solapados, responsables de la expresión divergente del gen *betT* y del operón *betIBA*. Ambos puntos de inicio de la transcripción estaban separados por 61 pares de bases y se inducía su transcripción mediante estrés osmótico, pero para su máxima expresión requirió de la presencia de colina en el medio de cultivo (English *et al.*, 1994; Koronakis *et al.*, 2000; Krug *et al.*, 2005). Se purificó una proteína de fusión quimérica de BetI con la glutatión S-transferasa (BetI*) y los ensayos de movilidad en gel mostraron que BetI* formaba un complejo con un fragmento de ADN de 41 pares de bases que portaba la región intergénica del promotor *betI*. Los análisis de improntas *in vitro* revelaron la presencia de dos secuencias simétricas que probablemente constituían el operador de BetI.

Como la proteína ArcA reprime la expresión de los genes *bet* en *E. coli* bajo condiciones de anaerobiosis, los genes *bet* sólo se expresan en aerobiosis. Una mutación en *arcA* provoca la desrepresión total de los genes *bet*. Un patrón similar al descrito se había comprobado anteriormente para otros genes, por ejemplo *sodA*, el cual está regulado directamente por ArcA (Choi *et al.*, 2002).

Los resultados de distintos laboratorios sugieren que la regulación de los genes *bet* mediante colina, pero no la regulación osmótica, depende de BetI, un miembro de la familia TetR. Esto se dedujo del requerimiento de colina, además de estrés osmótico, para que se expresase *betT* en un mutante en que *betI* se disponía en *trans*. Además, este efecto causado por la colina no se observó en células carentes de *betI*. Estos hallazgos indicaban que *betI* codificaba un represor que reducía la expresión de *betT* (Retzlaff y Distler, 1995).

6.7.4. El regulador ArpA de *Streptomyces*.

Los *Streptomyces* son bacterias del suelo filamentosas, gram positivos, caracterizados por su capacidad de síntesis de una amplia variedad de metabolitos secundarios, incluidos antibióticos y sustancias biológicamente activas, y por su proceso de diferenciación morfológica que culmina con la formación de cadenas de esporas (Chan *et al.*, 2004). La síntesis de metabolitos secundarios a veces se denomina diferenciación "fisiológica" porque ocurre durante la idiofase después del periodo principal de crecimiento vegetativo rápido y metabolismo asimilativo (Harrison y Aggarwal, 1990; Heuer *et al.*, 2004; Okamoto *et al.*, 1995).

El regulador final del proceso que acabamos de mencionar es, en *Streptomyces griseus*, una proteína homodimérica denominada "proteína receptora del factor A" (ArpA) (Hillen *et al.*, 1984; Hinrichs *et al.*, 1994; Hirata *et al.*, 2004; Hitomi *et al.*, 2003), que regula el cambio de la diferenciación fisiológica y morfológica. En efecto, ArpA reprime la expresión de varios genes (Figura 22K) pero en presencia de 2-isocaproilil-3R-hidroximetil- γ -butirolactona tiene lugar la transcripción del clúster de genes de la biosíntesis de la estreptomicina y se produce la esporulación (Horinouchi, 2002; Horinouchi y Beppu, 1990; Horinouchi y Beppu, 1994; Horinouchi *et al.*, 2001). Una cepa mutante deficiente en ArpA o que producía una proteína ArpA incapaz de unirse a sus secuencias de ADN operadoras, era superproductora de estreptomicina y formaba un micelio aéreo y esporas antes que la cepa silvestre (Nixon *et al.*, 1986; Noguchi *et al.*, 2000a).

In vitro, ArpA se une a su secuencia diana de ADN que consiste en un palíndromo de 22 p.b. (5'-GG(T/G)CGGT(A/T)(T/C)G(T/G)-3'). La adición al complejo ADN-ArpA del efector γ -butirolactona causaba la separación de ArpA del ADN. La sustitución de la Val-41, del dominio hélice-giro-hélice, por Ala, eliminó la capacidad de unión a ADN de ArpA pero no su capacidad para unir al efector γ -butirolactona, apoyando la implicación del dominio hélice-giro-hélice en la unión a ADN. Por otro lado, la sustitución del Trp-119 por Ala originó un mutante incapaz de unir γ -butirolactona, produciéndose así un mutante insensible a la presencia de factor-A. Estos datos apoyan el que ArpA esté constituida por un dominio hélice-giro-hélice de unión a ADN en su extremo N-terminal y de un dominio de unión a efector en su extremo C-terminal.

En el agrupamiento de genes para la biosíntesis de estreptomicina en *S. griseus*, StrR es un regulador específico de la ruta que actúa como activador de la transcripción de los otros genes del agrupamiento (Pragman *et al.*, 2004;).

La expresión del gen *strR* está bajo el control de la proteína dependiente de factor-A (AdpA) que se une a la región situada en el extremo 5' del promotor de *strR* y activa su transcripción (Tropel y van der Meer, 2004; Ueda *et al.*, 1999).

La expresión del gen *adpA* está bajo el control de ArpA. La secuencia operadora de ArpA solapa con las regiones -35 y -10 del promotor *adpA*. ArpA permanece unido a la región promotora en ausencia de factor-A y la adición de dicho factor-A al complejo ADN-ArpA libera inmediatamente a ArpA del ADN. En concordancia con este hecho, el análisis de ARN reveló que *adpA* se transcribía sólo en presencia de factor-A y *strR* se transcribía sólo en presencia de *adpA*. Además, los mutantes *adpA* eran incapaces de producir estreptomicina y la superexpresión de *adpA* en la cepa silvestre de *S. griseus* causaba la producción de estreptomicina en una fase más temprana de su crecimiento y en cantidades mayores. Este conjunto de fenómenos aclaran como el factor-A dispara la biosíntesis de estreptomicina.

La inactivación del gen *adsA* cromosómico que codifica el factor sigma alternativo σ^{AdsA} resultó en la pérdida de la capacidad de formación de hifas aéreas pero no de la producción de estreptomicina, indicando que este factor sigma está implicado en el desarrollo morfológico (Tatusov *et al.*, 2001; Westrich *et al.*, 1999).

Se han descrito varias proteínas receptoras de autorreguladores tipo γ -butirolactona en otras especies de *Streptomyces*.

Por ejemplo, CprA y CprB, implicadas en el metabolismo secundario y formación del micelio aéreo en *S. coelicolor* A3(2) (Noguchi *et al.*, 2000b). El receptor de la butanilida llamada virginia, BarA está implicado en la biosíntesis de virginiamicina en *S. virginiae* (Nilsson y Kilstrup, 1998), y el receptor de IM-2, FarA, está implicado en la producción de pigmento azul en otra cepa de *Streptomyces* (Ueda *et al.*, 2005).

FarA en *Streptomyces lavendulae* es el sensor de la γ -butirolactona IM-2 y transduce esta señal, desreprimiendo la biosíntesis de antibióticos y un pigmento azul. Además, FarA también parece ser necesario para la biosíntesis de IM-2 (Figura 22J).

El papel de la proteína ScbR en el circuito de "quorum-sensing" de *S. coelicolor* es similar al de FarA en *S. lavendulae*. ScbR está también implicado en 3 funciones: regulando la síntesis de SCB1 (la γ -butirolactona que actúa como señal), recibiendo la señal, y transduciendo dicha señal, desreprimiendo de este modo la producción de los antibióticos actinorrodina y U-prodigiosina (Figura 22M).

6.7.5. HapR regula los genes de virulencia en *Vibrio cholerae*.

El desarrollo de cólera en humanos está relacionado directamente con la producción de dos factores de virulencia: la fimbria (TCP), que media la colonización, y la toxina colérica (CT), responsable de la diarrea grave que caracteriza a la enfermedad. La expresión coordinada de TCP y CT ocurre mediante una cascada de regulación compleja (Figura 22S) en la que los reguladores AphA y AphB activan sinérgicamente la transcripción de *tcpPH* (Klein *et al.*, 1991). HapR, proteína de familia TetR, regula negativamente la expresión de AphA e indirectamente disminuye la producción de los factores de virulencia TCP y CT.

HapR se regula mediante un sistema de señales de "quorum sensing" que se detectan y se transmiten mediante LuxO. El aparato de "quorum-sensing" de *Vibrio cholerae* es inusualmente complejo, y está compuesto por 3 sistemas paralelos (Krukonis *et al.*, 2000; Millar *et al.*, 2002).

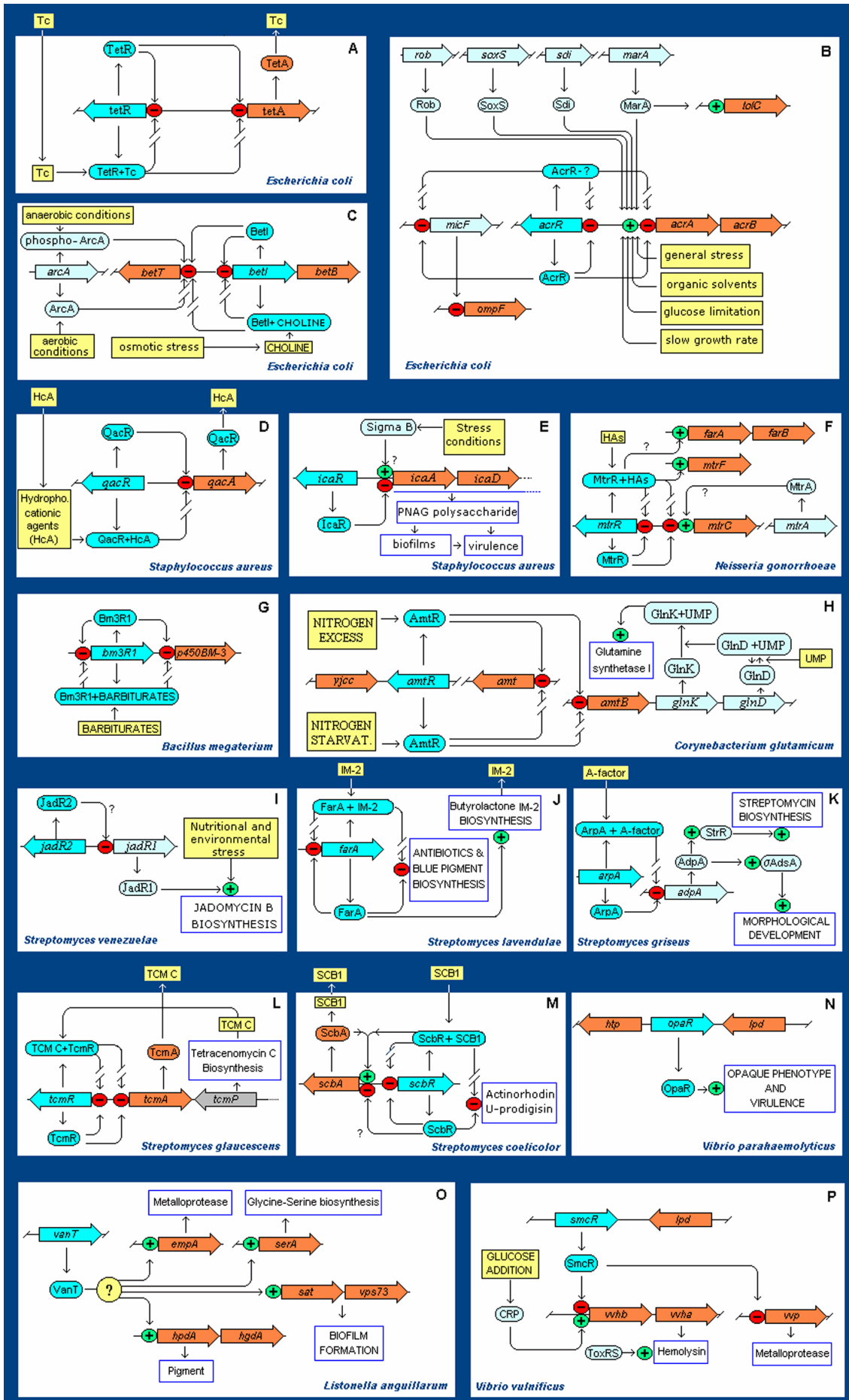
A diferencia de otras bacterias, en las que la densidad celular elevada dispara la expresión de los genes de virulencia, en *V. cholerae* la densidad celular baja es la condición que activa la producción de los factores de patogenicidad CT y TCP (Yamazaki *et al.*, 2000). A densidades celulares altas, HapR activa la expresión de una proteasa hemaglutinina (Hap) que promueve la liberación de *V. cholerae* del epitelio gastrointestinal (Bennik *et al.*, 2004) y muestra un efecto negativo sobre la formación de biofilms.

Teniendo en cuenta las funciones reguladoras de HapR y considerando que algunos biotipos patogénicos no expresan HapR y otros carecen del sitio de unión en *aphA*, parece que la expresión de HapR está relacionada con la disminución de la toxicidad y de la capacidad invasiva. Estas características ofrecen posibilidades fructíferas para el desarrollo de drogas capaces de modular la patogenicidad de *V. cholerae*.

6.7.6. Otros circuitos *quorum-sensing*.

Dentro del género *Vibrio* algunas proteínas de la familia TetR, por ejemplo LuxT, LuxR y LitR, están implicadas en circuitos de *quorum-sensing* complejos.

En *Vibrio harveyi* LuxT y LuxR participan en el circuito que regula la luminiscencia. Esta cepa detecta dos autoinductores: AI-1 y AI-2 (Chater y Bruton, 1985; Lau *et al.*, 1997). AI-1 es una acil-homoserin lactona, como en otras bacterias gram negativas. Sin embargo AI-2 es un nuevo tipo de autoinductor producido por muchas especies, que parece estar relacionado con la detección de la densidad celular interespecífica. Estos dos autoinductores usan el mismo sistema de fosfotransferencia para transducir la señal de regulación de la bioluminiscencia. La expresión de *luxR* está regulada por LitR (Figura 22R).



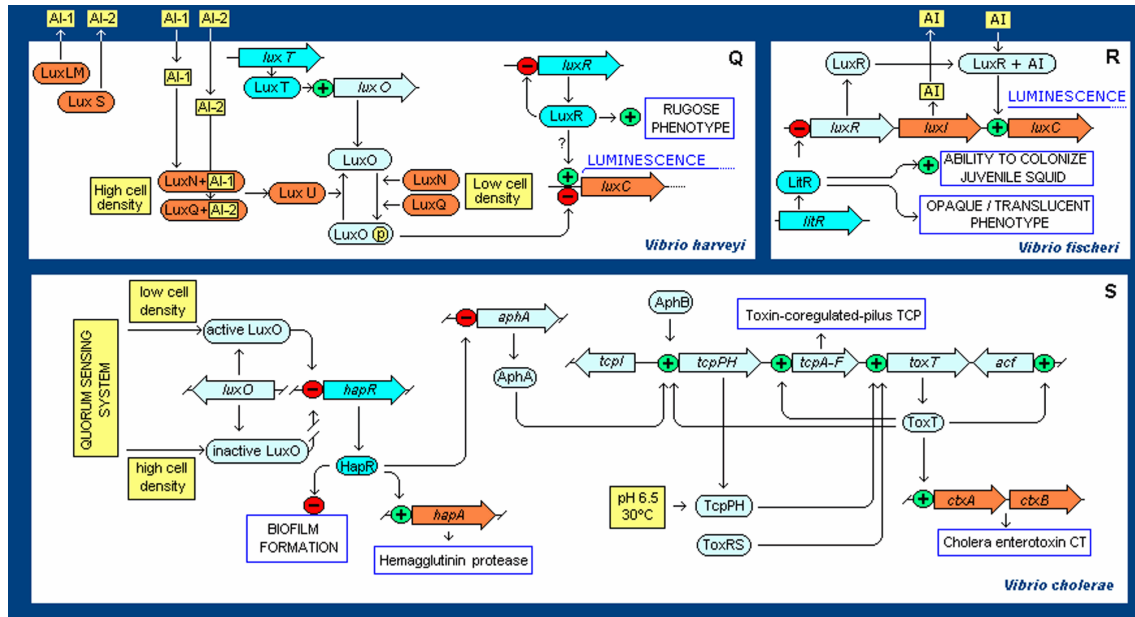


Figura 22. Redes de regulación que implican a miembros de la familia TetR. Aunque TetR y QacR no se pueden considerar integradas en una red, mostramos su tipo de control por encontrarse con frecuencia en otros miembros de la familia. Para las redes complejas se empleó el siguiente código de color: azul oscuro, miembro de la familia TetR; naranja, gen directamente regulado por un miembro de la familia TetR; azul claro, un regulador que controla la expresión de un miembro de la familia TetR, o que ayuda en la regulación del gen bajo el control del miembro de la familia TetR; cajas amarillas, señales y condiciones que influyen sobre el sistema; cajas abiertas, resultado final del sistema cuando el resultado es un fenotipo detectable. Referencias recomendadas para cada circuito: Panel A (Baumeister *et al.*, 1992a; 38, Orth *et al.*, 1998; Orth *et al.*, 2000; Galperin *et al.*, 2001; Wasylewski *et al.*, 1996; Waters *et al.*, 1983); Panel B (Aleksun and Levy, 1999; Amabile-Cuevas and Demple, 1991; Ariza *et al.*, 1994; Beinlich *et al.*, 2001; Bennik *et al.*, 2000; Gambino *et al.*, 1993; Kawamura-Sato *et al.*, 1999; Kobayashi *et al.*, 2001; Marger and Saier, 1993; McMurry *et al.*, 1998; Nishino and Yamaguchi, 2001; Yamamoto *et al.*, 2001); Panel C (Andresen *et al.*, 1998; Eshoo, 1988; Lamark *et al.*, 1995; Lamark *et al.*, 1996; Lamark *et al.*, 1992; Landfald and Strom, 1986; Ly *et al.*, 2004; Osteras *et al.*, 1998; Røkenes *et al.*, 1996); Panel D (Grkovic *et al.*, 1998; Grkovic *et al.*, 2001, A; Grkovic *et al.*, 2003; 249, Murray *et al.*, 2004; Rouch *et al.*, 1990; Schumacher *et al.*, 2001; Schumacher *et al.*, 2002b); Panel E (Allignet *et al.*, 2001; Allmeier *et al.*, 1992; Conlon *et al.*, 2002a; Conlon *et al.*, 2002b; Gotz, 2002; Jefferson *et al.*, 2003); Panel F (Maness and Sparling, 1973; Mavroidi *et al.*, 2001; Tropel and van der Meer, 2004; Rouquette *et al.*, 1999; Shafer *et al.*, 1995; Veal *et al.*, 1998); Panel G (English *et al.*, 1994; English *et al.*, 1996; English *et al.*, 1997; Guengerich, 1991; He *et al.*, 1995; Liang and Fulco, 1995; Liang *et al.*, 1995; Palmer *et al.*, 1998; Shaw *et al.*, 1997a; Urlinger *et al.*, 2000); Panel H (Burkovski, 2003; Jakoby *et al.*, 2000); Panel I (Doull *et al.*, 1994); Panel J (Chater and Bruton, 1985; Kitani *et al.*, 1999; Kitani *et al.*, 2001; Lee *et al.*, 2003; Waki *et al.*, 1997); Panel K (Hara and Beppu, 1982a; Hara and Beppu, 1982b; Horinouchi, 2002; Waki *et al.*, 1997); Panel L (Guilfoile and Hutchinson, 1992b); Panel M (Cho *et al.*, 2003; Orth *et al.*, 1998); Panel N (McCarter, 1998; McCarter, 1999; Schleif, 1987); Panel O (Milton *et al.*, 1997); Panel P (Cheng *et al.*, 1996; Choi *et al.*, 2003; McDougald *et al.*, 2000; McDougald *et al.*, 2001; Shao and Hor, 2001); Panel Q (Bassler *et al.*, 1994a; Bassler *et al.*, 1994b; Cao *et al.*, 1989; Fuqua *et al.*, 1996; Martin *et al.*, 1989; Miyamoto *et al.*, 1998); Panel R (Fidopiastis *et al.*, 2002; Fuqua *et al.*, 1996; Gray and Garey, 2001; Miyamoto *et al.*, 2000; Schauder and Bassler, 2001; Showalter *et al.*, 1990); Panel S (Fuqua *et al.*, 1996; Gray and Garey, 2001; Kim *et al.*, 1999; Krukoniš *et al.*, 2000; Lin *et al.*, 1993; Miller *et al.*, 2002).

6.8. Aplicaciones biotecnológicas y perspectivas futuras.

El sistema TetR es actualmente el más utilizado para la expresión condicional de genes en células eucariotas (Berens and Hillen, 2003). El sistema se basa en la alta afinidad (10^{-9} M) de TetR por su operador, *tetO*, la farmacocinética favorable de las tetraciclinas (difunden a través de las membranas biológicas) y su largo historial de uso clínico seguro. La clonación de los elementos *tetO* adyacentes a la caja TATA del gen diana (Gambino *et al.*, 1993) se utilizó con éxito para controlar los genes expresados por la ARN polimerasa I en *Leishmania donovani* y por ARN polimerasa III en levaduras y células vegetales (de Lorenzo and Schweizer, 2004; Takano *et al.*, 2000). Sin embargo, este sistema no es eficiente en células de mamífero. Por esta razón y basado en el conocimiento adquirido sobre cómo TetR se une a su secuencia operadora, se construyeron muchas versiones quiméricas en las que TetR se unió a dominios reguladores eucariotas, tales como el dominio de activación ácido (tTA) (Baron *et al.*, 1997; Garnier *et al.*, 2003; Tauch *et al.*, 2000;) o dominios de represión (tTS) (Belli *et al.*, 1998; Rodionov *et al.*, 2001). Basados en transreguladores híbridos, se obtuvieron ratones transgénicos capaces de producir la toxina diftérica o que expresaban reguladamente la toxina Shiga β para inducir apoptosis en células de mamífero.

También se ha usado el sistema *tet* en el estudio del cáncer y desórdenes neurológicos (Eshoo, 1988; Faust *et al.*, 2000; van Rooijen y de Vos, 1990). En el futuro, se esperan avances en la aproximación a procesos biológicos multifactoriales como el desarrollo y ciertas enfermedades, lo que podría ser relevante en el tratamiento de enfermedades complejas (Berens and Hillen, 2003).

Las bombas de eflujo de disolventes exhiben generalmente una amplia especificidad de sustrato pero algunas de ellas son altamente específicas y eliminan cierto número de productos químicos.

Una de tales bombas es TtgABC de *Pseudomonas putida* DOT-T1E, que extrae tolueno, *m*-xileno, propilbenceno además de estireno y otros hidrocarburos aromáticos (Terán *et al.*, 2003; Ramos *et al.*, 1998; Rojas *et al.*, 2003), además de varios antibióticos. Esta bomba de eflujo está bajo el control del represor TtgR miembro de la familia TetR (Duque *et al.*, 2001) y tiene aplicaciones en el desarrollo de bacterias resistentes a disolventes. Además es potencialmente apta para la biotransformación de compuestos insolubles en agua en compuestos con valor añadido. Uno de estos sistemas se ha explotado para producir catecoles a partir de xilenos/toluenos en sistema de fase-doble (Rojas *et al.*, 2004). Como estos sistemas de eflujo requieren altos niveles de energía, su expre-

sión en hospedadores heterólogos debe ser controlada estrechamente por su represor correspondiente, el cual debe de ser capaz de responder a la presencia del disolvente aromático. Estos tipos de bombas se presume que pueden ser útiles si se transfieren a una amplia variedad de bacterias que porten las maquinarias de biotransformación adecuadas.

Como se muestra en la Figura 22, los miembros de la familia TetR desempeñan papeles claves en la multirresistencia a drogas, procesos de virulencia y patogenicidad en algunas bacterias patógenas. El desarrollo de drogas que se unan irreversiblemente a los represores impidiendo su liberación de sus operadores correspondientes puede constituir una estrategia para luchar contra tales patógenos.

ANEXO II

Tabla 32: Residuos en las posiciones 22, 33, 34, 35,37, 38, 39 y 43 del alineamiento múltiple de las secuencias de la familia TetR (“Palabras de interacción con ADN”).

Nº	Palabra	AcNumber	ShortName	LongName	Organismo
1	LRGAYWHK	P34000	ACRR_ECOLI	Potential acrAB operon repressor	Escherichia coli O6 Shigella flexneri
		P31676	ENVR_ECOLI	Potential acrEF/envCD operon repressor	Escherichia coli O157:H7
		P39897	MTRR_NEIGO	Regulatory protein mtrR	Neisseria gonorrhoeae
		Q93PU7	Q93PU7	Putative regulatory protein TtgW	Pseudomonas putida
		Q9F939	Q9F939	Putative AcrR protein (Fragment)	Salmonella typhimurium
		Q59608	Q59608	MtrR protein	Neisseria gonorrhoeae
		Q93K42	Q93K42	AcrR protein	Klebsiella pneumoniae
		Q9AIU0	Q9AIU0	Regulatory protein TtgR	Pseudomonas putida
		Q9KJC4	Q9KJC4	ArpR	Pseudomonas putida
		Q9R9T9	Q9R9T9	Efflux pump regulator SrpR	Pseudomonas putida
		Q8VPB0	Q8VPB0	Regulatory protein AcrR	Proteus mirabilis
		Q8KLP4	Q8KLP4	Repressor	Stenotrophomonas maltophilia
		Q9AEG3	Q9AEG3	Putative AcrR protein	Enterobacter aerogenes
		Q9JY65	Q9JY65	Transcriptional regulator MtrR	Neisseria meningitidis serogroup B
		Q9JT49	Q9JT49	Transcriptional regulator	Neisseria meningitidis serogroup A
		Q9HY46	Q9HY46	Probable transcriptional regulator	Pseudomonas aeruginosa
		Q8ZLN6	Q8ZLN6	Transcriptional repressor for envCD (acrEF) (TetR/AcrR family)	Salmonella typhimurium
		Q8XEW7	Q8XEW7	AcrAB operon repressor (TetR/AcrR family) (Potential acrAB operon repressor)	Salmonella typhimurium
		Q8X5D3	Q8X5D3	AcrAB operon repressor	Salmonella typhi
		Q8UB16	Q8UB16	Transcriptional regulator, TetR family	Escherichia coli O157:H7 Agrobacterium tumefaciens str. C58
		Q92T04	Q92T04	Putative transcription regulator protein	Sinorhizobium meliloti
		Q8Y3G8	Q8Y3G8	Putative ACrAB operon repressor transcription regulator protein	Ralstonia solanacearum
		Q8ZC89	Q8ZC89	Putative tetR-family transcriptional regulatory protein (AcrAB operon repressor)	Yersinia pestis
		Q8FD26	Q8FD26	Potential acrEF/envCD operon repressor	Escherichia coli O6
		AAP15941	AAP15941	AcrAB operon repressor	Shigella flexneri 2a str. 2457T
		AAO69973	AAO69973	Potential acrAB operon repressor	Salmonella typhi
		16128448	acrR	acrAB operon repressor	Escherichia coli K12
		16131152	envR	putative transcriptional regulator	Escherichia coli K12
		15598770		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		15800193	acrR	acrAB operon repressor	Escherichia coli O157:H7 EDL933
		15803798	envR	putative transcriptional regulator	Escherichia coli O157:H7 EDL933
		15829771		acrAB operon repressor	Escherichia coli O157:H7
		15833390		putative transcriptional regulator	Escherichia coli O157:H7
26988121	ttgR	transcriptional regulator TtgR	Pseudomonas putida KT2440		
15963944		PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Sinorhizobium meliloti		
15891729		AGR_L_3219p	Agrobacterium tumefaciens str. C58 (Cereon)		
15677564		transcriptional regulator MtrR	Neisseria meningitidis MC58		
15794854	mtrR	transcriptional regulator	Neisseria meningitidis Z2491		
16123293	acrR	putative tetR-family transcriptional regulatory protein	Yersinia pestis CO92		
16763857	acrR	acrAB operon repressor (TetR/AcrR family)	Salmonella typhimurium LT2		
16766684	envR	transcriptional repressor for envCD (acrEF) (TetR/AcrR family)	Salmonella typhimurium LT2		
16759457	acrR	potential acrAB operon repressor	Salmonella enterica subsp.		

		17544731	acrR	PUTATIVE ACRAB OPERON REPRESSOR TRANSCRIPTION REGULATORY PROTEIN	enterica serovar Typhi Ralstonia solanacearum
		17936912	acrR	transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
		22124958	acrR	acrAB operon repressor	Yersinia pestis KIM
		24111847	acrR	acrAB operon repressor	Shigella flexneri 2a str. 301
		26246477	acrR	Potential acrAB operon repressor	Escherichia coli CFT073
		26249856	envR	Potential acrEF/envCD operon repressor	Escherichia coli CFT073
		28871439		transcriptional regulator, TetR family	Pseudomonas syringae pv. tomato str. DC3000
		29142771	acrR	potential acrAB operon repressor	Salmonella enterica subsp. enterica serovar Typhi Ty2
		30061964	acrR	acrAB operon repressor	Shigella flexneri 2a str. 2457T
2	LRGAHWHK	O68442	O68442	Regulatory protein	Agrobacterium tumefaciens
3	LRGAYWYK	Q8FC12	Q8FC12	Hypothetical protein	Escherichia coli O6
		26250370		Hypothetical protein	Escherichia coli CFT073
4	MPGAYWHK	Q8NNT1	Q8NNT1	Transcriptional regulator	Corynebacterium glutamicum
		Q8FNY4	Q8FNY4	Putative regulatory protein	Corynebacterium efficiens
		19553303		transcriptional regulator	Corynebacterium glutamicum ATCC 13032
		25028565		putative regulatory protein	Corynebacterium efficiens YS-314
5	FPGAYWHK	Q9K463	Q9K463	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
6	VKGSYWHR	Q8YQT9	Q8YQT9	Transcriptional regulator	Nostoc sp. PCC 7120
		Q8PI46	Q8PI46	Transcriptional regulator	Xanthomonas axonopodis pv. citri
		Q8P6U7	Q8P6U7	Transcriptional regulator	Xanthomonas campestris pv. campestris
		17231220		transcriptional regulator	Nostoc sp. PCC 7120
		21232299		transcriptional regulator	Xanthomonas campestris pv. campestris str. ATCC 33913
		21243779		transcriptional regulator	Xanthomonas axonopodis pv. citri str. 306
7	IKGAYLHK	O31500	YERO_BACSU	Hypothetical transcriptional regulator yerO	Bacillus subtilis
		16077738	yerO	similar to transcriptional regulator (TetR/AcrR family)	Bacillus subtilis subsp. subtilis str. 168
8	VKGAYVHK	P39601	YWCC_BACSU	Hypothetical transcriptional regulator ywcc	Bacillus subtilis
		16080873	ywcc	alternate gene name: ipa-33d	Bacillus subtilis subsp. subtilis str. 168
9	IKGAYFHK	Q9XCC7	Q9XCC7	Gamma-butyrolactone receptor protein TyIP	Streptomyces fradiae
		Q8VVP1	Q8VVP1	Autoregulator receptor protein	Streptomyces pristinaespiralis
		Q9ZN78	Q9ZN78	ArpA	Streptomyces griseus
		Q54189	Q54189	A-factor receptor	Streptomyces griseus
		Q9ZHP8	Q9ZHP8	Butyrolactone receptor	Streptomyces fradiae
		Q9RPK9	Q9RPK9	TarA	Streptomyces tendae
		O24741	O24741	FarA	Streptomyces sp.
		O86852	O86852	Gamma-BUTYROLACTONE binding protein	Streptomyces coelicolor
		BAC66444	BAC66444	Butyrolactone autoregulator receptor protein	Streptomyces clavuligerus
		BAC69981	BAC69981	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		BAC71417	BAC71417	Putative gamma-butyrolactone receptor protein	Streptomyces avermitilis
		29828812		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		29830248	avaR	putative gamma-butyrolactone receptor protein	Streptomyces avermitilis MA-4680
10	VKGAYFHK	Q60011	Q60011	Virginiae butanolide receptor	Streptomyces virginiae
		Q9JN89	Q9JN89	Hypothetical 24.1 kDa protein (Putative lactone-dependent transcriptional regulator (TetR-family),	Streptomyces coelicolor

		O67157	O67157	MmfR)	
		BAC76540	BAC76540	Transcriptional regulator (TETR/ACRR family)	Aquifex aeolicus
		15606341	acrR1	Putative TetR family receptor protein	Streptomyces rochei
		21234241	mmfR	transcriptional regulator (TetR/AcrR family)	Aquifex aeolicus VF5
11	LKGAYFHK	Q9LBV6 O66122 O66121 BAC76532	Q9LBV6 O66122 O66121 BAC76532	putative lactone-dependent transcriptional regulator (TetR-family), MmfR	Streptomyces coelicolor A3(2)
				BarZ	Streptomyces virginiae
				CPRL	Streptomyces coelicolor
				CprA (Transcriptional regulator)	Streptomyces coelicolor
				Putative TetR family receptor protein	Streptomyces rochei
12	IKGSYFHK	O85706	O85706	THLR (THLR, HTH transcriptional regulator TETR/ACRR family)	Clostridium acetobutylicum
		15004783	thlr	ThlR, HTH transcriptional regulator	Clostridium acetobutylicum
13	LRGAYFHK	Q8YF91	Q8YF91	TetR/AcrR family	Brucella melitensis
		Q8G2M8	Q8G2M8	Transcriptional regulator, TetR family	Brucella melitensis biovar Suis
		17987914		TRANSCRIPTIONAL REGULATOR, TETR FAMILY	Brucella melitensis 16M
		23501197		transcriptional regulator, TetR family	Brucella suis 1330
14	VSGAYFHK	Q93RS2	Q93RS2	Putative tetR-family regulatory protein	Streptomyces coelicolor
15	IQGGYFHK	AAM78022	AAM78022	Butyrolactone receptor	Streptomyces carzinostaticus subsp. neocarzinostaticus
16	MRGAYGHK	O87934 Q9RG61 CAD85025	O87934 Q9RG61 CAD85025	Regulator AmrR	Burkholderia pseudomallei
				Transcriptional regulator	Pseudomonas aeruginosa
				Bacterial regulatory proteins, TetR family	Nitrosomonas europaea
		15597216		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		30249103		Bacterial regulatory proteins, TetR family	Nitrosomonas europaea ATCC 19718
17	LRGAYGHK	Q8XRL5	Q8XRL5	Probable transcription regulator protein	Ralstonia solanacearum
		17549037	RSp0816	PROBABLE TRANSCRIPTION REGULATOR PROTEIN	Ralstonia solanacearum
18	LKGAYGHK	Q9JN93	Q9JN93	Hypothetical 21.9 kDa protein (Putative lactone-dependent transcriptional regulator (TetR-family), MmyR)	Streptomyces coelicolor
		BAC76314	BAC76314	Putative lactone-dependent transcriptional regulator MmyR	Streptomyces violaceoruber
		BAC69979	BAC69979	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29828810		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		21234245	mmyR	putative lactone-dependent transcriptional regulator (TetR-family), MmyR	Streptomyces coelicolor A3(2)
19	MKAAYYHK	Q9ZGB7	Q9ZGB7	LanK	Streptomyces cyanogenus
20	LKAAYYHK	Q9X8M4 BAC72413	Q9X8M4 BAC72413	Putative tetR-family regulator	Streptomyces coelicolor
				Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29831244		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
21	LKASYHK	O70020 Q9AIQ9 Q8GLC6 Q9RQQ0	O70020 Q9AIQ9 Q8GLC6 Q9RQQ0	IcaR	Staphylococcus epidermidis
				IcaR	Staphylococcus caprae
				Intercellular adhesion protein IcaR	Staphylococcus epidermidis
				ICAR (ICA operon transcriptional regulator ICAR)	Staphylococcus aureus subsp. aureus Mu50
					Staphylococcus aureus subsp. aureus N315
					Staphylococcus aureus subsp. aureus MW2
		15928251	icaR	ica operon transcriptional regulator	Staphylococcus aureus subsp. aureus N315
		15925655	icaR	IcaR	Staphylococcus aureus subsp. aureus Mu50
		21284314	icaR	ica operon transcriptional regulator	Staphylococcus aureus

22	LKPAYYHK	Q9RP98	Q9RP98	Regulatory protein		subsp. aureus MW2
		AAP25037	AAP25037	Transcriptional regulator, family	TetR	Streptomyces fradiae
		AAP08040	AAP08040	Transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		30261174		transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
23	MKPAYHHK	Q9RSH6	Q9RSH6	Transcriptional regulator, family	TETR	Bacillus cereus ATCC 14579
		15807142	DR2148	transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
24	LQGTYYHK	Q9L7W6	Q9L7W6	FenJ		Deinococcus radiodurans
		Q9KFW7	Q9KFW7	Hypothetical protein BH0353		Deinococcus radiodurans
		AAP25646	AAP25646	Transcriptional regulator, family	TetR	Bacillus subtilis
		15612916	BH0353	BH0353~unknown conserved protein		Bacillus halodurans
25	MKGTYYHK	Q59306	Q59306	transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		Q97N07	Q97N07	Transcriptional regulator family	TetR/AcrR	Bacillus anthracis str. Ames
		Q8RDG2	Q8RDG2	Transcriptional regulator		Clostridium pasteurianum
		15893330		Transcriptional regulator family	TetR/AcrR	Clostridium acetobutylicum
		20806602	AcrR	Transcriptional regulator		Thermoanaerobacter tengcongensis
26	MKGSYYHK	28209852	tetR	transcriptional regulator, tetR family		Thermoanaerobacter tengcongensis
		AAP24419	AAP24419	Transcriptional regulator, family	TetR	Clostridium tetani E88
		AAP07473	AAP07473	Transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		30260556		transcriptional regulator, family	TetR	Bacillus cereus ATCC 14579
27	MKSNYYHK	30018641		Transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		AAP24946	AAP24946	Transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		AAP10102	AAP10102	Transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		30261083		transcriptional regulator, family	TetR	Bacillus cereus ATCC 14579
28	VAGSYHHK	30021270		Transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		BAC73909	BAC73909	Putative TetR-family transcriptional regulator		Bacillus cereus ATCC 14579
		29832740		putative TetR-family transcriptional regulator		Bacillus anthracis str. Ames
29	IKGAYHHK	P28815	TETC_ECOLI	Transposon Tn10 tetC protein (ORFL)		Streptomyces avermitilis
		Q9S454	Q9S454	Tetracycline transcriptional regulator		Streptomyces avermitilis MA-4680
		Q93F25	Q93F25	Putative transcriptional regulator TetC		Escherichia coli
		Q9I454	Q9I454	Probable transcriptional regulator		Shigella flexneri
		Q8G4V3	Q8G4V3	Possible TetR-type transcriptional regulator		Escherichia coli
30	IKGSYHHK	15596487		probable transcriptional regulator		Salmonella typhi
		23465833		possible TetR-type transcriptional regulator		Salmonella typhimurium
		Q8VPQ1	Q8VPQ1	Putative transcriptional repressor		Shigella flexneri 2a
		Q9A3L1	Q9A3L1	Transcriptional regulator, family	TetR	Pseudomonas aeruginosa
31	VKGSYHHK	16127422		transcriptional regulator, family	TetR	Bifidobacterium longum
				transcriptional regulator, family	TetR	PAO1
32	MRGAYHHK	Q9I1K3	Q9I1K3	Probable transcriptional regulator		Bifidobacterium longum NCC2705
		Q985Z0	Q985Z0	Transcriptional regulator		Micrococcus sp. 28
		Q92WL6	Q92WL6	Putative transcriptional regulator		Caulobacter vibrioides

		Q8U7Y2	Q8U7Y2	protein Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
		15597466		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		13476207		transcriptional regulator	Mesorhizobium loti
		26990665		transcriptional regulator, TetR family	Pseudomonas putida KT2440
		15890665		AGR_L_1093p	Agrobacterium tumefaciens str. C58 (Cereon)
		16264071		putative transcriptional regulator protein	Sinorhizobium meliloti
		17938006		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
33	MKGAYHHK	Q8DSI0	Q8DSI0	Putative transcriptional regulator	Streptococcus mutans
		24380169		putative transcriptional regulator	Streptococcus mutans UA159
		27377507		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
34	TRGAYHHK	Q93QZ7	Q93QZ7	HydR	Clostridium difficile
		Q98IG0	Q98IG0	Probable transcriptional regulator	Mesorhizobium loti
		Q92T19	Q92T19	Putative transcription regulator protein	Sinorhizobium meliloti
		13472203		probable transcriptional regulator	Mesorhizobium loti
		15963922		PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	Sinorhizobium meliloti
35	VRGAFHHK	O07388	O07388	Transcriptional regulator (Fragment)	Mycobacterium avium
		O53165	O53165	Transcriptional regulator (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		Q8NQ97	Q8NQ97	Transcriptional regulator	Corynebacterium glutamicum
		Q8FTA6	Q8FTA6	Conserved hypothetical protein	Corynebacterium efficiens
		15608612		hypothetical protein Rv1474c	Mycobacterium tuberculosis H37Rv
		15840935		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
		19552754		transcriptional regulator	Corynebacterium glutamicum ATCC 13032
		25028219		conserved hypothetical protein	Corynebacterium efficiens YS-314
36	MKAGYHHK	Q93A32	Q93A32	Putative transcription factor	Myxococcus xanthus
37	MKASYHHK	Q9I497	Q9I497	Probable transcriptional regulator	Pseudomonas aeruginosa
		Q9HM72	Q9HM72	Vng2675c	Halobacterium sp. NRC-1
		15596438		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		15791395	VNG2675C	Vng2675c	Halobacterium sp. NRC-1
38	MKSSYHHK	Q9XA31	Q9XA31	Putative transcriptional regulator	Streptomyces coelicolor
		BAC72072	BAC72072	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29830903		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
39	MKSAYHHK	Q8FRU3	Q8FRU3	Putative transcription regulator	Corynebacterium efficiens
		25027221		putative transcription regulator	Corynebacterium efficiens YS-314
40	LKAAYHHK	O08306	O08306	Hypothetical 22.7 kDa protein	Pimelobacter simplex
41	MKSLYHHK	27378554		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
42	FRGNYHHK	Q92W40	Q92W40	Putative transcriptional regulator protein	Sinorhizobium meliloti
		16264253		putative transcriptional regulator protein	Sinorhizobium meliloti
43	VSGSFHHK	Q9RFR4	Q9RFR4	Putative transcriptional regulator Pfl2	Pseudomonas fluorescens
		Q9HZW2	Q9HZW2	Probable transcriptional regulator	Pseudomonas aeruginosa
		15598081		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
44	VSGSYHHK	P96839	P96839	Hypothetical 22.9 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15610693		hypothetical protein Rv3557c	Mycobacterium tuberculosis H37Rv
		15843169		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
45	IIGSYHHK	Q9I2Q9	Q9I2Q9	Probable transcriptional regulator	Pseudomonas aeruginosa
		15597033		probable transcriptional regulator	Pseudomonas aeruginosa

46	MIGSYHHK	Q9CCD1 15827452	Q9CCD1	Hypothetical protein ML0949 hypothetical protein	PAO1 Mycobacterium leprae
47	MAGSYHHK	26990461		transcriptional regulator, TetR family	Mycobacterium leprae Pseudomonas putida KT2440
48	PQGSYHHK	Q9K3M6 BAC69149 29827980	Q9K3M6 BAC69149	Putative tetR-family regulatory protein Putative TetR-family transcriptional regulator	Streptomyces coelicolor Streptomyces avermitilis
49	LNGAYPHK	Q98J38 13471975	Q98J38	Probable transcriptional regulator	Streptomyces avermitilis MA-4680
50	LNGAYTHK	27382016		probable transcriptional regulator blr6905	Mesorhizobium loti Mesorhizobium loti Bradyrhizobium japonicum USDA 110
51	LNSAYTHK	Q8RER3 19704369	Q8RER3	Transcriptional regulator, TetR family Transcriptional regulator, TetR family	Fusobacterium nucleatum subsp. nucleatum Fusobacterium nucleatum subsp. nucleatum ATCC 25586
52	VKGTYYNK	Q8XPD4 Q9S2V6 BAC73815 29832646 18309011	Q8XPD4 Q9S2V6 BAC73815	Probable transcription regulator Putative transcriptional regulator Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator probable transcription regulator	Clostridium perfringens Streptomyces coelicolor Streptomyces avermitilis Streptomyces avermitilis MA-4680 Clostridium perfringens str. 13
53	MVSSYYHK	Q9ZF45 Q9CFG3 15673496	Q9ZF45 Q9CFG3	Hypothetical 14.0 kDa protein (Fragment) Transcriptional regulator transcription regulator	Lactococcus lactis Lactococcus lactis subsp. lactis Lactococcus lactis subsp. lactis
54	MVGAYYHK	Q8F8X1 24213128	Q8F8X1	Transcriptional regulator, tetR family Transcriptional regulator, tetR family	Leptospira interrogans Leptospira interrogans serovar lai str. 56601
55	VVPAYYHK	Q9X7X9	Q9X7X9	Putative regulatory protein	Streptomyces coelicolor
56	TKGNYHHK	P23217 Q59802 Q8KVS4 14141854	QACR_STAAM	Transcriptional regulator qacR Transcriptional regulator QACR (Repressor QACR) Transcription regulator QacR (Fragment) transcriptional regulator	Staphylococcus aureus subsp. aureus Mu50 Staphylococcus haemolyticus Staphylococcus aureus Staphylococcus epidermidis Staphylococcus epidermidis
57	TEGNYHHK	Q8UF32 15888886 17935466	Q8UF32	Transcriptional regulator, TetR family AGR_C_2887p transcriptional regulator, TetR family	Staphylococcus aureus subsp. aureus Mu50 Agrobacterium tumefaciens str. C58 Agrobacterium tumefaciens str. C58 (Cereon) Agrobacterium tumefaciens str. C58 (U. Washington)
58	TPGNYHHK	Q9F9H1 Q51516 Q9AGB3 Q8VV55 Q9F9Z7 Q9Z3X8 Q9Z436 Q9KV25 Q9K017 Q9JV29 Q9I6C7 Q9HUC4 Q8XXM7	Q9F9H1 Q51516 Q9AGB3 Q8VV55 Q9F9Z7 Q9Z3X8 Q9Z436 Q9KV25 Q9K017 Q9JV29	PhaD PhaC1, phaC2 and phaD genes PhaD (Fragment) PhaD protein PhaD (Fragment) PhaD Hypothetical 23.1 kDa protein Transcriptional regulator, TetR family Transcriptional regulator, TetR family Putative tetR-family transcriptional regulator Probable transcriptional regulator Probable transcriptional regulator Putative transcription regulator protein	Alcanivorax borkumensis Pseudomonas aeruginosa Pseudomonas nitroreducens Pseudomonas chlororaphis Pseudomonas resinovorans Pseudomonas sp. 61-3 Pseudomonas putida Vibrio cholerae Neisseria meningitidis serogroup B Neisseria meningitidis serogroup A Pseudomonas aeruginosa Pseudomonas aeruginosa Ralstonia solanacearum

		Q8E819	Q8E819	Transcriptional regulator, TetR family	Shewanella oneidensis
		Q8DCW8 15640360 15595564	Q8DCW8	Transcriptional regulator transcriptional regulator, TetR family probable transcriptional regulator	Vibrio vulnificus Vibrio cholerae Pseudomonas aeruginosa PAO1
		15600252		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		26991683		transcriptional regulator, TetR family	Pseudomonas putida KT2440
		26991795		transcriptional regulator, TetR family	Pseudomonas putida KT2440
		15676708		transcriptional regulator, TetR family	Neisseria meningitidis MC58
		15793976	NMA1020	putative tetR-family transcriptional regulator	Neisseria meningitidis Z2491
		17546805	RSc2086	PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum
		24376147		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
		27364673 28867649		Transcriptional regulator transcriptional regulator, TetR family	Vibrio vulnificus CMCP6 Pseudomonas syringae pv. tomato str. DC3000
		28872258	phaD	transcriptional regulator PhaD	Pseudomonas syringae pv. tomato str. DC3000
		28899610		transcriptional regulator, TetR family	Vibrio parahaemolyticus RIMD 2210633
59	TRGLYYHK	Q8KLN7	Q8KLN7	Hypothetical protein	Streptococcus salivarius
60	TPAAYIHK	Q9S253 BAC74313	Q9S253 BAC74313	Putative transcriptional regulator Putative TetR-family transcriptional regulator	Streptomyces coelicolor Streptomyces avermitilis
		29833144		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		21220194	SCO1699	putative transcriptional regulator	Streptomyces coelicolor A3(2)
61	TQASYHHK	Q9S3L4	Q9S3L4	AmtR protein (Repressor of the high-affinity (methyl) ammonium uptake system)	Corynebacterium glutamicum
		Q8FR27 19552090	Q8FR27	Transcription regulator AmtR transcriptional regulator	Corynebacterium efficiens Corynebacterium glutamicum ATCC 13032
		25027495	amtR	transcription regulator AmtR	Corynebacterium efficiens YS-314
62	TQASYHYK	BAC74412	BAC74412	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29833243		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
63	LTGAQWHK	Q9KWD3	Q9KWD3	Riorf51 protein	Agrobacterium rhizogenes
64	LVGATFHK	AAM78020	AAM78020	Transcriptional regulator	Streptomyces carzinostaticus subsp. neocarzinostaticus
65	LNGAHFHK	O24739 Q93S03 BAC71415	O24739 Q93S03 BAC71415	BarB Putative regulatory protein Putative TetR-family transcriptional regulator	Streptomyces virginiae Streptomyces coelicolor Streptomyces avermitilis
		BAC76537	BAC76537	Putative TetR family receptor protein	Streptomyces rochei
		29830246		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
66	LPGAHFHK	Q56153 BAC71414	Q56153 BAC71414	JadR2 Putative TetR-family transcriptional regulator	Streptomyces venezuelae Streptomyces avermitilis
		29830245		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
67	LSGAHFQK	Q9XCC5	Q9XCC5	Hypothetical transcriptional regulator TyIQ	Streptomyces fradiae
68	MPGTYHHK	Q59431	UIDR_ECOLI	Uid operon repressor (Gus operon repressor)	Escherichia coli O6
		AAP17110	AAP17110	Repressor for uid operon	Escherichia coli O157:H7 Shigella flexneri Shigella flexneri 2a str. 2457T
		16129576 15802032	uidR uidR	repressor for uid operon repressor for uid operon	Escherichia coli K12 Escherichia coli O157:H7 EDL933

		15831580		repressor for uid operon	Escherichia coli O157:H7
		24113006	uidR	repressor for uid operon	Shigella flexneri 2a str. 301
		26247865	uidR	Uid operon repressor	Escherichia coli CFT073
		30063132	uidR	repressor for uid operon	Shigella flexneri 2a str. 2457T
69	MPGTYRHR	Q8PBA8	Q8PBA8	Transcriptional regulator	Xanthomonas campestris pv. campestris
		21230672		transcriptional regulator	Xanthomonas campestris pv. campestris str. ATCC 33913
70	MPGTYRHK	28377727		transcription regulator (putative)	Lactobacillus plantarum WCFS1
71	MMGTYRHK	Q99R60	Q99R60	Hypothetical protein SAV2578 (MW2498 protein)	Staphylococcus aureus subsp. aureus Mu50
		15928157		hypothetical protein, similar to transcription regulator acrR	Staphylococcus aureus subsp. aureus N315
		15925568		hypothetical protein	Staphylococcus aureus subsp. aureus MW2
		21284227	MW2498	ORFID:MW2498-hypothetical protein, similar to transcription regulator acrR	Staphylococcus aureus subsp. aureus N315
72	MVGTYRHK	AAO65340	AAO65340	Putative TetR/AcrR-family transcriptional regulator	Staphylococcus aureus subsp. aureus Mu50
73	IAGTYRHK	Q9ZBM6	Q9ZBM6	Putative transcriptional regulator	Staphylococcus aureus subsp. aureus N315
		Q9CCW8	Q9CCW8	Putative TetR/AcrR-family transcriptional regulator	Staphylococcus aureus subsp. aureus MW2
		15827081		putative TetR/AcrR-family transcriptional regulator	Staphylococcus aureus subsp. aureus N315
74	VAGTYRHK	O53612	O53612	Putative transcriptional regulator (Transcriptional regulator, TetR family)	Staphylococcus aureus subsp. aureus Mu50
		15607209		hypothetical protein Rv0067c	Staphylococcus aureus subsp. aureus MW2
		15839446		transcriptional regulator, TetR family	Streptomyces murayamaensis
75	LIGTYRHR	Q98JI8	Q98JI8	Probable transcription regulator	Mycobacterium leprae
		Q8UJI1	Q8UJI1	Transcriptional regulator, TetR family	Mycobacterium leprae
		Q8XQU7	Q8XQU7	Putative transcription regulator protein	Mycobacterium leprae
		Q9RL56	Q9RL56	Putative tetR family transcriptional regulator	Mycobacterium tuberculosis
		13471824		probable transcription regulator	Mycobacterium tuberculosis (Transcriptional regulator, TetR family)
		16119724		AGR_pAT_729p	Mycobacterium tuberculosis H37Rv
		17549342	RSp1121	PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Mycobacterium tuberculosis CDC1551
		17939078		transcriptional regulator, TetR family	Mycobacterium tuberculosis H37Rv
		27377327		transcriptional regulatory protein	Mycobacterium tuberculosis CDC1551
76	LVGTYRHR	Q9X9W8	Q9X9W8	Putative transcriptional regulator	Mycobacterium tuberculosis H37Rv
		21220391	SCO1904	putative transcriptional regulator	Mycobacterium tuberculosis CDC1551
77	FVGTYRHR	BAC68218	BAC68218	Putative TetR family transcriptional regulator	Mycobacterium tuberculosis CDC1551
		29827049		putative TetR family transcriptional regulator	Mycobacterium tuberculosis CDC1551
78	LSGTYRHR	BAC74063	BAC74063	Putative TetR-family transcriptional regulator	Mycobacterium tuberculosis CDC1551
		29832894		putative TetR-family transcriptional regulator	Mycobacterium tuberculosis CDC1551
79	LQGVYRHR	Q9RL15	Q9RL15	Putative tetR-family transcriptional regulator	Mycobacterium tuberculosis CDC1551
80	LQGTYRHR	28871207		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
81	VTATYRHR	Q9ADK8	Q9ADK8	Putative tetR-family transcriptional regulator	Mycobacterium tuberculosis CDC1551
82	MTATYRHR	BAC68001	BAC68001	Putative TetR-family transcriptional regulator	Mycobacterium tuberculosis CDC1551
		29826832		putative TetR-family transcriptional regulator	Mycobacterium tuberculosis CDC1551
83	MTATYRHK	28869835		transcriptional regulator, TetR	Mycobacterium tuberculosis CDC1551

84	ITGLYRHK	Q9KE70 15613550	Q9KE70 BH0987	family Transcriptional regulator (TetR/AcrR family) transcriptional regulator (TetR/AcrR family)	tomato str. DC3000 Bacillus halodurans Bacillus halodurans
85	VTGQYRHK	Q8PQQ6 Q8PDT8 21229725 21241040	Q8PQQ6 Q8PDT8	Transcriptional regulator family Transcriptional regulator family transcriptional regulator acrR family	Xanthomonas axonopodis pv. citri Xanthomonas campestris pv. campestris Xanthomonas campestris pv. campestris str. ATCC 33913 Xanthomonas axonopodis pv. citri str. 306
86	MTGSYRHK	Q9HX43 15599168	Q9HX43	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
87	IHGLFRHK	Q50720 15610541 15843000	YY05_MYCTU	Hypothetical transcriptional regula- tor Rv3405c hypothetical protein Rv3405c transcriptional regulator, TetR family	Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv Mycobacterium tuberculosis CDC1551
88	LHGTYRHK	Q8DTD2 24379822	Q8DTD2	Putative transcriptional regulator putative transcriptional regulator	Streptococcus mutans Streptococcus mutans UA159
89	VHGSYRHK	BAC73350 29832181 27380800	BAC73350	Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator transcriptional regulatory protein	Streptomyces avermitilis Streptomyces avermitilis MA-4680 Bradyrhizobium japonicum USDA 110
90	VHGAYAHR	Q9ALU8	Q9ALU8	Putative regulator	Enterobacter cloacae
91	VHGAYAHK	Q98JG6 13471847	Q98JG6	Probable transcriptional regulator probable transcriptional regulator	Mesorhizobium loti Mesorhizobium loti
92	IHGGYNHK	Q9F419	Q9F419	Putative transcription regulator	Mycobacterium avium subsp. paratuberculosis
93	VHGGYGHK	Q8UJN3 16119671 17939026	Q8UJN3	Transcriptional regulator, TetR family AGR_pAT_650p transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 Agrobacterium tumefaciens str. C58 (Cereon) Agrobacterium tumefaciens str. C58 (U. Washington)
94	VHGGYKHK	Q8P940 21231474 28870112	Q8P940	Transcriptional regulator transcriptional regulator	Xanthomonas campestris pv. campestris Xanthomonas campestris pv. campestris str. ATCC 33913 Pseudomonas syringae pv. tomato str. DC3000
95	LHGGYAHK	Q936Z4 Q9HUX9 15600024 28868964	Q936Z4 Q9HUX9	Putative transcriptional regulator Probable transcriptional regulator probable transcriptional regulator transcriptional regulator, TetR family	Pseudomonas sp. ADP Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1 Pseudomonas syringae pv. tomato str. DC3000
96	VHGGYAHR	27379063		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
97	VVGGYAHK	Q9HZ91 15598329 26988704	Q9HZ91	Probable transcriptional regulator probable transcriptional regulator transcriptional regulator, TetR family	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1 Pseudomonas putida KT2440
98	VKGSYNHK	Q9RAJ1	Q9RAJ1	Inactive regulatory protein	Mycobacterium sp. GP1
99	IKGTYNHK	Q8EMS1 23100225	Q8EMS1	Transcriptional regulator (TetR/AcrR family) transcriptional regulator	Oceanobacillus iheyensis Oceanobacillus iheyensis HTE831
100	MKGTYNHK	Q8DA45 27365684 28898059	Q8DA45	Transcriptional regulator Transcriptional regulator putative transcriptional regulator	Vibrio vulnificus Vibrio vulnificus CMCP6 Vibrio parahaemolyticus RIMD 2210633
101	IKGGYNHK	Q8YNJ8 17232059	Q8YNJ8	Transcriptional regulator transcriptional regulator	Nostoc sp. PCC 7120 Nostoc sp. PCC 7120
102	VQGTYNHR	27378292		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
103	VKGTYKHK	Q9HZP1	Q9HZP1	Probable transcriptional regulator	Pseudomonas aeruginosa

		15598153		probable transcriptional regulator	<i>Pseudomonas aeruginosa</i>
		26988694		transcriptional regulator, TetR family	<i>Pseudomonas putida</i>
		28870976		transcriptional regulator, TetR family	<i>Pseudomonas syringae</i> pv. tomato str. DC3000
104	IRGTYKHK	Q929J6 Q8Y593 16804215	Q929J6 Q8Y593	Hypothetical protein lin2279 Hypothetical protein lmo2176 similar to transcriptional regulator (tetR family)	<i>Listeria innocua</i> <i>Listeria monocytogenes</i>
		16801343		similar to transcriptional regulator (tetR family)	<i>Listeria innocua</i>
105	IIGSYNHK	Q97FX1	Q97FX1	Transcriptional regulator (TetR/AcrR family)	<i>Clostridium acetobutylicum</i>
		15895863		Transcriptional regulator (TetR/AcrR family)	<i>Clostridium acetobutylicum</i>
106	ILGSYNHK	BAC68295	BAC68295	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29827126		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
107	IFGSYNHK	BAC68276	BAC68276	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29827107		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
108	VGGAYNHK	26989599		transcriptional regulator, TetR family	<i>Pseudomonas putida</i> KT2440
109	IRGASHHK	O33453	O33453	CYMR	<i>Pseudomonas putida</i>
110	VRGAQHHR	Q9K3V9	Q9K3V9	Putative tetR-family transcriptional regulator	<i>Streptomyces coelicolor</i>
		BAC71574	BAC71574	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29830405		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
111	TRGAMHHK	Q9XDW2	Q9XDW2	PsbI	<i>Rhodospseudomonas palustris</i>
112	TRGAYHQK	P72624 O53119	P72624 O53119	Hypothetical protein sll1392 Hypothetical 26.4 kDa protein (Putative TETR-family transcriptional regulator)	<i>Synechocystis</i> sp. PCC 6803 <i>Mycobacterium leprae</i>
		O53623	O53623	Transcriptional regulator (Transcriptional regulator, TetR family)	<i>Mycobacterium tuberculosis</i>
		16329218 15607220		hypothetical protein hypothetical protein Rv0078	<i>Synechocystis</i> sp. PCC 6803 <i>Mycobacterium tuberculosis</i> H37Rv
		15828443		putative TetR-family transcriptional regulator	<i>Mycobacterium leprae</i>
		15839457		transcriptional regulator, TetR family	<i>Mycobacterium tuberculosis</i> CDC1551
113	VPGSYAHK	Q93TU7	Q93TU7	Transcriptional repressor	<i>Rhodococcus</i> sp. NCIMB 9784
114	MPASYAHK	BAC68485	BAC68485	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29827316		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
115	MGGSYAHK	Q9RRV9	Q9RRV9	Transcriptional regulator, TETR family	<i>Deinococcus radiodurans</i>
		15807366	DR2376	transcriptional regulator, TetR family	<i>Deinococcus radiodurans</i>
116	MHASYAHR	Q53901	Q53901	ActII protein (Putative transcriptional regulatory protein)	<i>Streptomyces coelicolor</i>
117	MVASYSHQ	Q8F8U7	Q8F8U7	Transcriptional regulator, tetR family	<i>Leptospira interrogans</i>
		24213155		Transcriptional regulator, tetR family	<i>Leptospira interrogans</i> serovar lai str. 56601
118	MAASYDHK	Q8F7W7	Q8F7W7	Transcriptional regulator, tetR family	<i>Leptospira interrogans</i>
		24213525		Transcriptional regulator, tetR family	<i>Leptospira interrogans</i> serovar lai str. 56601
119	MAGTYRYK	P43506 Q8E9F7	BM3R_BACME Q8E9F7	Transcriptional repressor Bm3R1 Transcriptional regulator, TetR family	<i>Bacillus megaterium</i> <i>Shewanella oneidensis</i>
		Q8DDC4	Q8DDC4	Transcriptional regulator, TetR family	<i>Vibrio vulnificus</i>
		24375808		transcriptional regulator, TetR family	<i>Shewanella oneidensis</i> MR-1

		27364504		Transcriptional regulator	TetR	Vibrio vulnificus CMCP6
		28896814		Transcriptional regulator, family	TetR	Vibrio parahaemolyticus RIMD 2210633
120	MIGTYRYK	Q9RA03	Q9RA03	Hypothetical repressor protein	KstR	Rhodococcus erythropolis
121	MVGTYRYK	P96856	P96856	Hypothetical 21.9 kDa protein (Transcriptional regulator, family)	TetR	Mycobacterium tuberculosis
		15610710		hypothetical protein Rv3574		Mycobacterium tuberculosis H37Rv
		15843187		transcriptional regulator, family	TetR	Mycobacterium tuberculosis CDC1551
122	MLGTYRYK	Q9KXK1	Q9KXK1	Putative tetR-family regulatory protein		Streptomyces coelicolor
		BAC73566	BAC73566	Putative TetR-family transcriptional regulator		Streptomyces avermitilis
		29832397		putative TetR-family transcriptional regulator		Streptomyces avermitilis MA-4680
123	IAGTYRYK	AAP26772	AAP26772	Transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		AAP09884	AAP09884	Transcriptional repressor Bm3R1		Bacillus cereus ATCC 14579
		30262909		transcriptional regulator, family	TetR	Bacillus anthracis str. Ames
		27382250		transcriptional regulatory protein		Bradyrhizobium japonicum USDA 110
124	MTGTYRYK	30021052		Transcriptional repressor Bm3R1		Bacillus cereus ATCC 14579
		Q9KV10	Q9KV10	Transcriptional regulator, family	TetR	Vibrio cholerae
		15640196		transcriptional regulator, family	TetR	Vibrio cholerae
125	ITGTYRYK	O08335	O08335	30S ribosomal protein S21		Bacillus subtilis
		16079771	yrhI	similar to transcriptional regulator (TetR/AcrR family)		Bacillus subtilis subsp. subtilis str. 168
126	MVGQYRYK	Q9AA02	Q9AA02	Transcriptional regulator, family	TetR	Caulobacter vibrioides
		Q8X7E0	Q8X7E0	Putative transcriptional repressor (Putative transcription regulatory protein)		Escherichia coli O157:H7
		Q8FHT3	Q8FHT3	Putative transcriptional repressor		Escherichia coli O6
		15801886		putative transcriptional repressor		Escherichia coli O157:H7 EDL933
		15831116		putative transcription regulatory protein		Escherichia coli O157:H7
		16125062		transcriptional regulator, family	TetR	Caulobacter crescentus CB15
127	MVGLYRYK	26247624		Putative transcriptional repressor		Escherichia coli CFT073
		Q8FLS9	Q8FLS9	Putative transcription regulator		Corynebacterium efficiens
		25029337		putative transcription regulator		Corynebacterium efficiens YS-314
128	MLGLYRYK	Q8F2I1	Q8F2I1	Transcriptional regulator, family	tetR	Leptospira interrogans
		24215490		Transcriptional regulator, family	tetR	Leptospira interrogans serovar lai str. 56601
129	MAAAYRYK	Q985C9	Q985C9	Transcriptional regulator		Mesorhizobium loti
		13476418		transcriptional regulator		Mesorhizobium loti
130	MAGAYRYK	Q9S1P1	Q9S1P1	Putative transcriptional regulator		Streptomyces coelicolor
131	MIATYHYK	Q983S2	Q983S2	Transcriptional regulator		Mesorhizobium loti
		13476778		transcriptional regulator		Mesorhizobium loti
132	MTGTYHYK	Q8YR12	Q8YR12	Transcriptional regulator		Nostoc sp. PCC 7120
		Q8F3Y9	Q8F3Y9	Transcriptional regulator, family	tetR	Leptospira interrogans
		17231138		transcriptional regulator		Nostoc sp. PCC 7120
		24214960		Transcriptional regulator, family	tetR	Leptospira interrogans serovar lai str. 56601
133	MVGLYTYK	Q8RG45	Q8RG45	Regulatory protein betI		Fusobacterium nucleatum subsp. nucleatum
		EAA23252	EAA23252	Regulatory protein betI		Fusobacterium nucleatum subsp. vincentii ATCC 49256
		19703808		Regulatory protein betI		Fusobacterium nucleatum subsp. nucleatum ATCC 25586
134	VVGSYQYK	Q8KU48	Q8KU48	EF0114		Enterococcus faecalis
		AAO80431	AAO80431	Transcriptional regulator,	TetR	Enterococcus faecalis

		29375207		family transcriptional regulator, TetR family	Enterococcus faecalis V583
135	TIGSYQYK	Q9I3U1 Q8YLK2 Q8Y170	Q9I3U1 Q8YLK2 Q8Y170	Probable transcriptional regulator Transcriptional regulator Putative transcription regulator protein	Pseudomonas aeruginosa Nostoc sp. PCC 7120 Ralstonia solanacearum
		15596600		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		17232788 17545539	RSc0820	transcriptional regulator PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Nostoc sp. PCC 7120 Ralstonia solanacearum
136	TVGSYQYK	Q98HF5 Q98B81 13472558 13474736 27379341	Q98HF5 Q98B81	Transcriptional regulator Transcriptional regulator transcriptional regulator transcriptional regulator transcriptional regulatory protein	Mesorhizobium loti Mesorhizobium loti Mesorhizobium loti Mesorhizobium loti Bradyrhizobium japonicum USDA 110
		27380101		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
137	TVGSYQFK	Q8YTB0 17230305	Q8YTB0	Transcriptional regulator transcriptional regulator	Nostoc sp. PCC 7120 Nostoc sp. PCC 7120
138	TIGTYEYK	Q9K6D2	Q9K6D2	Transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
		Q8EM51	Q8EM51	Transcriptional regulator (TetR/AcrR family)	Oceanobacillus iheyensis
		AAP29227	AAP29227	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP12203	AAP12203	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		15616359	BH3797	transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
		30265364		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		23100464		transcriptional regulator	Oceanobacillus iheyensis HTE831
		30023371		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
139	TIGTYRYK	Q9A7Q6	Q9A7Q6	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16125910		transcriptional regulator, TetR family	Caulobacter crescentus CB15
140	TIGSYRFK	Q8CK21	Q8CK21	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
		BAC74300	BAC74300	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29833131		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		31340544	SCO1712	putative TetR-family transcriptional regulator	Streptomyces coelicolor A3(2)
141	SIGSYSYK	Q97D02	Q97D02	Transcriptional regulator, AcrR family	Clostridium acetobutylicum
		15896919		Transcriptional regulator, AcrR family	Clostridium acetobutylicum
142	MIGSYQFK	Q98F48	Q98F48	Transcriptional regulator (tetR-family)	Mesorhizobium loti
		13473366		transcriptional regulator (tetR-family)	Mesorhizobium loti
143	FIGTYQFK	Q8NMS5	Q8NMS5	Transcriptional regulator	Corynebacterium glutamicum
		Q8FMW1 23308944	Q8FMW1	Putative transcription regulator transcriptional regulator	Corynebacterium efficiens Corynebacterium glutamicum ATCC 13032
		25028944		putative transcription regulator	Corynebacterium efficiens YS-314
144	MVGNYNYK	O66658 15605841	O66658	Hypothetical protein AQ_321 putative protein	Aquifex aeolicus Aquifex aeolicus VF5
145	MLSNYNYK	29347892		transcriptional regulator	Bacteroides thetaiotaomicron VPI-5482
146	MVSNYHYK	29345917		transcriptional regulator (TetR/AcrR family)	Bacteroides thetaiotaomicron VPI-5482
147	VVGNYKYK	Q97KU5	Q97KU5	Transcriptional regulator (TetR/AcrR family)	Clostridium acetobutylicum
		15894108		Transcriptional regulator (TetR/AcrR family)	Clostridium acetobutylicum
148	MVGNYYYK	Q9KZ96	Q9KZ96	Putative tetR-family transcriptional	Streptomyces coelicolor

		BAC71119	BAC71119	regulator		
		29829950		Putative TetR-family transcriptional regulator		<i>Streptomyces avermitilis</i>
149	IAGNYYYK	BAC72494	BAC72494	putative TetR-family transcriptional regulator		<i>Streptomyces avermitilis</i> MA-4680
		29831325		Putative TetR-family transcriptional regulator		<i>Streptomyces avermitilis</i>
150	LAGSYTYK	BAC69092	BAC69092	putative TetR-family transcriptional regulator		<i>Streptomyces avermitilis</i> MA-4680
		29827923		Putative TetR-family transcriptional regulator		<i>Streptomyces avermitilis</i>
151	LMGGYTYK	27382977		transcriptional regulatory protein		<i>Bradyrhizobium japonicum</i> USDA 110
152	MAGTYLYK	P44923	Y893_HAEIN	Hypothetical transcriptional regulator HI0893		<i>Haemophilus influenzae</i>
153	MKGTYYVK	16272831		transcriptional repressor		<i>Haemophilus influenzae</i> Rd
		Q98D84	Q98D84	Transcriptional regulator (TetR family)		<i>Mesorhizobium loti</i>
		Q8U7H2	Q8U7H2	Transcriptional regulator, TetR family		<i>Agrobacterium tumefaciens</i> str. C58
		Q92LW3	Q92LW3	Putative transcription regulator protein		<i>Sinorhizobium meliloti</i>
		Q8YAY4	Q8YAY4	Transcriptional regulator, TETR family		<i>Brucella melitensis</i>
		Q8FXG0	Q8FXG0	Transcriptional regulator, TetR family		<i>Brucella melitensis</i> biovar Suis
		13474033		transcriptional regulator (TetR family)		<i>Mesorhizobium loti</i>
		15966653		PUTATIVE TRANSCRIPTION REGULATORY PROTEIN		<i>Sinorhizobium meliloti</i>
		15890508		AGR_L_784p		<i>Agrobacterium tumefaciens</i> str. C58 (Cereon)
		17938166		transcriptional regulator, TetR family		<i>Agrobacterium tumefaciens</i> str. C58 (U. Washington)
		17989462		TRANSCRIPTIONAL REGULATOR, TETR FAMILY		<i>Brucella melitensis</i> 16M
		23499882		transcriptional regulator, TetR family		<i>Brucella suis</i> 1330
		27381734		transcriptional regulatory protein		<i>Bradyrhizobium japonicum</i> USDA 110
154	MKGTYYLYK	Q9RXD6	Q9RXD6	Transcriptional regulator, TetR family		<i>Deinococcus radiodurans</i>
		Q8TMB8	Q8TMB8	Hypothetical protein MA2746		<i>Methanosarcina acetivorans</i>
		15805406	DR0378	transcriptional regulator, TetR family		<i>Deinococcus radiodurans</i>
		20091569		conserved hypothetical protein		<i>Methanosarcina acetivorans</i> C2A
155	VKGTYYVK	Q8U9D0	Q8U9D0	Transcriptional regulator, TetR family		<i>Agrobacterium tumefaciens</i> str. C58
		15891152		AGR_L_2069p		<i>Agrobacterium tumefaciens</i> str. C58 (Cereon)
		17937503		transcriptional regulator, TetR family		<i>Agrobacterium tumefaciens</i> str. C58 (U. Washington)
156	VDGTYYLYK	P94548	P94548	Hypothetical protein ysiA		<i>Bacillus subtilis</i>
		Q9K8A4	Q9K8A4	Transcriptional regulator (TetR/AcrR family)		<i>Bacillus halodurans</i>
		Q8EPI4	Q8EPI4	Transcriptional regulator (TetR/AcrR family)		<i>Oceanobacillus iheyensis</i>
		AAP28454	AAP28454	Transcriptional regulator, TetR family		<i>Bacillus anthracis</i> str. Ames
		AAP11434	AAP11434	Transcriptional regulator, TetR family		<i>Bacillus cereus</i> ATCC 14579
		16079907	ysiA	similar to transcriptional regulator (TetR/AcrR family)		<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168
		15615664	BH3102	transcriptional regulator (TetR/AcrR family)		<i>Bacillus halodurans</i>
		30264591		transcriptional regulator, TetR family		<i>Bacillus anthracis</i> str. Ames
		23099576		transcriptional regulator		<i>Oceanobacillus iheyensis</i> HTE831

		30022602		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
157	IQGTYLYK	O07001 AAP24837	O07001 AAP24837	Hypothetical protein yvdt Transcriptional regulator, TetR family	Bacillus subtilis Bacillus anthracis str. Ames
		AAP07841	AAP07841	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		16080501	yvDT	similar to transcriptional regulator (TetR/AcrR family)	Bacillus subtilis subsp. subtilis str. 168
		30260974		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30019009		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
158	VQGTLYLYK	Q9KFA5	Q9KFA5	Transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
		15613143	BH0580	transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
159	IKGTYLYK	Q98I65 AAN10243	Q98I65 AAN10243	Probable transcriptional regulator Transcriptional regulator	Mesorhizobium loti Streptomyces viridifaciens
160	VKGTLYLYK	13472298 Q9UYP3	Q9UYP3	probable transcriptional regulator Hypothetical protein PAB0976	Mesorhizobium loti Pyrococcus abyssi
161	LKGTLYLYK	14521663 Q8Y1J2	Q8Y1J2	hypothetical protein Putative transcription regulator protein	Pyrococcus abyssi Ralstonia solanacearum
		17545417	RSc0698	PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum
162	IRGTYYIK	O66635	O66635	Transcriptional regulator (TETR/ACRR family)	Aquifex aeolicus
		15605818	acrR3	transcriptional regulator (TetR/AcrR family)	Aquifex aeolicus VF5
163	VKGAYIYK	P40950	YUXN_BACSU	Hypothetical transcriptional regulator yuxN	Bacillus subtilis
		16080356	yuxN	similar to hypothetical proteins from B. subtilis	Bacillus subtilis subsp. subtilis str. 168
164	VKGAYLYK	AAP29317	AAP29317	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP12296	AAP12296	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30265454		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30023464		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
165	VRGAYFYK	Q9S1R1	Q9S1R1	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
166	IHGTYTYK	Q9KFQ0	Q9KFQ0	Transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
		15612991	BH0428	transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
167	IHGTYNYK	AAP25822	AAP25822	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30261959		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
168	IYGTIVYK	Q8ES85 AAP26590	Q8ES85 AAP26590	Hypothetical conserved protein Transcriptional regulator, TetR family	Oceanobacillus iheyensis Bacillus anthracis str. Ames
		AAP09713	AAP09713	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30262727		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		23098212		hypothetical protein	Oceanobacillus iheyensis HTE831
		30020881		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
169	MHGAYLYK	27379144		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
170	LKGSYHYK	P76189	YDHM_ECOLI	Hypothetical transcriptional regulator ydhM	Escherichia coli O6
		Q9I1S1 Q8XFC3	Q9I1S1 Q8XFC3	Probable transcriptional regulator Putative transcriptional repressor (TetR/AcrR family) (Putative TetR-family transcriptional regulator)	Pseudomonas aeruginosa Salmonella typhimurium
		Q8X632 Q8ZE12	Q8X632 Q8ZE12	Orf, hypothetical protein TetR-family transcriptional regulatory protein (Hypothetical protein)	Salmonella typhi Escherichia coli O157:H7 Yersinia pestis
		Q8E926	Q8E926	Transcriptional regulator, TetR	Shewanella oneidensis

		Q8D6M3 AAP17143	Q8D6M3 AAP17143	family Transcriptional regulator Hypothetical protein	Vibrio vulnificus Shigella flexneri 2a str. 2457T
		AAO68955	AAO68955	Putative TetR-family transcriptional regulator	Salmonella typhi
		16129607 15597392	b1649	orf, hypothetical protein probable transcriptional regulator	Escherichia coli K12 Pseudomonas aeruginosa PAO1
		15802063		orf, hypothetical protein	Escherichia coli O157:H7 EDL933
		15831612 26989198		hypothetical protein transcriptional regulator, TetR family	Escherichia coli O157:H7 Pseudomonas putida KT2440
		16122601		TetR-family transcriptional regulatory protein	Yersinia pestis CO92
		16764785	ydhM	putative transcriptional repressor (TetR/AcrR family)	Salmonella typhimurium LT2
		16760476		putative TetR-family transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi
		22125851 24375946		hypothetical protein transcriptional regulator, TetR family	Yersinia pestis KIM Shewanella oneidensis MR-1
		26247897	ydhM	Hypothetical transcriptional regulator ydhM	Escherichia coli CFT073
		27366939 28901270		Transcriptional regulator putative transcriptional regulator	Vibrio vulnificus CMCP6 Vibrio parahaemolyticus RIMD 2210633
		29141764		putative TetR-family transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi Ty2
		30063165		hypothetical protein	Shigella flexneri 2a str. 2457T
171	VKGSYHYK	Q9I5A1 15596036	Q9I5A1	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
172	LKGAYHYK	Q98LR7 Q97D74	Q98LR7 Q97D74	Probable transcriptional regulator Transcriptional regulator, AcrR family	Mesorhizobium loti Clostridium acetobutylicum
		13471041 15896840		probable transcriptional regulator Transcriptional regulator, AcrR family	Mesorhizobium loti Clostridium acetobutylicum
173	LKGSYYYK	Q8Y043	Q8Y043	Probable transcription regulator protein	Ralstonia solanacearum
		17545920	RSc1201	PROBABLE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum
		28870251		transcriptional regulator, TetR family	Pseudomonas syringae pv. tomato str. DC3000
174	IKGTYHYK	Q9I429 15596512	Q9I429	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
175	VKGTYYHYK	AAP26340	AAP26340	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP09373	AAP09373	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30262477		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30020541		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
176	VQGTYYHYK	Q8TM84	Q8TM84	Transcriptional regulator, TetR family	Methanosarcina acetivorans
		20091605		transcriptional regulator, TetR family	Methanosarcina acetivorans C2A
177	VKGAYHYK	Q92WI6	Q92WI6	Putative transcriptional regulator protein	Sinorhizobium meliloti
		Q9S261	Q9S261	Putative tetR transcriptional regulator	Streptomyces coelicolor
		BAC74324	BAC74324	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		16264101		putative transcriptional regulator protein	Sinorhizobium meliloti
		29833155		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
178	IKGQYYYYK	P42097	YPB3_LACLA	Hypothetical 22.2 kDa protein in lacX 3'region (ORF3)	Lactococcus lactis subsp. lactis
179	IKGTYYYYK	Q8ESG2	Q8ESG2	Transcriptional regulator (TetR/AcrR family)	Oceanobacillus iheyensis

		Q8EN16	Q8EN16	Transcriptional (TetR/AcrR family) transcriptional regulator	regulator	Oceanobacillus iheyensis
		23098132				Oceanobacillus iheyensis HTE831
		23100129		transcriptional regulator		Oceanobacillus iheyensis HTE831
180	IKGSYYFK	Q8UGP2	Q8UGP2	Transcriptional family AGR_C_1821p	regulator, TetR	Agrobacterium tumefaciens str. C58
		15888336				Agrobacterium tumefaciens str. C58 (Cereon)
		17934901		transcriptional regulator, family	TetR	Agrobacterium tumefaciens str. C58 (U. Washington)
181	TKGSYYFK	BAC69234	BAC69234	Putative TetR-family transcriptional regulator		Streptomyces avermitilis
		29828065		putative TetR-family transcriptional regulator		Streptomyces avermitilis MA-4680
182	IKGSYSYK	Q9KBN8	Q9KBN8	Hypothetical protein BH1887		Bacillus halodurans
		15614450	BH1887	BH1887~unknown conserved protein		Bacillus halodurans
183	IKGSYTYK	Q8RE54	Q8RE54	Transcriptional family	regulator, TetR	Fusobacterium nucleatum subsp. nucleatum
		19704607		Transcriptional family	regulator, TetR	Fusobacterium nucleatum subsp. nucleatum ATCC 25586
184	IKGSYQYK	Q97IV4	Q97IV4	Transcriptional family	regulator, AcrR	Clostridium acetobutylicum
		15894814		Transcriptional family	regulator, AcrR	Clostridium acetobutylicum
185	IRGSYQYK	Q8ELR5	Q8ELR5	Transcriptional (TetR/AcrR family)	regulator	Oceanobacillus iheyensis
		AAO82305	AAO82305	Transcriptional family	regulator, TetR	Enterococcus faecalis
		23100608		transcriptional regulator		Oceanobacillus iheyensis HTE831
		29377081		transcriptional family	regulator, TetR	Enterococcus faecalis V583
186	LKGSFQYK	P96415	P96415	Transcriptional family	regulator, TetR	Mycobacterium tuberculosis
		15607373		hypothetical protein Rv0232		Mycobacterium tuberculosis H37Rv
		15839612		transcriptional family	regulator, TetR	Mycobacterium tuberculosis CDC1551
187	VKGTYNKYK	Q9EUW7	Q9EUW7	Putative transcriptional (Fragment)	regulator	Listeria innocua
		Q8VMX5	Q8VMX5	Putative transcriptional (Fragment)	regulator	Listeria monocytogenes
		O53054	O53054	ORF Z protein		Listeria ivanovii
		Q927Q4	Q927Q4	Hypothetical protein lin2734		Listeria innocua
		Q8Y466	Q8Y466	Hypothetical protein lmo2589		Listeria monocytogenes
		16804627		similar to transcription TetR/AcrR family	regulator	Listeria monocytogenes EGD-e
		16801795		similar to transcription TetR/AcrR family	regulator	Listeria innocua
188	IKGTYNKYK	Q8GA63	Q8GA63	Putative UidR transcriptional regulator		Escherichia coli
189	VKGTFNKYK	Q9K7E7	Q9K7E7	NADH dehydrogenase		Bacillus halodurans
		AAP26049	AAP26049	Transcriptional family	regulator, TetR	Bacillus anthracis str. Ames
		15615977	BH3415	NADH dehydrogenase		Bacillus halodurans
		30262186		transcriptional family	regulator, TetR	Bacillus anthracis str. Ames
190	IKGSYNKYK	Q8EXV3	Q8EXV3	Putative TetR-family transcriptional regulator		Leptospira interrogans
		24217165		putative TetR-family transcriptional regulator		Leptospira interrogans serovar lai str. 56601
191	VKGSYNKYK	Q8ESU6	Q8ESU6	Transcriptional (TetR/AcrR family)	regulator	Oceanobacillus iheyensis
		23097975		transcriptional regulator		Oceanobacillus iheyensis HTE831
192	VKGTYYKYK	Q59213	Q59213	BM1P1		Bacillus megaterium
		O68276	O68276	Putative DNA-binding protein Bm1P1		Bacillus megaterium
193	VKGTYYGYK	O07254	O07254	Hypothetical 25.9 kDa protein (Transcriptional regulator,	TetR	Mycobacterium tuberculosis

		15607471		family) hypothetical protein Rv0330c	Mycobacterium tuberculosis H37Rv
		15839717		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
194	VSGTYNYK	Q9AA36	Q9AA36	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16125025		transcriptional regulator, TetR family	Caulobacter crescentus CB15
195	ITGTYNYK	Q8XP87	Q8XP87	Probable transcriptional regulator	Clostridium perfringens
		18309060		probable transcriptional regulator	Clostridium perfringens str. 13
196	ITGTYNYR	Q8NR95	Q8NR95	Transcriptional regulator	Corynebacterium glutamicum
		19552381		transcriptional regulator	Corynebacterium glutamicum ATCC 13032
197	IIGTYNYK	Q98N96	Q98N96	Hypothetical protein mlr0237	Mesorhizobium loti
		13470510		unknown protein	Mesorhizobium loti
198	IVGTYNYK	28212055		putative transcription regulator	Clostridium tetani E88
199	MKGTYYFK	P32398	YHGD_BACSU	Hypothetical transcriptional regulator yhgD	Bacillus subtilis
		AAP25059	AAP25059	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP08064	AAP08064	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		16078079	yhgD	alternate gene name: yixD~similar to transcriptional regulator (TetR/AcrR family)	Bacillus subtilis subsp. subtilis str. 168
		30261196		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30019232		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
200	IKGTYTAK	Q9KAZ1	Q9KAZ1	Transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
		15614708	BH2145	transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
201	MHGLYYYYK	Q8GCB5	Q8GCB5	BtrW protein	Bacillus circulans
202	MKGLYYYYK	Q9RKY8	Q9RKY8	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
		BAC74628	BAC74628	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29833459		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
203	MKALYYYYK	Q9WZT0	Q9WZT0	Transcriptional regulator, TETR family	Thermotoga maritima
		O27135	O27135	Transcriptional regulator related protein	Methanothermobacter thermautotrophicus str. Delta H
		15643586		transcriptional regulator, TetR family	Thermotoga maritima
		15679074		transcriptional regulator related protein	Methanothermobacter thermautotrophicus str. Delta H
204	VKALYYYYK	Q8TK36	Q8TK36	Transcriptional regulator, TetR family	Methanosarcina acetivorans
		20092389		transcriptional regulator, TetR family	Methanosarcina acetivorans C2A
205	VYGLYWYK	O08377	O08377	Hypothetical 24.5 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15608672		hypothetical protein Rv1534	Mycobacterium tuberculosis H37Rv
		15841000		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
206	IYGLYFYK	Q97VV1	Q97VV1	Transcriptional regulator, putative	Sulfolobus solfataricus
		15899244		Transcriptional regulator, putative	Sulfolobus solfataricus
207	IYGVYYYYK	Q976W7	Q976W7	Hypothetical protein ST0072	Sulfolobus tokodaii
		15920251		194aa long conserved hypothetical protein	Sulfolobus tokodaii
208	ITGLYYYYK	AAP26746	AAP26746	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP09863	AAP09863	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30262883		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30021031		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579

209	MDAAYYK	Q9A9I8 16125244	Q9A9I8	Transcriptional family transcriptional family	regulator, regulator,	TetR TetR	Caulobacter vibrioides Caulobacter crescentus CB15
210	INGIYKYK	Q8RFA3 EAA24963 19704148	Q8RFA3 EAA24963	Transcriptional family Transcriptional family	regulator, regulator,	TetR TetR	Fusobacterium nucleatum subsp. nucleatum Fusobacterium nucleatum subsp. vincentii ATCC 49256 Fusobacterium nucleatum subsp. nucleatum ATCC 25586
211	MKATYNYK	Q9KS52 Q8XRL6 15641419 17549036	Q9KS52 Q8XRL6 RSp0815	Transcriptional family Probable transcription protein	regulator, regulator,	TetR TetR	Vibrio cholerae Ralstonia solanacearum Vibrio cholerae
212	MKATYSYK	28870736		PROBABLE TRANSCRIPTION REGULATORY PROTEIN transcriptional family	regulator,	TetR	Pseudomonas syringae pv. tomato str. DC3000
213	MKATYAYK	Q8P7D1 21232111 27379433	Q8P7D1	Transcriptional family transcriptional family transcriptional family	regulator regulator regulatory protein		Xanthomonas campestris pv. campestris Xanthomonas campestris pv. campestris str. ATCC 33913 Bradyrhizobium japonicum USDA 110
214	MKATYHYK	O53641 15607300 15839538	O53641	Hypothetical protein (Transcriptional family) hypothetical protein Rv0158	24.1 kDa regulator,	protein TetR	Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv Mycobacterium tuberculosis CDC1551
215	MKATYYYK	AAP26269 AAP09304 30262406 30020472	AAP26269 AAP09304	Transcriptional family Transcriptional family transcriptional family Transcriptional family	regulator, regulator, regulator, regulator,	TetR TetR TetR TetR	Bacillus anthracis str. Ames Bacillus cereus ATCC 14579 Bacillus anthracis str. Ames Bacillus cereus ATCC 14579
216	MKATYVYK	Q8DVA4 24379070	Q8DVA4	Putative transcriptional family putative transcriptional family	regulator regulator		Streptococcus mutans Streptococcus mutans UA159
217	MKATYLYK	Q8TMA4 Q8TMA0 Q8TLX1 Q8PU11 20091584 20091589 20091723 21228648	Q8TMA4 Q8TMA0 Q8TLX1 Q8PU11 Q8TMA2	Transcriptional family Transcriptional family Transcriptional family Hypothetical protein MM2546 transcriptional family transcriptional family transcriptional family Conserved hypothetical protein	regulator, regulator, regulator, regulator, regulator, regulator, regulator,	TetR TetR TetR TetR TetR TetR TetR TetR	Methanosarcina acetivorans Methanosarcina acetivorans Methanosarcina acetivorans Methanosarcina mazei Methanosarcina acetivorans C2A Methanosarcina acetivorans C2A Methanosarcina acetivorans C2A Methanosarcina acetivorans C2A Methanosarcina mazei Goe1 Methanosarcina acetivorans
218	MKAAVLYK	Q8TMA2 20091587	Q8TMA2	Transcriptional family transcriptional family	regulator, regulator,	TetR TetR	Methanosarcina acetivorans Methanosarcina acetivorans C2A
219	MKKTYSYK	Q97FH1 15896023	Q97FH1	Transcriptional family Transcriptional family	regulators, regulators,	AcrR AcrR	Clostridium acetobutylicum Clostridium acetobutylicum
220	MKQTYSYK	Q8PJN3 21243230	Q8PJN3	Transcriptional family transcriptional family	regulator regulator	tetR/acrR tetR/acrR	Xanthomonas axonopodis pv. citri Xanthomonas axonopodis pv. citri str. 306
221	MKTTYSYK	Q8YPB9 17231771	Q8YPB9	Transcriptional family transcriptional family	regulator regulator		Nostoc sp. PCC 7120 Nostoc sp. PCC 7120

222	MKRTYAYK	O32228 16080411	O32228 yvaF	YVAF protein yvaF	Bacillus subtilis Bacillus subtilis subsp. subtilis str. 168
223	MKRTYGYK	Q8PIQ0	Q8PIQ0	Transcriptional regulator	Xanthomonas axonopodis pv. citri
		Q8P7C7	Q8P7C7	Transcriptional regulator	Xanthomonas campestris pv. campestris
		21232115		transcriptional regulator	Xanthomonas campestris pv. campestris str. ATCC 33913
		21243572		transcriptional regulator	Xanthomonas axonopodis pv. citri str. 306
224	MRRTYTYK	29349178		transcriptional regulator	Bacteroides thetaiotaomicron VPI-5482
225	MTATYAYK	Q9A9Q5	Q9A9Q5	Transcriptional regulator, family	TetR Caulobacter vibrioides
		16125171		transcriptional regulator, family	TetR Caulobacter crescentus CB15
226	MSSTVYK	28377011		transcription regulator	Lactobacillus plantarum WCFS1
227	MKTVYRYK	Q9CB41	Q9CB41	Possible TetR-family transcriptional regulator	Mycobacterium leprae
		O53757	O53757	Putative regulatory protein (Tran- scriptional regulator, TetR family)	Mycobacterium tuberculosis
		15607613		hypothetical protein Rv0472c	Mycobacterium tuberculosis H37Rv
		15828330		possible TetR-family transcriptional regulator	Mycobacterium leprae
		15839861		transcriptional regulator, family	TetR Mycobacterium tuberculosis CDC1551
228	MKGAFRYK	28379529		transcription regulator	Lactobacillus plantarum WCFS1
229	VKAQYAYK	Q9KXX1	Q9KXX1	Putative transcriptional regulator	Streptomyces coelicolor
230	VKERYAYK	Q9EWE9	Q9EWE9	Putative regulatory protein	Streptomyces coelicolor
231	IKQLYAYK	Q9ADP7	Q9ADP7	Putative tetR family regulatory protein	Streptomyces coelicolor
232	FTGTYKFK	Q8EH16	Q8EH16	Transcriptional regulator, family	TetR Shewanella oneidensis
		24372993		transcriptional regulator, family	TetR Shewanella oneidensis MR-1
233	FTNTYKYK	Q8EBA4	Q8EBA4	Transcriptional regulator, family	TetR Shewanella oneidensis
		24375127		transcriptional regulator, family	TetR Shewanella oneidensis MR-1
234	LKGSYHFG	P42105	YXAF_BACSU	Hypothetical transcriptional regula- tor yxaF	Bacillus subtilis
		Q8UJT9	Q8UJT9	Transcriptional regulator, family	TetR Agrobacterium tumefaciens str. C58
		16081051	yxaF	similar to hypothetical proteins	Bacillus subtilis subsp. subtilis str. 168
		16119613		AGR_pAT_558p	Agrobacterium tumefaciens str. C58 (Cereon)
		17938970		transcriptional regulator, family	TetR Agrobacterium tumefaciens str. C58 (U. Washington)
235	LKGSYHYG	Q9I643	Q9I643	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
236	IRGSYHYG	O53647	O53647	Possible regulatory protein (Tran- scriptional regulator, putative) hypothetical protein Rv0196	Mycobacterium tuberculosis
		15607337			Mycobacterium tuberculosis H37Rv
		15839576		transcriptional regulator, putative	Mycobacterium tuberculosis CDC1551
237	LKGSYYHG	O34619	O34619	LMRA	Bacillus subtilis
		16077337	lmrA	transcriptional regulator	Bacillus subtilis subsp. subtilis str. 168
238	LKGSYYYG	Q97DX5	Q97DX5	Transcriptional regulator, family	AcrR Clostridium acetobutylicum
		Q8TU95	Q8TU95	Transcriptional regulator, family	TetR Methanosarcina acetivorans
		15896588		Transcriptional regulator, family	AcrR Clostridium acetobutylicum
		20089076		transcriptional regulator, family	TetR Methanosarcina acetivorans C2A
239	LRGSYFHG	Q983S7	Q983S7	Transcriptional regulator	Mesorhizobium loti
		13476773		transcriptional regulator	Mesorhizobium loti
240	LKSQFHYG	Q9RJM0	Q9RJM0	Putative regulatory protein	Streptomyces coelicolor

241	ILGSYHHG	Q8XRN5 17549017	Q8XRN5 RSp0796	Putative transcription regulator protein PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum Ralstonia solanacearum
242	ILGSYHFG	BAC74878 29833709	BAC74878	Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces avermitilis Streptomyces avermitilis MA-4680
243	VLGSYHHI	Q9F2Q9	Q9F2Q9	Putative tetR-family transcriptional regulatory protein	Streptomyces coelicolor
244	VPRTFRHK	P39885	TCMR_STRGA	Tetracenomycin C transcriptional repressor	Streptomyces glaucescens
245	VRRTFRHK	Q9AJ68 Q9ZBK0 BAC69621 29828452	Q9AJ68 Q9ZBK0 BAC69621	Putative transcriptional repressor ButR Putative transcriptional regulator Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces cinnamomensis Streptomyces coelicolor Streptomyces avermitilis Streptomyces avermitilis MA-4680
246	VERTFRHK	Q9RJL4	Q9RJL4	Putative tetR family transcriptional regulator	Streptomyces coelicolor
247	VRRTSRHK	Q9S5E8 Q9KYM3	Q9S5E8 Q9KYM3	Orf1, orf2, orf3, orf4, orf5 genes Putative TetR-family transcriptional regulator	Streptomyces coelicolor Streptomyces coelicolor
248	IRRTYRHR	Q8E4M0 Q8DZ12 22537457 25011423	Q8E4M0 Q8DZ12	Hypothetical protein Transcriptional regulator, TetR family transcriptional regulator, TetR family unknown	Streptococcus agalactiae serogroup III Streptococcus agalactiae serogroup V Streptococcus agalactiae 2603V/R Streptococcus agalactiae NEM316
249	MERTYRHR	27381254		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
250	LRRTFGHR	Q54185	Q54185	Hypothetical protein SCO5956	Streptomyces coelicolor
251	MRRTYGHR	BAC70525 29829356	BAC70525	Hypothetical protein hypothetical protein	Streptomyces avermitilis Streptomyces avermitilis MA-4680
252	VRRTFRYK	Q9ZBA0 P95037 15607831 15840095	Q9ZBA0 P95037	Hypothetical 20.9 kDa protein Hypothetical 22.3 kDa protein (Transcriptional regulator, TetR family) hypothetical protein Rv0691c transcriptional regulator, TetR family	Rhodococcus erythropolis Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv Mycobacterium tuberculosis CDC1551
253	VRRSFRYK	Q9XDF0 Q9RK42 Q9EWU3	Q9XDF0 Q9RK42 Q9EWU3	Putative transcriptional repressor Putative transcriptional regulator Putative regulatory protein	Streptomyces griseus subsp. griseus Streptomyces coelicolor Streptomyces coelicolor
254	VQRTFRYK	Q9RDR0 BAC73508 29832339	Q9RDR0 BAC73508	Putative transcriptional regulator Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces coelicolor Streptomyces avermitilis Streptomyces avermitilis MA-4680
255	IRRTYRYK	Q8DWK7 24378566	Q8DWK7	Putative transcriptional regulator putative transcriptional regulator	Streptococcus mutans Streptococcus mutans UA159
256	VKRTYRYK	28378381		transcription regulator, N-terminal fragment (putative)	Lactobacillus plantarum WCFS1
257	TKRTFRYK	Q9RJM3	Q9RJM3	Putative tetR family transcriptional regulator	Streptomyces coelicolor
258	LRRTFYK	Q8UK34 16119518 17938874	Q8UK34	Transcriptional regulator, TetR family AGR_pAT_416p transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 Agrobacterium tumefaciens str. C58 (Cereon) Agrobacterium tumefaciens str. C58 (U. Washington)
259	LRRTFHYK	Q8Y1Q4 17545354	Q8Y1Q4 RSc0635	Probable transcription regulator protein PROBABLE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum Ralstonia solanacearum
260	IRRSFHYK	27376363		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
261	IKRTFNYK	Q8NTC8	Q8NTC8	Transcriptional regulator	Corynebacterium glutamicum

		19551625		transcriptional regulator	<i>Corynebacterium glutamicum</i> ATCC 13032
262	VKRTFNYK	Q8FSJ4 25026953	Q8FSJ4	Hypothetical protein hypothetical protein	<i>Corynebacterium efficiens</i> <i>Corynebacterium efficiens</i> YS-314
263	VKRTFAYK	AAO06913	AAO06913	Putative transcriptional regulator	<i>Streptomyces hygroscopicus</i>
264	VKVTFKYK	Q9CGX8	Q9CGX8	Transcriptional regulator	<i>Lactococcus lactis</i> subsp. <i>lactis</i>
		15672946	yjjB	transcription regulator	<i>Lactococcus lactis</i> subsp. <i>lactis</i>
265	MKTTFKYK	Q8ESC2	Q8ESC2	Transcriptional regulator (TetR family)	<i>Oceanobacillus iheyensis</i>
		23098174		transcriptional regulator	<i>Oceanobacillus iheyensis</i> HTE831
266	VKPTFRYK	Q9FCH5	Q9FCH5	Putative TetR-family transcriptional regulator	<i>Streptomyces coelicolor</i>
		BAC71729	BAC71729	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29830560		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
267	ATRTFRYK	Q938D7	Q938D7	Putative transcription regulator	<i>Mycobacterium smegmatis</i>
268	APRTYRYK	Q9KYU4	Q9KYU4	Putative tetR-family transcriptional regulator	<i>Streptomyces coelicolor</i>
		BAC71405	BAC71405	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29830236		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
269	ALRTYRHK	P96248	P96248	Hypothetical 23.1 kDa protein (Transcriptional regulator, TetR family)	<i>Mycobacterium tuberculosis</i>
		15610966		hypothetical protein Rv3830c	<i>Mycobacterium tuberculosis</i> H37Rv
		15843454		transcriptional regulator, TetR family	<i>Mycobacterium tuberculosis</i> CDC1551
270	ALRTYRYK	Q9K4H1	Q9K4H1	Putative tetR-family transcriptional regulator	<i>Streptomyces coelicolor</i>
271	ARRTYKYK	Q9I5B1 15596025	Q9I5B1	Probable transcriptional regulator probable transcriptional regulator	<i>Pseudomonas aeruginosa</i> <i>Pseudomonas aeruginosa</i> PAO1
272	IRATYDHK	P46330	YXBF_BACSU	Hypothetical transcriptional regulator yxbF	<i>Bacillus subtilis</i>
		16081036	yxbF	alternate gene name: yxaT	<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168
273	VRSTYDHK	Q97DN9	Q97DN9	Transcriptional regulators, AcrR family	<i>Clostridium acetobutylicum</i>
		15896674		Transcriptional regulators, AcrR family	<i>Clostridium acetobutylicum</i>
274	IRSTYDHK	Q8CN97 27468852	Q8CN97	Transcriptional regulator transcriptional regulator	<i>Staphylococcus epidermidis</i> <i>Staphylococcus epidermidis</i> ATCC 12228
275	VRSTYSHK	Q9F8C4 Q8E6U3	Q9F8C4 Q8E6U3	Hypothetical 21.0 kDa protein Hypothetical protein	<i>Streptococcus agalactiae</i> <i>Streptococcus agalactiae</i> serogroup III
		Q8E1C6	Q8E1C6	Transcriptional regulator, TetR family	<i>Streptococcus agalactiae</i> serogroup V
		22536614		transcriptional regulator, TetR family	<i>Streptococcus agalactiae</i> 2603V/R
		25010539		Unknown	<i>Streptococcus agalactiae</i> NEM316
276	IRSTYSHK	Q8DTP9 24379700	Q8DTP9	Putative transcriptional regulator putative transcriptional regulator	<i>Streptococcus mutans</i> <i>Streptococcus mutans</i> UA159
277	VRSTYAHK	Q9A0B7 Q8P1F5	Q9A0B7 Q8P1F5	Hypothetical protein SPy0846 Hypothetical protein spyM18_0906	<i>Streptococcus pyogenes</i> <i>Streptococcus pyogenes</i> MGAS8232
		Q8K7X5	Q8K7X5	Hypothetical protein SpyM3_0572	<i>Streptococcus pyogenes</i> MGAS315
		15674880		conserved hypothetical protein	<i>Streptococcus pyogenes</i> M1 GAS
		19745921		conserved hypothetical protein	<i>Streptococcus pyogenes</i> MGAS8232
		21910108		conserved hypothetical protein	<i>Streptococcus pyogenes</i> MGAS315
		28896194		conserved hypothetical protein	<i>Streptococcus pyogenes</i> SSI-1
278	IRATYAHK	Q9CDI1	Q9CDI1	Transcriptional regulator	<i>Lactococcus lactis</i> subsp.

		AAO81117	AAO81117	Transcriptional regulator, TetR family		<i>lactis</i> <i>Enterococcus faecalis</i>
		15674221	yxcB	transcription regulator		<i>Lactococcus lactis</i> subsp. <i>lactis</i>
		29375893		transcriptional regulator, TetR family		<i>Enterococcus faecalis</i> V583
279	VRATYAHK	Q99R67	Q99R67	Hypothetical protein SAV2571 (MW2492 protein)		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50 <i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315 <i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2
		Q9RL08	Q9RL08	Putative tetR-family transcriptional regulator		<i>Streptomyces coelicolor</i>
		15928150		hypothetical protein, similar to transcriptional regulator (TetR/AcrR family)		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315
		15925561		hypothetical protein		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50
		21284221	MW2492	ORFID:MW2492~hypothetical protein, similar to transcriptional regulator (TetR/AcrR family)		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2
280	IRATYSHK	Q8DVN7 24378929	Q8DVN7	Putative transcriptional regulator putative transcriptional regulator		<i>Streptococcus mutans</i> <i>Streptococcus mutans</i> UA159
281	VRATYSHK	AAP28503	AAP28503	Transcriptional regulator, TetR family		<i>Bacillus anthracis</i> str. Ames
		AAP11477	AAP11477	Transcriptional regulator, TetR family		<i>Bacillus cereus</i> ATCC 14579
		30264640		transcriptional regulator, TetR family		<i>Bacillus anthracis</i> str. Ames
		30022645		Transcriptional regulator, TetR family		<i>Bacillus cereus</i> ATCC 14579
282	IRGTYSHK	Q8ELN7	Q8ELN7	Transcriptional regulator (tetR-family)		<i>Oceanobacillus iheyensis</i>
		23100639		transcriptional regulator		<i>Oceanobacillus iheyensis</i> HTE831
283	VRGTYAHK	AAP26926	AAP26926	Transcriptional regulator, TetR family		<i>Bacillus anthracis</i> str. Ames
		AAP10030	AAP10030	Transcriptional regulator, TetR family		<i>Bacillus cereus</i> ATCC 14579
		30263063		transcriptional regulator, TetR family		<i>Bacillus anthracis</i> str. Ames
		30021198		Transcriptional regulator, TetR family		<i>Bacillus cereus</i> ATCC 14579
284	VRSTYCHK	Q97NZ1	Q97NZ1	Transcriptional regulator, TetR family		<i>Streptococcus pneumoniae</i>
		Q8DNK2	Q8DNK2	Conserved hypothetical protein		<i>Streptococcus pneumoniae</i> R6
		15901686		transcriptional regulator, TetR family		<i>Streptococcus pneumoniae</i> TIGR4
		15903715		Conserved hypothetical protein		<i>Streptococcus pneumoniae</i> R6
285	VRSTYTHP	Q97TQ2	Q97TQ2	TetR family HTH transcriptional regulator		<i>Clostridium acetobutylicum</i>
		15004750		TetR family HTH transcriptional regulator		<i>Clostridium acetobutylicum</i>
286	VRSTYTHK	Q99R45	Q99R45	Hypothetical protein SAV2593		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50 <i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315 <i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2
		Q8NUN6	Q8NUN6	MW2513 protein		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315 <i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2
		15928172		hypothetical protein, similar to transcriptional regulator tetR-family		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315
		15925583		hypothetical protein		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50
		21284242	MW2513	ORFID:MW2513~hypothetical protein, similar to transcriptional regulator tetR-family		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2
287	IRGTYLHK	Q99RQ2	Q99RQ2	Hypothetical protein SAV2375 (MW2297 protein)		<i>Staphylococcus aureus</i> subsp. <i>aureus</i> Mu50 <i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315 <i>Staphylococcus aureus</i>

		15927955		hypothetical protein, similar to transcriptional regulator tetR-family	Staphylococcus aureus subsp. aureus MW2
		15925365		hypothetical protein	Staphylococcus aureus subsp. aureus N315
		21284026	MW2297	ORFID:MW2297~hypothetical protein, similar to transcriptional regulator tetR-family	Staphylococcus aureus subsp. aureus Mu50
288	VRGTYLHK	Q8P0U1	Q8P0U1	Putative transcriptional regulator	Streptococcus pyogenes MGAS8232
		Q99ZE9	Q99ZE9	Putative transcriptional regulator (TetR/AcrR family)	Streptococcus pyogenes MGAS315
		Q8CQH0 15675217	Q8CQH0	Transcriptional regulator tetR-family putative transcriptional regulator (TetR/AcrR family)	Staphylococcus epidermidis Streptococcus pyogenes M1 GAS
		19746186		putative transcriptional regulator	Streptococcus pyogenes MGAS8232
		21910429		putative transcriptional regulator (TetR/AcrR family)	Streptococcus pyogenes MGAS315
		27467011		transcriptional regulator tetR-family	Staphylococcus epidermidis ATCC 12228
		28378292		transcription regulator	Lactobacillus plantarum WCFS1
		28896005		putative transcriptional regulator (TetR/AcrR family)	Streptococcus pyogenes SSI-1
289	VRGTYVHK	Q8ESJ8	Q8ESJ8	Transcriptional regulator (tetR-family)	Oceanobacillus iheyensis
		23098091		transcriptional regulator	Oceanobacillus iheyensis HTE831
290	LKGYLHK	27380762		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
291	VRSTYIHK	Q8ELC0 23100766	Q8ELC0	Hypothetical protein	Oceanobacillus iheyensis
				hypothetical protein	Oceanobacillus iheyensis HTE831
292	IRSTYLHK	Q8DTZ2 24379598	Q8DTZ2	Putative transcriptional regulator putative transcriptional regulator	Streptococcus mutans UA159
293	IRATYLHK	26988668		hypothetical protein	Pseudomonas putida KT2440
294	LRMTYRHK	Q9CDH1	Q9CDH1	Transcriptional regulator	Lactococcus lactis subsp. lactis
		15674231	yxD	transcription regulator	Lactococcus lactis subsp. lactis
295	IRATYRHK	Q9KZX2	Q9KZX2	Putative transcriptional regulator	Streptomyces coelicolor
296	VRTTYKHK	Q99R83	Q99R83	Hypothetical protein SAV2553	Staphylococcus aureus subsp. aureus Mu50
		Q8NUR3	Q8NUR3	MW2474 protein	Staphylococcus aureus subsp. aureus N315
		15928132		hypothetical protein, similar to transcriptional regulator tetR-family	Staphylococcus aureus subsp. aureus MW2
		15925543		hypothetical protein	Staphylococcus aureus subsp. aureus N315
		21284203	MW2474	ORFID:MW2474~hypothetical protein, similar to transcriptional regulator tetR-family	Staphylococcus aureus subsp. aureus Mu50
297	IRTTYKHK	AAP24934	AAP24934	Transcriptional regulator, TetR-related family	Bacillus anthracis str. Ames
		AAP07948	AAP07948	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30261071		transcriptional regulator, TetR-related family	Bacillus anthracis str. Ames
		30019116		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
298	MRKTYKHK	Q8EZX8	Q8EZX8	Transcriptional regulator, tetR family	Leptospira interrogans
		24216544		Transcriptional regulator, tetR family	Leptospira interrogans serovar lai str. 56601
299	IRSTYSHL	Q9KEY0	Q9KEY0	Transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
		15613282	BH0719	transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
300	VRSTYAHl	Q8G7U2	Q8G7U2	Probable TetR-like transcriptional regulator	Bifidobacterium longum
		23464769		probable TetR-like transcriptional	Bifidobacterium longum

301	IRSTYSHQ	Q8ETB5 23097801	Q8ETB5	regulator Transcriptional (TetR/AcrR family) transcriptional regulator	regulator	NCC2705 Oceanobacillus iheyensis
302	VRSTYKHQ	28379520		transcription regulator (putative)		Oceanobacillus iheyensis HTE831
303	MKRTYRHK	Q9KR96 Q8D4B9 15641750	Q9KR96 Q8D4B9	Transcriptional regulator, family Transcriptional regulator transcriptional regulator, family	TetR TetR	Vibrio cholerae Vibrio vulnificus Vibrio cholerae
304	MKQTYNHK	27367753 Q9A250 16127946	Q9A250	Transcriptional regulator Transcriptional regulator, family transcriptional regulator, family	TetR TetR	Vibrio vulnificus CMCP6 Caulobacter vibrioides Caulobacter crescentus CB15
305	MKRTYNHK	Q8PQS1 Q8PDV2 Q8D6D2 21229711 21241025 27367030 28870710	Q8PQS1 Q8PDV2 Q8D6D2	Transcriptional regulator tetR family Transcriptional regulator tetR family Transcriptional regulator transcriptional regulator tetR family transcriptional regulator tetR family		Xanthomonas axonopodis pv. citri Xanthomonas campestris pv. campestris Vibrio vulnificus Xanthomonas campestris pv. campestris str. ATCC 33913 Xanthomonas axonopodis pv. citri str. 306 Vibrio vulnificus CMCP6
306	VKRTYNHK	27380233		transcriptional regulator, family transcriptional regulatory protein	TetR	Pseudomonas syringae pv. tomato str. DC3000 Bradyrhizobium japonicum USDA 110
307	MKQTYKHK	Q98D70 13474047	Q98D70	Transcriptional regulator transcriptional regulator		Mesorhizobium loti Mesorhizobium loti
308	MKQTYSHK	Q8P877 Q8EC67 21231803 24374788	Q8P877 Q8EC67	Transcriptional regulator Transcriptional regulator, family transcriptional regulator tetR/acrR family transcriptional regulator, family	tetR/acrR TetR TetR	Xanthomonas campestris pv. campestris Shewanella oneidensis Xanthomonas campestris pv. campestris str. ATCC 33913 Shewanella oneidensis MR-1
309	VKRTYQHK	O53323 15610309 15842750	O53323	Putative regulatory protein (Tran- scriptional regulator, TetR family) hypothetical protein Rv3173c transcriptional regulator, family		Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv Mycobacterium tuberculosis CDC1551
310	IKRTYEHK	AAP27126 30263263	AAP27126	Transcriptional regulator, family transcriptional regulator, family	TetR TetR	Bacillus anthracis str. Ames Bacillus anthracis str. Ames
311	IKRTYSHK	28901417		putative transcriptional regulator		Vibrio parahaemolyticus RIMD 2210633
312	IRQTYNHK	Q98KH5 Q8YI41 Q8FZS0 13471487 17986887 23502274	Q98KH5 Q8YI41 Q8FZS0	Probable transcriptional regulator Transcriptional regulator, family Transcriptional regulator, family probable transcriptional regulator TRANSCRIPTIONAL REGULA- TOR, TETR FAMILY transcriptional regulator, family	TETR TETR TetR	Mesorhizobium loti Brucella melitensis Brucella melitensis biovar Suis Mesorhizobium loti Brucella melitensis 16M Brucella suis 1330
313	VRQTYNHK	Q8P228 Q8E3R3 Q8DY45 19745641 22537788 25011733	Q8P228 Q8E3R3 Q8DY45	Hypothetical protein spyM18_0584 Hypothetical protein Transcriptional regulator, family, putative conserved hypothetical protein transcriptional regulator, family, putative Unknown		Streptococcus pyogenes MGAS8232 Streptococcus agalactiae serogroup III Streptococcus agalactiae serogroup V Streptococcus pyogenes MGAS8232 Streptococcus agalactiae 2603V/R Streptococcus agalactiae

314	IRSTYNHK	Q97MJ5 15893494	Q97MJ5	Transcriptional regulators, family Transcriptional regulators, family	AcrR AcrR	NEM316 Clostridium acetobutylicum Clostridium acetobutylicum
315	VRRSYNHK	Q8E6F0 25010682	Q8E6F0	Hypothetical protein unknown		Streptococcus agalactiae serogroup III Streptococcus agalactiae NEM316
316	VRQTYYHK	Q9K7G7 AAO82741 AAP24955 AAP07967 15615956 30261092 29377517 30019135	Q9K7G7 AAO82741 AAP24955 AAP07967 BH3394	Hypothetical protein BH3394 Transcriptional regulator, family Hypothetical protein Transcriptional regulator, family BH3394-unknown conserved hypothetical protein transcriptional regulator, family Transcriptional regulator, family	TetR TetR TetR	Bacillus halodurans Enterococcus faecalis Bacillus anthracis str. Ames Bacillus cereus ATCC 14579 Bacillus halodurans Bacillus anthracis str. Ames Enterococcus faecalis V583 Bacillus cereus ATCC 14579
317	IRQTYYHK	Q92EH8 Q8Y9P4 16802523 16799557	Q92EH8 Q8Y9P4	Hypothetical protein lin0482 Hypothetical protein lmo0480 similar to putative transcriptional regulator similar to putative transcriptional regulator		Listeria innocua Listeria monocytogenes Listeria monocytogenes EGD-e Listeria innocua
318	VRATYYHK	Q8EL79 23100807	Q8EL79	Transcriptional regulator (tetR-family) transcriptional regulator		Oceanobacillus iheyensis Oceanobacillus iheyensis HTE831
319	VKRTYYHK	Q983X0 Q92NC8 13476730 15966033	Q983X0 Q92NC8	Transcription regulator Putative transcription regulator protein transcription regulator PUTATIVE TRANSCRIPTION REGULATORY PROTEIN transcriptional regulatory protein		Mesorhizobium loti Sinorhizobium meliloti Mesorhizobium loti Sinorhizobium meliloti
320	VKRMYYHK	27377946				Bradyrhizobium japonicum USDA 110
321	VKQTYHHK	Q8FAH4 AAN45699 24115482 26251111	Q8FAH4 AAN45699	Putative conserved protein Orf, partial conserved hypothetical protein orf, partial conserved hypothetical protein Putative conserved protein		Escherichia coli O6 Shigella flexneri Shigella flexneri 2a str. 301
322	IKKTYHHK	Q8EH37 24372971	Q8EH37	Transcriptional regulator, family transcriptional regulator, family	TetR TetR	Escherichia coli CFT073 Shewanella oneidensis Shewanella oneidensis MR-1
323	MKLTYSHK	Q9HXW2 Q8PH01 Q8P5K6 15598874 26988247 21232763 21244186	Q9HXW2 Q8PH01 Q8P5K6	Probable transcriptional regulator Transcriptional regulator Transcriptional regulator probable transcriptional regulator transcriptional regulator, tetR family transcriptional regulator transcriptional regulator		Pseudomonas aeruginosa Xanthomonas axonopodis pv. citri Xanthomonas campestris pv. campestris Pseudomonas aeruginosa PAO1 Pseudomonas putida KT2440 Xanthomonas campestris pv. campestris str. ATCC 33913 Xanthomonas axonopodis pv. citri str. 306
324	MKATYSHK	Q8YYN6 17228305	Q8YYN6	Transcriptional regulator transcriptional regulator		Nostoc sp. PCC 7120 Nostoc sp. PCC 7120
325	MKATYAHK	Q8PIQ4 21243568	Q8PIQ4	Transcriptional regulator transcriptional regulator		Xanthomonas axonopodis pv. citri Xanthomonas axonopodis pv. citri str. 306
326	MKSSYAHK	Q8ELQ0 23100624	Q8ELQ0	Hypothetical protein hypothetical protein		Oceanobacillus iheyensis Oceanobacillus iheyensis HTE831
327	MESSYNHK	27377365		transcriptional regulatory protein		Bradyrhizobium japonicum USDA 110
328	LKQSYSHK	Q929T1 Q8Y5H4 16804127	Q929T1 Q8Y5H4	Hypothetical protein lin2193 Hypothetical protein lmo2088 similar to transcription regulators		Listeria innocua Listeria monocytogenes Listeria monocytogenes

		16801258		similar to transcription regulators	EGD-e
329	LKLSYRHK	Q8UEX9	Q8UEX9	Transcriptional regulator, TetR family	<i>Listeria innocua</i>
		15888937		AGR_C_2996p	<i>Agrobacterium tumefaciens</i> str. C58
		17935520		transcriptional regulator, TetR family	<i>Agrobacterium tumefaciens</i> str. C58 (Cereon)
330	TRSSYTHK	Q9CH58	Q9CH58	Transcriptional regulator (YjaD)	<i>Agrobacterium tumefaciens</i> str. C58 (U. Washington)
		Q97N38	Q97N38	Transcriptional regulator, TetR family	<i>Lactococcus lactis</i> subsp. <i>lactis</i>
		Q8NZ06	Q8NZ06	Putative transcriptional regulator	<i>Streptococcus pyogenes</i> MGAS8232
		Q99XJ5	Q99XJ5	Putative transcriptional regulator (TetR/AcrR family)	<i>Streptococcus pyogenes</i> MGAS315
		Q8E2M8	Q8E2M8	Hypothetical protein	<i>Streptococcus agalactiae</i> serogroup III
		Q8DWS8	Q8DWS8	Transcriptional regulator, TetR family, putative	<i>Streptococcus agalactiae</i> serogroup V
		Q8DMW6	Q8DMW6	Conserved hypothetical protein	<i>Streptococcus pneumoniae</i> R6
		15672863	yjaD	transcription regulator	<i>Lactococcus lactis</i> subsp. <i>lactis</i>
		15675913		putative transcriptional regulator (TetR/AcrR family)	<i>Streptococcus pyogenes</i> M1 GAS
		15902037		transcriptional regulator, TetR family	<i>Streptococcus pneumoniae</i> TIGR4
		15904081		Conserved hypothetical protein	<i>Streptococcus pneumoniae</i> R6
		19747013		putative transcriptional regulator	<i>Streptococcus pyogenes</i> MGAS8232
		21911368		putative transcriptional regulator (TetR/AcrR family)	<i>Streptococcus pyogenes</i> MGAS315
		22538269		transcriptional regulator, TetR family, putative	<i>Streptococcus agalactiae</i> 2603V/R
		25012129		unknown	<i>Streptococcus agalactiae</i> NEM316
		28896740		putative transcriptional regulator (TetR/AcrR family)	<i>Streptococcus pyogenes</i> SSI-1
331	TRSGYTHK	Q8DRT2	Q8DRT2	Putative transcriptional regulator	<i>Streptococcus mutans</i>
		24380464		putative transcriptional regulator	<i>Streptococcus mutans</i> UA159
332	IRRSYTHK	CAD85857	CAD85857	Bacterial regulatory proteins, TetR family	<i>Nitrosomonas europaea</i>
		30249894		Bacterial regulatory proteins, TetR family	<i>Nitrosomonas europaea</i> ATCC 19718
333	VAMTYRYK	Q93HA7	Q93HA7	TetR-family regulatory protein	<i>Streptomyces avermitilis</i>
		BAC70092	BAC70092	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29828923		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
334	VAMTYRHK	Q9FBX5	Q9FBX5	Putative TetR-family transcriptional regulator	<i>Streptomyces coelicolor</i>
335	IKMTYKHK	Q9HZH4	Q9HZH4	Probable transcriptional regulator	<i>Pseudomonas aeruginosa</i>
		15598230		probable transcriptional regulator	<i>Pseudomonas aeruginosa</i> PAO1
336	VKMTYRHK	Q986H0	Q986H0	Transcriptional regulator	<i>Mesorhizobium loti</i>
		Q984Q2	Q984Q2	Transcriptional regulator	<i>Mesorhizobium loti</i>
		13476127		transcriptional regulator	<i>Mesorhizobium loti</i>
		13476546		transcriptional regulator	<i>Mesorhizobium loti</i>
337	VKMAYRHK	Q9EWI2	Q9EWI2	Putative tetR-family transcriptional regulator	<i>Streptomyces coelicolor</i>
338	ITMTYNHK	Q8ETH2	Q8ETH2	Transcriptional regulator	<i>Oceanobacillus iheyensis</i>
		23097744		transcriptional regulator	<i>Oceanobacillus iheyensis</i> HTE831
339	VKMSYRQK	Q8XTK0	Q8XTK0	Putative transcription regulator protein	<i>Ralstonia solanacearum</i>
		17548329	RSp0108	PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	<i>Ralstonia solanacearum</i>
340	IKANYVHK	Q8XQQ7	Q8XQQ7	Putative transcription regulator protein	<i>Ralstonia solanacearum</i>
		17549385	RSp1164	PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	<i>Ralstonia solanacearum</i>
341	MRANYLHK	Q8NL93	Q8NL93	Transcriptional regulator	<i>Corynebacterium glu-</i>

		Q8FSY2 19554245	Q8FSY2	Putative transcription regulator transcriptional regulator	tamicum Corynebacterium efficiens Corynebacterium glu- tamicum ATCC 13032
		25029466		putative transcription regulator	Corynebacterium efficiens YS-314
342	VKGTYYYHI	Q98IH8 13472185	Q98IH8	Probable transcriptional regulator probable transcriptional regulator	Mesorhizobium loti Mesorhizobium loti
343	IRMTYYHI	Q8DVH5 24378997	Q8DVH5	Putative transcriptional regulator putative transcriptional regulator	Streptococcus mutans Streptococcus mutans UA159
344	VRQTYYYHL	Q8DT13 24379967	Q8DT13	Putative transcriptional regulator putative transcriptional regulator	Streptococcus mutans Streptococcus mutans UA159
345	IRQTYYYHI	28211491		transcriptional regulator	Clostridium tetani E88
346	VRNTYYIYM	Q8XNF6 18309359	Q8XNF6	Probable transcriptional regulator probable transcriptional regulator	Clostridium perfringens Clostridium perfringens str. 13
347	VRNSYYHL	Q8G7D2 23464937	Q8G7D2	Hypothetical protein possibly in TetR transcriptional regulator family hypothetical protein possibly in TetR transcriptional regulator family	Bifidobacterium longum Bifidobacterium longum NCC2705
348	VRGTYAHV	Q97TK1 15004803 28377217	Q97TK1	HTH transcriptional regulator TetR/AcrR family HTH transcriptional regulator TetR/AcrR family transcription regulator (putative)	Clostridium acetobutylicum Clostridium acetobutylicum Lactobacillus plantarum WCFS1
349	VRGTYLHI	Q93D90 Q8CUW4 23098448 24380276	Q93D90 Q8CUW4 psaR	PsaR (Putative transcriptional regu- lator) Transcriptional regulator (TetR/AcrR family) transcriptional regulator putative transcriptional regulator	Streptococcus mutans Oceanobacillus iheyensis Oceanobacillus iheyensis HTE831 Streptococcus mutans UA159
350	VRGSYSHI	28379402		transcription regulator (putative)	Lactobacillus plantarum WCFS1
351	VRGTYTHA	Q8XME4 18309727	Q8XME4	Probable transcriptional regulator probable transcriptional regulator	Clostridium perfringens Clostridium perfringens str. 13
352	VRATYVHL	BAC75168 29833999	BAC75168	Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces avermitilis Streptomyces avermitilis MA-4680
353	VRKTYLHI	28377674		transcription regulator (putative)	Lactobacillus plantarum WCFS1
354	IRATYRHI	Q92DF2 Q8Y8M8 16802909	Q92DF2 Q8Y8M8	Hypothetical protein lin0861 Hypothetical protein lmo0868 lmo0868	Listeria innocua Listeria monocytogenes Listeria monocytogenes EGD-e
355	VRATYRHI	16799935 28379758		lin0861 unknown	Listeria innocua Lactobacillus plantarum WCFS1
356	HRATYRHL	BAC68261 29827092	BAC68261	Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces avermitilis Streptomyces avermitilis MA-4680
357	ARGSYWHL	27375712		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
358	MRQTYKER	Q9AE61 Q93DM9	Q9AE61 Q93DM9	Putative TetR-regulatory alkB1 Putative regulator	Rhodococcus erythropolis Rhodococcus sp. Q15
359	LRQTYNER	Q9AE65 Q93DM6 O05892 Q8VJ43 15610385	Q9AE65 Q93DM6 O05892 Q8VJ43	ORF1, alkB2 (Fragment) Putative regulator Hypothetical protein Rv3249c Hypothetical protein MT3347 hypothetical protein Rv3249c	Rhodococcus erythropolis Rhodococcus sp. Q15 Mycobacterium tuberculosis Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv
		15842838		hypothetical protein	Mycobacterium tuberculosis CDC1551
360	MRQTYNEK	Q93SY4 Q9K4B0 BAC70760 29829591	Q93SY4 Q9K4B0 BAC70760	Putative TetR-regulatory protein Putative tetR-family transcriptional regulator Putative TetR-family transcriptional regulator putative TetR-family transcriptional	Prauserella rugosa Streptomyces coelicolor Streptomyces avermitilis Streptomyces avermitilis

361	LRQTYNRK	Q9X7X3	Q9X7X3	regulator	MA-4680
362	LRLTYNQR	27382169		Putative transcriptional regulator	Streptomyces coelicolor
				transcriptional regulatory protein	Bradyrhizobium japonicum
363	VRQTYNYQ	Q9CIV9	Q9CIV9	Hypothetical protein yceG	USDA 110
		30023983	yceG	hypothetical protein	Lactococcus lactis subsp. lactis
364	VRQTYDYK	28377106		transcription regulator (putative)	Lactococcus lactis subsp. lactis
365	IRQTYNVK	Q8EVP6	Q8EVP6	Transcriptional regulator AcrR family	Lactobacillus plantarum WCFS1
		26553966		transcriptional regulator AcrR family	Mycoplasma penetrans
366	MKQSYNAK	28379778		transcription regulator (putative)	Mycoplasma penetrans
367	LRSTYTTR	Q939Q2	Q939Q2	JadR*	Lactobacillus plantarum WCFS1
368	TRSTYDYK	Q99RS1	Q99RS1	Hypothetical protein SAV2354	Streptomyces venezuelae
					Staphylococcus aureus subsp. aureus Mu50
		Q8NV49	Q8NV49	MW2275 protein	Staphylococcus aureus subsp. aureus N315
		15927934		hypothetical protein, similar to transcriptional regulator (TetR/AcrR family)	Staphylococcus aureus subsp. aureus MW2
		15925344		hypothetical protein	Staphylococcus aureus subsp. aureus Mu50
		21284004	MW2275	ORFID:MW2275~hypothetical protein, similar to transcriptional regulator (TetR/AcrR family)	Staphylococcus aureus subsp. aureus MW2
369	IRSTYDSK	AAP25000	AAP25000	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30261137		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
370	VKRTYDRK	P96812	P96812	Hypothetical 21.9 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15607277		hypothetical protein Rv0135c	Mycobacterium tuberculosis H37Rv
		15839516		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
371	VKKTYDRK	O86514	O86514	Putative transcriptional regulator	Streptomyces coelicolor
372	VVTTFAHK	Q8KUE8	Q8KUE8	Transcriptional repressor	Actinosynnema pretiosum subsp. auranticum
373	VATTFAHK	BAC68590	BAC68590	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827421		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
374	VTTTFAHK	Q9RK21	Q9RK21	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
375	IVTTFSHK	Q93J28	Q93J28	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
376	VTTTFKHK	BAC69974	BAC69974	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29828805		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
377	TQPAYWHK	P03038	TER1_ECOLI	Tetracycline repressor protein class A from transposon 1721	Escherichia coli
		P03039	TER3_ECOLI	Tetracycline repressor protein class C	Escherichia coli
		P51560	TER7_VIBAN	Tetracycline repressor protein class G	Listonella anguillarum
		Q8VVJ2	Q8VVJ2	TetR protein	Corynebacterium glutamicum
		Q9Z480	Q9Z480	Transcriptional repressor TetR	Agrobacterium tumefaciens str. C58

	Q939B1	Q939B1	Tetracycline resistance regulatory protein TetR	Aeromonas salmonicida
	Q939B5	Q939B5	Tetracycline resistance regulatory protein TetR	Aeromonas salmonicida subsp. salmonicida
	Q9X683	Q9X683	Tetracycline resistance repressor	Pseudomonas sp.
	Q8L2K1	Q8L2K1	TetR class A variant	Shigella sonnei
	Q9ZH63	Q9ZH63	Tetracycline repressor protein (Tetracycline resistance regulator protein)	Salmonella typhimurium DT104 Pseudomonas sp. Salmonella enterica subsp. enterica serovar Typhimurium
	Q8GN25	Q8GN25	Truncated tetracycline resistance repressor protein	Salmonella enteritidis
	Q8U893	Q8U893	Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
	CAC41337 CAC42502 CAD11598 AAM21660	CAC41337 CAC42502 CAD11598 AAM21660	Repressor protein Repressor protein Repressor Tetracycline resistance repressor protein	Escherichia coli Escherichia coli Escherichia coli Shigella sonnei
	AAN06706 AAO38185 CAD57191 AAG43219 15890777	AAN06706 AAO38185 CAD57191 AAG43219	Tetracycline resistance repressor TetR Tetracycline resistance repressor Tetracycline repressor protein AGR_L_1301p	Shigella sonnei Acinetobacter baumannii Aeromonas salmonicida IncN plasmid R46 Agrobacterium tumefaciens str. C58 (Cereon)
	17937894	tetR	transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
378	TQPSYWHK	Q939W4	Q939W4 Tetracycline repressor protein TetR(31)	Aeromonas salmonicida
379	TQPTYWHK	P04483	TER2_ECOLI	Escherichia coli
	P09164	TER4_ECOLI	TER4_ECOLI	Escherichia coli Salmonella enterica subsp. enterica serovar Ordonez
	P21337	TER5_ECOLI	TER5_ECOLI	Escherichia coli
	P51561	TER8_PASMU	TER8_PASMU	Pasteurella multocida
	P51562	TER8_PASPI	TER8_PASPI	Photobacterium damsela

			H	subsp. piscicida
	Q9Z358	Q9Z358	Tetracycline resistance repressor protein (TET repressor protein)	Mannheimia haemolytica
				Pasteurella multocida
				Pasteurella aerogenes
	Q93F26	Q93F26	Tet repressor	Shigella flexneri 2a
	Q934L0	Q934L0	Tetracycline repressor protein TetR(E)	Vibrio salmonicida
	Q9ZI84	Q9ZI84	Repressor protein	Proteus mirabilis
	Q8L0M9	Q8L0M9	Putative tetracycline repressor protein	Neisseria meningitidis
	Q9S455	Q9S455	TetR (Tetracycline repressor protein)	Salmonella typhi
				Shigella flexneri
	CAC44637 BAA78834 BAB91577 18466622	CAC44637 BAA78834 BAB91577 tetR	Tet repressor protein Repressor of tet operon TetR protein tetracycline repressor protein	Pasteurella aerogenes Plasmid R100 Salmonella typhimurium Salmonella enterica subsp. enterica serovar Typhi
380	TQPGYWHK	Q9X545	Q9X545	Tetracycline repressor protein TetR Corynebacterium glutamicum
381	TQPSYWYK	Q9S1N8	Q9S1N8	Putative tetR-family transcriptional regulator Streptomyces coelicolor
382	VNPTYWHK	Q98IT4	Q98IT4	Probable tetracycline repressor protein Mesorhizobium loti
		13472079	probable tetracycline repressor protein	Mesorhizobium loti
383	TQPAYHHK	Q928H4 Q8Y4H7 16804502	Q928H4 Q8Y4H7	Hypothetical protein lin2558 Hypothetical protein lmo2464 similar to transcription regulator Listeria innocua Listeria monocytogenes EGD-e
		16801620 28377824	similar to transcription regulator transcription regulator (putative)	Listeria innocua Lactobacillus plantarum WCFS1
384	TQPNYHHK	AAO80605	AAO80605	Transcriptional regulator, TetR family Enterococcus faecalis
		29375381	transcriptional regulator, TetR family	Enterococcus faecalis V583
385	TQPAYYHK	Q8ENL7	Q8ENL7	Transcriptional regulator (TetR/AcrR family) Oceanobacillus iheyensis
		23099919	transcriptional regulator	Oceanobacillus iheyensis HTE831
386	LAPAYWHK	Q9EWD6	Q9EWD6	Putative tetR-family transcriptional regulator Streptomyces coelicolor
		BAC71091	BAC71091	Putative TetR-family transcriptional regulator Streptomyces avermitilis
		29829922	putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
387	MAPAYWYK	AAO80602	AAO80602	Transcriptional regulator, TetR Enterococcus faecalis

		29375378		family transcriptional regulator, TetR	Enterococcus faecalis V583
388	IQPAYKHK	Q9CJ57	Q9CJ57	Transcriptional regulator	Lactococcus lactis subsp. lactis
		15672131	ybeD	transcription regulator	Lactococcus lactis subsp. lactis
389	TQPAYKHK	O53310	O53310	Hypothetical 23.1 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15610296		hypothetical protein Rv3160c	Mycobacterium tuberculosis H37Rv
		15842737		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
390	IQPAYRHR	27378018		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
391	VQPVYAHR	Q8U899	Q8U899	Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
		15890783		AGR_L_1315p	Agrobacterium tumefaciens str. C58 (Cereon)
		17937888		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
392	TQPVYSHK	Q9Z551	Q9Z551	Putative transcriptional regulator	Streptomyces coelicolor
		Q9RDS4	Q9RDS4	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
		BAC67914	BAC67914	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		BAC73436	BAC73436	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29826745		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		29832267		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
393	TQPAYAYK	Q9A3T4	Q9A3T4	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16127348		transcriptional regulator, TetR family	Caulobacter crescentus CB15
394	MKPVYQHK	Q9KIL9	Q9KIL9	F58r (Fragment)	Streptomyces coelicolor A3(2)
		Q9CCH0	Q9CCH0	Putative TetR-family transcriptional regulator	Mycobacterium leprae
		O05858	O05858	Hypothetical 25.2 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		Q9FBI8	Q9FBI8	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
		BAC70805	BAC70805	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		15610344		hypothetical protein Rv3208	Mycobacterium tuberculosis H37Rv
		15827358		putative TetR-family transcriptional regulator	Mycobacterium leprae
		15842794		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
		29829636		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
395	MKPIYRHK	Q93H95	Q93H95	TetR-family transcriptional regulator protein	Streptomyces avermitilis
		Q9FCA3	Q9FCA3	Putative transcriptional regulatory protein	Streptomyces coelicolor
		BAC74838	BAC74838	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29833669		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
396	IKPVYEHK	P96381	P96381	Hypothetical 21.7 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15608159		hypothetical protein Rv1019	Mycobacterium tuberculosis H37Rv
		15840447		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
397	VKPVYEHK	Q8NRT0	Q8NRT0	Transcriptional regulator	Corynebacterium glutamicum
		Q9K3T5	Q9K3T5	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
		Q8FQT6	Q8FQT6	Putative transcription regulator	Corynebacterium efficiens

		BAC71280	BAC71280	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29830111		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		19552187		transcriptional regulator	Corynebacterium glutamicum ATCC 13032
		25027589		putative transcription regulator	Corynebacterium efficiens YS-314
398	VKPIYEHK	Q8G735	Q8G735	Probable TetR-type transcriptional regulator	Bifidobacterium longum
		23465035		probable TetR-type transcriptional regulator	Bifidobacterium longum NCC2705
399	IAPVYQHK	BAC68158	BAC68158	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29826989		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
400	LKPVDHR	Q9I4B2	Q9I4B2	Probable transcriptional regulator	Pseudomonas aeruginosa
		Q930S6	Q930S6	Putative transcriptional regulator	Sinorhizobium meliloti
		15596423		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		16262572		putative transcriptional regulator	Sinorhizobium meliloti
401	LKPTYDHR	Q8PKE7	Q8PKE7	Transcriptional regulator tetR family	Xanthomonas axonopodis pv. citri
		21242963		transcriptional regulator tetR family	Xanthomonas axonopodis pv. citri str. 306
402	IRPVYDHK	Q9RYK4	Q9RYK4	Transcriptional regulator, TETR family	Deinococcus radiodurans
		15807968	DRA0308	transcriptional regulator, TetR family	Deinococcus radiodurans
403	VKPKYRHK	Q9CBQ1	Q9CBQ1	Putative TetR-family transcriptional regulator	Mycobacterium leprae
		O53295	O53295	Hypothetical 27.6 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15610187		hypothetical protein Rv3050c	Mycobacterium tuberculosis H37Rv
		15827931		putative TetR-family transcriptional regulator	Mycobacterium leprae
		15842616		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
404	VKPTYRHA	Q92XY9	Q92XY9	Putative transcriptional regulator	Sinorhizobium meliloti
		16263552		Putative transcriptional regulator	Sinorhizobium meliloti
405	LTPSYAHK	Q92DG8	Q92DG8	Hypothetical protein lin0845	Listeria innocua
		Q8Y8P4	Q8Y8P4	Hypothetical protein lmo0852	Listeria monocytogenes
		16802893		similar to transcription regulator	Listeria monocytogenes EGD-e
		16799919		similar to transcription regulator	Listeria innocua
				TetR/AcrR family	
406	LKPSYAHK	AAP28148	AAP28148	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP11127	AAP11127	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30264285		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30022295		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
407	IKPTYAHK	27382134		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
408	MAPSYKHR	Q8NQX5	Q8NQX5	Transcriptional regulator	Corynebacterium glutamicum
409	MAPSYKHK	Q9FBW1	Q9FBW1	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
410	MTPSYKHQ	Q9RL58	Q9RL58	Putative tetR family transcriptional regulator	Streptomyces coelicolor
411	VPPAYKHN	Q8ZSI4	Q8ZSI4	Transcriptional regulator	Nostoc sp. PCC 7120
		17158662	ahr7526	transcriptional regulator	Nostoc sp. PCC 7120
412	TPPAYRYK	Q8PEV3	Q8PEV3	Transcriptional regulatory protein	Xanthomonas axonopodis pv. citri
		21244954		transcriptional regulatory protein	Xanthomonas axonopodis pv. citri str. 306
413	TPPAYRLK	BAC68592	BAC68592	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827423		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
414	LGPAYRHK	O06169	O06169	Hypothetical 23.2 kDa protein (Transcriptional regulator, TetR	Mycobacterium tuberculosis

		15609643		family) hypothetical protein Rv2506	Mycobacterium tuberculosis H37Rv
		15842034		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
415	VGPGYRHK	Q9L078	Q9L078	Putative tetR-family regulatory protein	Streptomyces coelicolor
		BAC72991	BAC72991	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29831822		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
416	LGPAYRYR	Q9KXS8	Q9KXS8	Putative regulatory protein	Streptomyces coelicolor
		AAP21651	AAP21651	Shy4	Streptomyces hygroscopicus subsp. yingchengensis
		BAC69176	BAC69176	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29828007		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
417	TEAAYRHK	P06969	TTK_ECOLI	Ttk protein	Escherichia coli
		Q93V18	Q93V18	Hypothetical 24.2 kDa protein	Shigella flexneri
		Q9KVD2	Q9KVD2	Transcriptional regulator, TetR family	Salmonella typhimurium
		Q8ZL46	Q8ZL46	Putative transcriptional regulator (TetR/ArcR family)	Vibrio cholerae
		Q8Z2H4	Q8Z2H4	Putative TetR-family transcriptional regulator	Salmonella typhi
		Q8X580	Q8X580	Putative transcriptional regulator	Escherichia coli O157:H7
		Q8Y3E9	Q8Y3E9	Probable transcription regulator protein	Ralstonia solanacearum
		Q8ZJP6	Q8ZJP6	TetR-family regulatory protein (Putative transcriptional regulator)	Yersinia pestis
		Q8FC86	Q8FC86	Ttk protein	Escherichia coli O6
		Q8E9L9	Q8E9L9	Transcriptional regulator, TetR family	Shewanella oneidensis
		Q8DDX7	Q8DDX7	Transcriptional regulator, TetR family	Vibrio vulnificus
		AAP19065	AAP19065	Putative transcriptional regulator	Shigella flexneri 2a str. 2457T
		AAO71268	AAO71268	Putative TetR-family transcriptional regulator	Salmonella typhi
		16131512	ttk	putative transcriptional regulator	Escherichia coli K12
		15640244		transcriptional regulator, TetR family	Vibrio cholerae
		15804182	ttk	putative transcriptional regulator	Escherichia coli O157:H7 EDL933
		15833770		putative transcriptional regulator	Escherichia coli O157:H7
		16120399	ttk	TetR-family regulatory protein	Yersinia pestis CO92
		16767017	ttk	putative transcriptional regulator (TetR/ArcR family)	Salmonella typhimurium LT2
		16762577	ttk	putative TetR-family transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi
		17544750	RSc0031	PROBABLE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum
		22124015	ttk	putative transcriptional regulator	Yersinia pestis KIM
		24114910	ttk	putative transcriptional regulator	Shigella flexneri 2a str. 301
		24375735		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
		26250287	ttk	Ttk protein	Escherichia coli CFT073
		27364277		Transcriptional regulator, TetR family	Vibrio vulnificus CMCP6
		28896954		transcriptional regulator, TetR family	Vibrio parahaemolyticus RIMD 2210633
		29144066	ttk	putative TetR-family transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi Ty2
		30065083	ttk	putative transcriptional regulator	Shigella flexneri 2a str. 2457T
418	TEGAYRHK	Q9A771	Q9A771	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16126098		transcriptional regulator, TetR family	Caulobacter crescentus CB15
419	TEAAYRYK	P29280	TTK_HAEIN	Ttk protein homolog	Haemophilus influenzae
		Q9CLQ9	Q9CLQ9	Hypothetical protein PM1155	Pasteurella multocida
		16272893	ttk	transcriptional regulator	Haemophilus influenzae Rd
		15603020		unknown	Pasteurella multocida
420	THSAYRHK	27381351		transcriptional regulatory protein	Bradyrhizobium japonicum

421	IEGAYRHK	O67927 15607114	O67927 acrR2	Transcriptional regulator (TETR/ACRR family) transcriptional regulator (TetR/AcrR family)	USDA 110 Aquifex aeolicus Aquifex aeolicus VF5
422	TEGLFKHK	Q8F6M1 24213982	Q8F6M1	Transcriptional regulator, tetR family Transcriptional regulator, tetR family	Leptospira interrogans Leptospira interrogans serovar lai str. 56601
423	TEALFKHK	27377970		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
424	TEALFRYK	Q8EX89 24217386	Q8EX89	Transcriptional regulator, TetR family transcriptional regulator, TetR family	Leptospira interrogans Leptospira interrogans serovar lai str. 56601
425	ISTAYMHK	Q9A905 16125441	Q9A905	Transcriptional regulator, TetR family transcriptional regulator, TetR family	Caulobacter vibrioides Caulobacter crescentus CB15
426	VSSSYRHK	O87854	O87854	Putative transcriptional regulator	Streptomyces coelicolor
427	LARAYRHK	P95251 15609100 15841435	P95251	Hypothetical 44.3 kDa protein (Transcriptional regulator, TetR family) hypothetical protein Rv1963c transcriptional regulator, TetR family	Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv Mycobacterium tuberculosis CDC1551
428	MAAAYRHK	Q9X7X6	Q9X7X6	Putative regulatory protein	Streptomyces coelicolor
429	FAAAYRHR	27382272		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
430	LHQAYRHR	Q9A5V8 16126578	Q9A5V8	Transcriptional regulator, TetR family transcriptional regulator, TetR family	Caulobacter vibrioides Caulobacter crescentus CB15
431	LHTAYKHK	Q97FG6 15896028	Q97FG6	Transcriptional regulators, AcrR family Transcriptional regulators, AcrR family	Clostridium acetobutylicum Clostridium acetobutylicum
432	LHAAYKHK	Q926U3 Q8Y3M1 16804851	Q926U3 Q8Y3M1	Hypothetical protein lin2947 Hypothetical protein lmo2814 similar to transcriptional regulator	Listeria innocua Listeria monocytogenes Listeria monocytogenes EGD-e
433	LHAASRHK	16802006 BAC72796 29831627	BAC72796	similar to transcriptional regulator Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Listeria innocua Streptomyces avermitilis Streptomyces avermitilis MA-4680
434	FAPSYNHM	Q9CJ74 15672109	Q9CJ74 ybdA	Transcription regulator transcription regulator	Lactococcus lactis subsp. lactis Lactococcus lactis subsp. lactis
435	FSPSYNHM	Q97TR9 15004732	Q97TR9	HTH transcriptional regulator TetR family HTH transcriptional regulator TetR family	Clostridium acetobutylicum Clostridium acetobutylicum
436	FYPSYNHI	AAO81799 29376575	AAO81799	Transcriptional regulator, TetR family transcriptional regulator, TetR family	Enterococcus faecalis Enterococcus faecalis V583
437	LPPSYNHL	O34892 16078967	O34892 yobS	Transcription regulator yobS	Bacillus subtilis Bacillus subtilis subsp. subtilis str. 168
438	LSPSYNHL	AAP28909 AAP11872 30265046 30023040	AAP28909 AAP11872	Transcriptional regulator, putative Transcriptional regulator, TetR family transcriptional regulator, putative Transcriptional regulator, TetR family	Bacillus anthracis str. Ames Bacillus cereus ATCC 14579 Bacillus anthracis str. Ames Bacillus cereus ATCC 14579
439	IGPSYNHL	O53789 15607821 15840084	O53789	Putative transcriptional regulator (Transcriptional regulator, putative) hypothetical protein Rv0681 transcriptional regulator, putative	Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv Mycobacterium tuberculosis CDC1551
440	TGPSYNHK	Q9S250	Q9S250	Putative transcriptional regulator	Streptomyces coelicolor

		BAC74310	BAC74310	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29833141		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		21220197	SCO1702	putative transcriptional regulator	Streptomyces coelicolor A3(2)
441	MVTSYWYK	Q11023	YD53_MYCTU	Hypothetical transcriptional regulator Rv1353c	Mycobacterium tuberculosis H37Rv
		15608493		hypothetical protein Rv1353c	Mycobacterium tuberculosis CDC1551
		15840808		transcriptional regulator, TetR/AcrR family	Mycobacterium tuberculosis CDC1551
442	MVTSYWHR	BAC74151	BAC74151	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29832982		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
443	MVMSYWHR	Q92SE9	Q92SE9	Hypothetical protein R00451	Sinorhizobium meliloti
		15964204		HYPOTHETICAL PROTEIN	Sinorhizobium meliloti
444	MPMSYWYK	Q9AJL5	Q9AJL5	VarR	Streptomyces virginiae
		O52558	O52558	Repressor	Amycolatopsis mediterranei
445	MAMSYWYK	Q9F0Y2	Q9F0Y2	Pip	Streptomyces coelicolor
		Q9ADN3	Q9ADN3	Putative tetR family transcriptional regulatory protein	Streptomyces coelicolor
		BAC71905	BAC71905	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		BAC76550	BAC76550	Putative TetR family receptor protein	Streptomyces rochei
		29830736		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
446	IPSSYWHK	O07229	O07229	Putative regulatory protein MTCY63.07 (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis H37Rv
		15607443		hypothetical protein Rv0302	Mycobacterium tuberculosis CDC1551
		15839687		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
447	VPSAYWHK	Q9FCM0	Q9FCM0	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
448	ITSSYWHK	BAC68146	BAC68146	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29826977		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
449	VRSSYWYR	Q98CA4	Q98CA4	Hypothetical protein mll5233	Mesorhizobium loti
		13474363		unknown protein	Mesorhizobium loti
450	IRTSYWFR	Q92US8	Q92US8	Putative transcriptional regulator TetR family protein	Sinorhizobium meliloti
		Q8UCF0	Q8UCF0	Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
		15889802		AGR_C_4609p	Agrobacterium tumefaciens str. C58 (Cereon)
		16264759		putative transcriptional regulator TetR family protein	Sinorhizobium meliloti
		17936418		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
451	VRTSYWFR	26991435		transcriptional regulator, TetR family	Pseudomonas putida KT2440
452	VKGPYWRR	Q930U4	Q930U4	Hypothetical protein RA0100	Sinorhizobium meliloti
		16262553		hypothetical protein	Sinorhizobium meliloti
453	ITGISHYK	P17446	BETI_ECOLI	Regulatory protein betI	Escherichia coli
		Q8X6C3	Q8X6C3	Probably transcriptional repressor of bet genes	Escherichia coli O157:H7
		Q8FK17	Q8FK17	Regulatory protein betI	Escherichia coli O6
		16128298	betI	probably transcriptional repressor of bet genes	Escherichia coli K12
		15800035	betI	probably transcriptional repressor of	Escherichia coli O157:H7

		15829613		bet genes probably transcriptional repressor of bet genes	EDL933 Escherichia coli O157:H7
454	ITGINHYK	26246331 28900969	betI	Regulatory protein betI putative transcriptional regulator BetI	Escherichia coli CFT073 Vibrio parahaemolyticus RIMD 2210633
455	INGISHYK	Q9HTJ0 Q8ZGV8	Q9HTJ0 Q8ZGV8	Transcriptional regulator BetI TetR-family regulatory protein	Pseudomonas aeruginosa Yersinia pestis
		15600567	betI	transcriptional regulator BetI	Pseudomonas aeruginosa PAO1
		16121463 28867670	betI betI	TetR-family regulatory protein regulatory protein BetI	Yersinia pestis CO92 Pseudomonas syringae pv. tomato str. DC3000
456	VAGISHYK	Q9L4K2	Q9L4K2	BetI protein	Halomonas elongata
457	VSGINHYK	Q8D3K4 27368025	Q8D3K4	Transcriptional regulator Transcriptional regulator	Vibrio vulnificus Vibrio vulnificus CMCP6
458	LTGLTHYK	Q9F3N6	Q9F3N6	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
459	LAGLRHYK	Q98CC7 13474340	Q98CC7	Transcriptional regulator transcriptional regulator	Mesorhizobium loti Mesorhizobium loti
460	VAGLRHYK	Q92S61	Q92S61	Putative transcription regulator protein	Sinorhizobium meliloti
		15964318		PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	Sinorhizobium meliloti
461	ISGARHYR	Q9RVM4 15806026	Q9RVM4 DR1003	Transcriptional regulator transcriptional regulator	Deinococcus radiodurans Deinococcus radiodurans
462	VVGLNHHK	Q9HUI1 15600177	Q9HUI1	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
463	IVGLSHHK	26990438		transcriptional regulator, TetR family	Pseudomonas putida KT2440
464	MRGINFHK	Q98FP9 13473165	Q98FP9	Transcriptional regulator transcriptional regulator	Mesorhizobium loti Mesorhizobium loti
465	VRGMNYHK	BAC68141	BAC68141	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29826972		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
466	LRGLSYYK	Q9K3P4	Q9K3P4	Putative tetR-family transcriptional regulatory protein	Streptomyces coelicolor
		BAC69213	BAC69213	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29828044		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
467	LKGILHYK	Q92YD5 Q8UDR2	Q92YD5 Q8UDR2	Putative transcriptional regulator Transcriptional regulator, TetR family	Sinorhizobium meliloti Agrobacterium tumefaciens str. C58
		16263402 15889345		putative transcriptional regulator AGR_C_3729p	Sinorhizobium meliloti Agrobacterium tumefaciens str. C58 (Cereon)
		17935944	betI	transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
468	TKGLFHYK	Q92A46 Q8Y5U3 Q9X0C0	Q92A46 Q8Y5U3 Q9X0C0	Hypothetical protein lin2076 Hypothetical protein lmo1962 Transcriptional regulator, TETR family	Listeria innocua Listeria monocytogenes Thermotoga maritima
		15643788		transcriptional regulator, TetR family	Thermotoga maritima
		16804001		similar to transcription regulators (TetR family)	Listeria monocytogenes EGD-e
		16801142		similar to transcription regulators (TetR family)	Listeria innocua
		28377103		transcription regulator (putative)	Lactobacillus plantarum WCFS1
469	TKGLFYYK	Q8DUB5 24379466	Q8DUB5	Putative transcription regulator putative transcription regulator	Streptococcus mutans Streptococcus mutans UA159
470	TKRLFYYK	27382028		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
471	MPALHHYK	O69786 15964703	BETI_RHIME betI	Regulatory protein betI TRANSCRIPTION REGULATOR PROTEIN	Sinorhizobium meliloti Sinorhizobium meliloti
472	VPALHHYK	Q8YFY3 Q9X8G9	Q8YFY3 Q9X8G9	Regulatory protein BETI Putative transcriptional regulator	Brucella melitensis Streptomyces coelicolor
		Q8G1Z7	Q8G1Z7	Transcriptional regulator BetI	Brucella melitensis biovar

		17987662		REGULATORY PROTEIN BETI	Suis
		23501443	betI	transcriptional regulator BetI	Brucella melitensis 16M
473	MSALHHYK	Q985M2	Q985M2	Transcriptional regulator	Brucella suis 1330
		13476325		transcriptional regulator	Mesorhizobium loti
474	MAALHHYK	Q8UH57	Q8UH57	Transcriptional regulator, TetR family	Mesorhizobium loti
		15888168		AGR_C_1514p	Agrobacterium tumefaciens str. C58
		17934736		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (Cereon)
475	IAALHHYK	O33187	O33187	Hypothetical 22.3 kDa protein (Transcriptional regulator, TetR family)	Agrobacterium tumefaciens str. C58 (U. Washington)
		15608823		hypothetical protein Rv1685c	Mycobacterium tuberculosis H37Rv
		15841142		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
476	VSALHHYK	BAC72427	BAC72427	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29831258		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
477	MKPLYHYK	Q9FA56	Q9FA56	Putative regulator (Fragment)	Azoarcus evansii
478	IRPLYHYK	Q9KZ14	Q9KZ14	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
		BAC73697	BAC73697	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29832528		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
479	IRALYHYK	Q49662	Q49662	B1308_C2_181	Mycobacterium leprae
		Q9CCL4	Q9CCL4	Putative TetR-family transcriptional regulator	Mycobacterium leprae
		P96900	P96900	Hypothetical 24.8 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis H37Rv
		15610431		hypothetical protein Rv3295	Mycobacterium tuberculosis H37Rv
		15827303		putative TetR-family transcriptional regulator	Mycobacterium leprae
		15842886		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
480	VRSLYHYK	Q8G6X9	Q8G6X9	Possible TetR-type transcriptional regulator	Bifidobacterium longum
		23465091		possible TetR-type transcriptional regulator	Bifidobacterium longum NCC2705
481	IRGLYHYK	Q9RD77	Q9RD77	Putative TetR-family transcriptional regulatory protein	Streptomyces coelicolor
		BAC75138	BAC75138	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29833969		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
482	IQGLYHYK	Q8TTR2	Q8TTR2	Transcriptional regulator, TetR family	Methanosarcina acetivorans
		20089261		transcriptional regulator, TetR family	Methanosarcina acetivorans C2A
483	VKQLYHYK	27379320		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
484	MKSLQHYK	28379749		transcription regulator (putative)	Lactobacillus plantarum WCFS1
485	MKSRYHYK	Q8XS68	Q8XS68	Probable transcription regulator protein	Ralstonia solanacearum
		17548834	RSp0613	PROBABLE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum
486	LGSQYHYK	BAC76583	BAC76583	Putative TetR family receptor protein	Streptomyces rochei
487	EPSQYHYK	Q9RJ10	Q9RJ10	Putative transcriptional regulatory protein	Streptomyces coelicolor
488	VKRMYYK	Q9I6P5	Q9I6P5	Probable transcriptional regulator	Pseudomonas aeruginosa
		Q92XC6	Q92XC6	Hypothetical transcriptional regulator protein SMB20033	Sinorhizobium meliloti
		Q8XQ91	Q8XQ91	Putative transcription regulator	Ralstonia solanacearum
		BAC69489	BAC69489	transcription regulator protein	Streptomyces avermitilis
		15595440		Putative TetR-family transcriptional regulator	Streptomyces avermitilis
				probable transcriptional regulator	Pseudomonas aeruginosa

		26990239		transcriptional regulator, TetR family	PAO1 Pseudomonas putida KT2440
		16263784		hypothetical transcriptional regulator protein	Sinorhizobium meliloti
		29828320		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		17549614	RSp1395	PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum
		27381933		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
		28869545		transcriptional regulator, TetR family	Pseudomonas syringae pv. tomato str. DC3000
489	IKRMYYYK	Q8U7B9	Q8U7B9	Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
		15890456		AGR_L_676p	Agrobacterium tumefaciens str. C58 (Cereon)
		17938219		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
490	VERMYYYK	Q8Y0N6	Q8Y0N6	Probable transcription regulator protein	Ralstonia solanacearum
		17545726	RSc1007	PROBABLE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum
491	VKRMYYHK	Q92YT3	Q92YT3	Putative TetR/AcrR-family transcriptional regulator	Sinorhizobium meliloti
		Q92T62	Q92T62	Putative transcription regulator protein	Sinorhizobium meliloti
		16263233		putative TetR/AcrR-family transcriptional regulator	Sinorhizobium meliloti
		15963871		PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Sinorhizobium meliloti
492	IKRMYHYK	Q8UBA1	Q8UBA1	Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
		15891811		AGR_L_1690GMp	Agrobacterium tumefaciens str. C58 (Cereon)
		17936826		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
493	IDRMYYYK	26990654		transcriptional regulator, TetR family	Pseudomonas putida KT2440
494	LKRMHYHK	Q8NN32	Q8NN32	Transcriptional regulator	Corynebacterium glutamicum
		19553580		transcriptional regulator	Corynebacterium glutamicum ATCC 13032
495	LKRMHYHK	Q8FN60	Q8FN60	Putative transcription regulator	Corynebacterium efficiens
		25028843		putative transcription regulator	Corynebacterium efficiens YS-314
496	IIRMYHYK	Q9A556	Q9A556	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16126845		transcriptional regulator, TetR family	Caulobacter crescentus CB15
497	LDRMLYYK	Q9A4F8	Q9A4F8	Transcriptional regulator, TetR family	Caulobacter vibrioides
		BAC71472	BAC71472	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		16127108		transcriptional regulator, TetR family	Caulobacter crescentus CB15
		29830303		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
498	LQGMHYFK	Q93L29	Q93L29	Putative transcription factor	Burkholderia cepacia
499	LQSLHHFK	Q9RCV4	Q9RCV4	Putative TetR-family regulator	Streptomyces coelicolor
500	TEASFKRK	Q9Z601	Q9Z601	RegN	Nostoc punctiforme
501	TEATFKKK	Q927F2	Q927F2	Hypothetical protein lin2837	Listeria innocua
		Q8Y3Y9	Q8Y3Y9	Hypothetical protein lmo2690	Listeria monocytogenes
		16804727		similar to transcription regulator, TetR family	Listeria monocytogenes EGD-e
		16801897		similar to transcription regulator, TetR family	Listeria innocua
502	TEATFRNK	O34643	O34643	YVKB	Bacillus subtilis
		16080573	yvkb	similar to transcriptional regulator (TetR/AcrR family)	Bacillus subtilis subsp. subtilis str. 168
503	TEVTFRCK	Q8RCU4	Q8RCU4	Transcriptional regulator	Thermoanaerobacter tengcongensis
		20806833	AcrR2	Transcriptional regulator	Thermoanaerobacter teng-

504	TEVTFRKK	O27815	O27815	Conserved protein	congensis
		15679775		conserved protein	Methanothermobacter thermoautotrophicus str. Delta H
505	VERTFRKK	Q9CIW9	Q9CIW9	Transcriptional regulator	Methanothermobacter thermoautotrophicus str. Delta H
		15672219	ycdF	transcription regulator	Lactococcus lactis subsp. lactis
506	TEGTFRYK	P73186	P73186	Hypothetical protein sll1286	Lactococcus lactis subsp. lactis
		Q97E45	Q97E45	Transcriptional regulator, AcrR family	Synechocystis sp. PCC 6803
		Q8ZR43	Q8ZR43	Putative regulatory protein	Clostridium acetobutylicum
		Q8Z8M3	Q8Z8M3	Putative tetR-family transcriptional regulator	Salmonella typhimurium
		AAP26397	AAP26397	Transcriptional regulator, TetR family	Salmonella typhi
		AAO69889	AAO69889	Putative TetR-family transcriptional regulator	Bacillus anthracis str. Ames
		AAP09441	AAP09441	Transcriptional regulator, TetR family	Salmonella typhi
		16329804		unknown protein	Bacillus cereus ATCC 14579
		15896516		Transcriptional regulator, AcrR family	Synechocystis sp. PCC 6803
		16763957		putative regulatory protein	Clostridium acetobutylicum
		16759540		putative tetR-family transcriptional regulator	Salmonella typhimurium LT2
		30262534		transcriptional regulator, TetR family	Salmonella enterica subsp. enterica serovar Typhi
29142687		putative TetR-family transcriptional regulator	Bacillus anthracis str. Ames		
30020609		Transcriptional regulator, TetR family	Salmonella enterica subsp. enterica serovar Typhi Ty2		
507	TEGTFRHK	Q92D91	Q92D91	Hypothetical protein lin0926	Bacillus cereus ATCC 14579
		Q8YVN3	Q8YVN3	Transcriptional regulator	Listeria innocua
		Q8Y8H7	Q8Y8H7	Hypothetical protein lmo0926	Nostoc sp. PCC 7120
		16802966		similar to transcription regulator (TetR/AcrR family)	Listeria monocytogenes
16799997		similar to transcription regulator (TetR/AcrR family)	Listeria monocytogenes EGD-e		
508	TEGTFYK	17229433		transcriptional regulator	Listeria innocua
		Q9PA05	Q9PA05	Hypothetical protein Xf2715	Nostoc sp. PCC 7120
509	TEGTFKRK	15839304		conserved hypothetical protein	Xylella fastidiosa
		Q9RWD4	Q9RWD4	Transcriptional regulator, TETR family	Xylella fastidiosa 9a5c
		15805761	DR0735	transcriptional regulator, TetR family	Deinococcus radiodurans
510	TEGTYRRK	28379259		transcription regulator (putative)	Deinococcus radiodurans
511	TEKTFRHK	Q9KBG3	Q9KBG3	Transcriptional regulator (TetR/AcrR family)	Lactobacillus plantarum WCFS1
		15614528	BH1965	transcriptional regulator (TetR/AcrR family)	Bacillus halodurans
512	TTGTFHHK	Q8EB50	Q8EB50	Transcriptional regulator, TetR family	Bacillus halodurans
		24375182		transcriptional regulator, TetR family	Shewanella oneidensis
513	TEGTFHNK	28377199		transcription regulator (putative)	Shewanella oneidensis MR-1
514	TDGTFSHK	28900148		transcriptional regulator	Lactobacillus plantarum WCFS1
515	TTGTIYAYK	Q97N19	Q97N19	Transcriptional regulator, AcrR family	Vibrio parahaemolyticus RIMD 2210633
		15893317		Transcriptional regulator, AcrR family	Clostridium acetobutylicum
516	TVGTYAYK	Q8DWD0	Q8DWD0	Putative transcriptional regulator (TetR/AcrR family)	Clostridium acetobutylicum
		24378652		putative transcriptional regulator (TetR/AcrR family)	Streptococcus mutans
517	TTGTFNYK	Q8TTS1	Q8TTS1	Transcriptional regulator, TetR family	Streptococcus mutans UA159
		Q8TTQ8	Q8TTQ8	Transcriptional regulator, TetR family	Methanosarcina acetivorans
		Q8PWH3	Q8PWH3	Transcriptional regulator, TetR family	Methanosarcina acetivorans
				Transcriptional regulator, TetR family	Methanosarcina mazei

		20089252		transcriptional regulator, TetR	Methanosarcina acetivorans C2A
		20089265		transcriptional regulator, TetR	Methanosarcina acetivorans C2A
		21227716		transcriptional regulator, TetR	Methanosarcina mazei Goe1 family
518	TTGTFLYK	BAC68912	BAC68912	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827743		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
519	ARGTFNYK	Q9RUI4	Q9RUI4	Transcriptional repressor, TETR family	Deinococcus radiodurans
		15806419	DR1402	transcriptional repressor, TetR family	Deinococcus radiodurans
520	TPGTYQFK	BAC73045	BAC73045	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29831876		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
521	THGTFAHQ	Q97G98	Q97G98	Transcriptional regulator, TetR/AcrR family	Clostridium acetobutylicum
		15895736		Transcriptional regulator, TetR/AcrR family	Clostridium acetobutylicum
522	TQATYKYK	Q8E7A5	Q8E7A5	Hypothetical protein	Streptococcus agalactiae serogroup III
		Q8E1U3	Q8E1U3	Transcriptional regulator, TetR family	Streptococcus agalactiae serogroup V
		22536442		transcriptional regulator, TetR family	Streptococcus agalactiae 2603V/R
		25010324		unknown	Streptococcus agalactiae NEM316
523	TQATFKYK	Q8DW40	Q8DW40	Putative transcriptional regulator	Streptococcus mutans
		24378746		putative transcriptional regulator	Streptococcus mutans UA159
524	TQSNLYK	28379503		transcription regulator (putative)	Lactobacillus plantarum WCFS1
525	LKATHRFR	P32265	NFXB_PSEAE	Transcriptional regulatory protein nfxB	Pseudomonas aeruginosa
		AAO47352	AAO47352	Truncated NfxB	Pseudomonas aeruginosa
		15599796	nfxB	transcriptional regulator NfxB	Pseudomonas aeruginosa PAO1
		26989539	nfxB	transcriptional regulator NfxB	Pseudomonas putida KT2440
526	MRATHRHR	Q936C8	Q936C8	Repressor protein	Streptomyces clavuligerus
		Q9F147	Q9F147	PqrA	Streptomyces coelicolor A3(2)
		Q9L1B3	Q9L1B3	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
527	MRATHRQR	BAC74492	BAC74492	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29833323		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
528	LSATHRHR	Q9ADD9	Q9ADD9	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
529	MVATHRYR	Q9KB32	Q9KB32	Hypothetical protein BH2103	Bacillus halodurans
		15614666	BH2103	BH2103-unknown conserved protein in others	Bacillus halodurans
530	MIATHRYR	AAP08788	AAP08788	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30019956		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
531	MVATFRYK	Q8ETD3	Q8ETD3	Transcriptional regulator (TetR/AcrR family)	Oceanobacillus iheyensis
		23097783		transcriptional regulator	Oceanobacillus iheyensis HTE831
532	MRATYRYR	Q11063	YC55_MYCTU	Hypothetical transcriptional regulator Rv1255c	Mycobacterium tuberculosis
		15608395		hypothetical protein Rv1255c	Mycobacterium tuberculosis H37Rv
		15840701		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
533	VRATYRYQ	Q98GC7	Q98GC7	Transcriptional regulator	Mesorhizobium loti
		13472936		transcriptional regulator	Mesorhizobium loti
534	IRATYRYQ	Q8ZAG2	Q8ZAG2	Putative TetR-family regulatory protein (Transcriptional regulator)	Yersinia pestis
		16123975		putative TetR-family regulatory	Yersinia pestis CO92

535	MSATYRFQ	22124305 Q9X546	Q9X546	protein transcriptional regulator Tetracycline repressor protein ho- molog Trh	Yersinia pestis KIM Corynebacterium glu- tamicum
536	LIATYRRR	Q8KRU1	Q8KRU1	PauB	Pseudonocardia autotrophica
537	MVATYRRR	Q9WX19	Q9WX19	Putative transcriptional regulator	Streptomyces coelicolor
538	MVATYRNR	Q8NQ14	Q8NQ14	Transcriptional regulator	Corynebacterium glu- tamicum
		19552852		transcriptional regulator	Corynebacterium glu- tamicum ATCC 13032
539	LRATYRNR	Q8PP81	Q8PP81	Transcriptional regulator tetR/acrR family	Xanthomonas axonopodis pv. citri
		Q8P335	Q8P335	Transcriptional regulator tetR/acrR family	Xanthomonas campestris pv. campestris
		21230229		transcriptional regulator tetR/acrR family	Xanthomonas campestris pv. campestris str. ATCC 33913
		21241577		transcriptional regulator tetR/acrR family	Xanthomonas axonopodis pv. citri str. 306
540	LNATYRNR	Q9KYT9	Q9KYT9	Putative tetR-family transcriptional regulator (Fragment)	Streptomyces coelicolor
		BAC71411	BAC71411	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29830242		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
541	LNATYRHR	Q9F2L2	Q9F2L2	Putative tetR family transcriptional regulatory protein	Streptomyces coelicolor
542	LSGYRNR	Q9RJL5	Q9RJL5	Putative tetR family transcriptional regulator	Streptomyces coelicolor
543	LQGYRNR	Q9EWZ1	Q9EWZ1	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
		BAC69252	BAC69252	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29828083		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
544	LLATLRHR	BAC68379	BAC68379	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		BAC68385	BAC68385	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827210		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		29827216		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
545	VKATYRRK	Q10774	YF56_MYCTU	Hypothetical transcriptional regula- tor Rv1556	Mycobacterium tuberculosis
		15608694		hypothetical protein Rv1556	Mycobacterium tuberculosis H37Rv
		15841023		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
546	IKATYRRK	Q8CMN8	Q8CMN8	Transcriptional regulator AcrR family	Staphylococcus epidermidis
		27469328		transcriptional regulator AcrR family	Staphylococcus epidermidis ATCC 12228
547	IKATYKRK	AAP28418	AAP28418	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP11410	AAP11410	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30264555		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30022578		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
548	MKATYRRK	BAC73657	BAC73657	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29832488		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
549	IRATYRRK	O33179	O33179	Hypothetical 20.3 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15608914		hypothetical protein Rv1776c	Mycobacterium tuberculosis H37Rv
		15841245		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
550	MKSTYRQK	Q9XAN6	Q9XAN6	Putative transcriptional regulator	Streptomyces coelicolor
		BAC69422	BAC69422	Putative TetR-family transcriptional regulator	Streptomyces avermitilis

		29828253		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
551	MKATYRQK	Q9RKC0	Q9RKC0	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
		BAC71331	BAC71331	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29830162		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
552	MPATYRRK	Q9F3C3	Q9F3C3	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
553	FKAAYRRK	P72185	P72185	Repressor protein (HEMR protein)	Propionibacterium freudenreichii
554	MKAAYRRK	Q54288	Q54288	Regulator of antibiotic transport complexes	Streptomyces hygroscopicus
		O50423	O50423	Transcriptional regulator (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		Q8CJQ4	Q8CJQ4	Transcriptional regulator	Streptomyces coelicolor
		AAO65329	AAO65329	Putative transcriptional regulator	Streptomyces murayamaensis
		BAC70165	BAC70165	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		15608307		hypothetical protein Rv1167c	Mycobacterium tuberculosis H37Rv
		15840610		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
		29828996		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
555	LKSAYRRK	O53317	O53317	Putative regulatory protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15610303		hypothetical protein Rv3167c	Mycobacterium tuberculosis H37Rv
		15842744		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
556	IKTAYRRK	Q9EWI0	Q9EWI0	Putative tetR-family transcriptional regulator (Fragment)	Streptomyces coelicolor
		BAC72794	BAC72794	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29831625		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
557	MKNAYRRK	Q97DF9	Q97DF9	Transcriptional regulators, AcrR family	Clostridium acetobutylicum
		15896755		Transcriptional regulators, AcrR family	Clostridium acetobutylicum
558	IKGSYRRK	Q9RMY4	Q9RMY4	PXO2-48 (Transcriptional regulator, TetR family)	Bacillus anthracis
		AAP27253	AAP27253	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP08792	AAP08792	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		30263390		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30019960		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
559	VKGGYRRR	27376455		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
560	LKSTYRRK	O32899	O32899	Hypothetical 8.7 kDa protein	Mycobacterium leprae
561	LKTTYRRK	Q9L2A4	Q9L2A4	Putative transcriptional regulator	Streptomyces coelicolor
		BAC70542	BAC70542	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29829373		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
562	IKTTYRRA	Q8G725	Q8G725	Possible TetR-type transcriptional regulator	Bifidobacterium longum
		23465045		possible TetR-type transcriptional regulator	Bifidobacterium longum NCC2705
563	IKTTYRRR	AAN37481	AAN37481	Putative TetR-family transcriptional regulator	Arthrobacter sp. BP2
564	LKTTYRRP	O54180	O54180	Putative transcriptional regulator	Streptomyces coelicolor
565	VRTTYRRV	Q8GGM1	Q8GGM1	Putative tetR-family transcriptional regulatory	Streptomyces atroolivaceus
		Q9RKM8	Q9RKM8	Probable tetR family transcriptional regulator	Streptomyces coelicolor
566	LKTTYRRV	BAC72242	BAC72242	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29831073		putative TetR-family transcriptional	Streptomyces avermitilis

567	MRMTYRRV	Q9RJ96	Q9RJ96	regulator	MA-4680
				Putative TetR-family transcriptional regulator	Streptomyces coelicolor
568	MRVTYRRK	Q9L240	Q9L240	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
569	MQGYRRK	O05388	O05388	20 kDa polypeptide	Bacillus licheniformis
570	MVGYRRK	Q9KYF2	Q9KYF2	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
		BAC71822	BAC71822	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29830653		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
571	MKHTYRRR	Q92XG4	Q92XG4	Hypothetical protein RA1290	Sinorhizobium meliloti
		16263743		hypothetical protein	Sinorhizobium meliloti
572	MKGTFRRR	P95103	P95103	Hypothetical 22.1 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		Q9RD23	Q9RD23	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
		BAC74757	BAC74757	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		15610192		hypothetical protein Rv3055	Mycobacterium tuberculosis H37Rv
		15842622		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
		29833588		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
573	LRPTYRRT	Q10528	YM50_MYCTU	Hypothetical transcriptional regulator Rv2250c	Mycobacterium tuberculosis
		15609387		hypothetical protein Rv2250c	Mycobacterium tuberculosis H37Rv
		15841743		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
574	VRPTYRRK	Q8KQM0	Q8KQM0	ORFE	Saccharopolyspora erythraea
575	LKPTYSRK	27379796		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
576	MPPVYRRR	Q9RJX7	Q9RJX7	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
577	FKPVYRRP	28379515		transcription regulator	Lactobacillus plantarum WCFS1
578	VETSYRRR	Q9RJB8	Q9RJB8	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
579	VRTQYRQK	28379741		transcription regulator (putative)	Lactobacillus plantarum WCFS1
580	VDSSYRRL	Q9RL21	Q9RL21	Putative transcriptional regulator	Streptomyces coelicolor
581	LSSTYRRL	Q9EX36	Q9EX36	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
582	IATTYRRL	BAC76557	BAC76557	Putative TetR family receptor protein	Streptomyces rochei
583	VPSTYRRL	28870331		transcriptional regulator, putative	Pseudomonas syringae pv. tomato str. DC3000
584	IKPTYRWK	Q51730	Q51730	PhIF	Pseudomonas fluorescens
		Q9S615	Q9S615	Transcriptional regulatory protein PhIF	Pseudomonas fluorescens
		Q9RF02	Q9RF02	PhIF	Pseudomonas fluorescens
		Q98G69	Q98G69	Transcription regulator TetR/AcrR family homolog, YdeS	Mesorhizobium loti
		13472994		transcription regulator TetR/AcrR family homolog ydeS	Mesorhizobium loti
585	IKRTYRWK	Q8ZSI7	Q8ZSI7	Transcriptional regulator	Nostoc sp. PCC 7120
		17158659	all7523	transcriptional regulator	Nostoc sp. PCC 7120
586	IKQTYRWK	Q9RKP8	Q9RKP8	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
		BAC73632	BAC73632	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29832463		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
587	VKATYKWK	Q97DJ6	Q97DJ6	Transcriptional regulator, AcrR family	Clostridium acetobutylicum
		P96676	P96676	YDES protein	Bacillus subtilis
		AAP27079	AAP27079	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP10105	AAP10105	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		16077599	ydeS	similar to transcriptional regulator	Bacillus subtilis subsp.

		15896718		(TetR/AcrR family) Transcriptional regulator, AcrR family	subtilis str. 168 Clostridium acetobutylicum
		30263216		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30021273		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
588	IKMTYKWR	27376574		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
589	FKTTYKWK	Q9RK47	Q9RK47	Putative transcriptional regulator	Streptomyces coelicolor
590	MKETYKWR	Q98KM6	Q98KM6	Probable transcriptional regulator	Mesorhizobium loti
		13471436		probable transcriptional regulator	Mesorhizobium loti
591	TKESYKWR	Q8UDU9	Q8UDU9	Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
		Q92NV1	Q92NV1	Putative transcription regulator protein	Sinorhizobium meliloti
		Q8YI22	Q8YI22	Transcriptional regulator, TETR family	Brucella melitensis
		Q8FZU0	Q8FZU0	Transcriptional regulator, TetR family	Brucella melitensis biovar Suis
		15965822		PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	Sinorhizobium meliloti
		15889308		AGR_C_3662p	Agrobacterium tumefaciens str. C58 (Cereon)
		17935906		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
		17986906		TRANSCRIPTIONAL REGULATOR, TETR FAMILY	Brucella melitensis 16M
		23502252		transcriptional regulator, TetR family	Brucella suis 1330
592	LKSSYIFK	Q9ZN97	Q9ZN97	HOMOLOGUE of ACRR potential operon repressor of ESCHERICHIA COLI (ACRR potential operon repressor homolog)	Escherichia coli
593	IKSSYSFK	Q8TQQ3	Q8TQQ3	Transcription regulation protein	Methanosarcina acetivorans
		20090346	tetR	transcription regulation protein	Methanosarcina acetivorans C2A
594	VKAYYRFK	Q9PBQ5	Q9PBQ5	Transcriptional regulator (AcrR family)	Xylella fastidiosa
		15838676		transcriptional regulator (AcrR family)	Xylella fastidiosa 9a5c
595	VKAYYKFK	Q9A7D7	Q9A7D7	Transcriptional regulator, TetR family	Caulobacter vibrioides
		Q8XUI1	Q8XUI1	Putative transcription regulator protein	Ralstonia solanacearum
		Q8PQ91	Q8PQ91	Transcriptional regulator tetR/acrR family	Xanthomonas axonopodis pv. citri
		Q8PDC0	Q8PDC0	Transcriptional regulator tetR/acrR family	Xanthomonas campestris pv. campestris
		16126030		transcriptional regulator, TetR family	Caulobacter crescentus CB15
		26990018		transcriptional regulator, tetR family	Pseudomonas putida KT2440
		17547926	RSc3207	PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	Ralstonia solanacearum
		21229897	ameR	transcriptional regulator tetR/acrR family	Xanthomonas campestris pv. campestris str. ATCC 33913
		21241210	ameR	transcriptional regulator tetR/acrR family	Xanthomonas axonopodis pv. citri str. 306
		28868050		transcriptional regulator, TetR family	Pseudomonas syringae pv. tomato str. DC3000
596	MKASYKLK	P94423	P94423	DNA for 25-36 DEGREE region containing the AMYE-SRFA region, complete CDS	Bacillus subtilis
		16077453	ycnC	similar to transcriptional regulator (TetR/AcrR family)	Bacillus subtilis subsp. subtilis str. 168
597	IKASYSLK	Q8NLK1	Q8NLK1	Transcriptional regulator	Corynebacterium glutamicum
		Q8FLS2	Q8FLS2	Putative transcription regulator	Corynebacterium efficiens
		19554126		transcriptional regulator	Corynebacterium glutamicum ATCC 13032
		25029344		putative transcription regulator	Corynebacterium efficiens YS-314
598	VKRSYQLK	BAC67852	BAC67852	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29826683		putative TetR-family transcriptional	Streptomyces avermitilis

599	MKRITYKLK	29347374		regulator hypothetical protein	MA-4680 Bacteroides thetaiotaomicron VPI-5482
600	MKRITYEVK	29348098		conserved hypothetical protein	Bacteroides thetaiotaomicron VPI-5482
601	IKTTYKLK	27380850		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
602	TKQTYRLK	Q9F6W0	Q9F6W0	CasR	Rhizobium etli
603	ITKTYRLK	Q9KIH5	Q9KIH5	Putative regulator	Rhizobium etli
604	TTKTYRLK	27383068		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
605	TKATYACK	Q8XSM4	Q8XSM4	Probable transcription regulator protein	Ralstonia solanacearum
		17548664	RSp0443	PROBABLE TRANSCRIPTION REGULATOR PROTEIN	Ralstonia solanacearum
606	MKILYRLR	Q9A825	Q9A825	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16125786		transcriptional regulator, TetR family	Caulobacter crescentus CB15
607	LSSAYRYR	Q50617	YI16_MYCTU	Hypothetical transcriptional regula- tor Rv1816	Mycobacterium tuberculosis
		BAC69178	BAC69178	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		15608953		hypothetical protein Rv1816	Mycobacterium tuberculosis H37Rv
		15841286		hypothetical protein	Mycobacterium tuberculosis CDC1551
		29828009		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
608	LANAYGYR	BAC69069	BAC69069	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827900		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
609	LVQAYHYR	Q9KYZ0	Q9KYZ0	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
610	LTTTYSFK	Q8KU49	Q8KU49	EF0113	Enterococcus faecalis
		AAO80430	AAO80430	Transcriptional regulator, TetR family	Enterococcus faecalis
		29375206		transcriptional regulator, TetR family	Enterococcus faecalis V583
611	LTTAYSLK	BAC67855	BAC67855	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29826686		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
612	VPATYTYK	P71822	Y767_MYCTU	Hypothetical transcriptional regula- tor Rv0767c	Mycobacterium tuberculosis
		15607907		hypothetical protein Rv0767c	Mycobacterium tuberculosis H37Rv
		15840182		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
613	IPATYVYK	Q8VUA8	Q8VUA8	Putative transcriptional regulator	Lactococcus lactis
614	IPATYIYK	Q97LU7	Q97LU7	Transcriptional regulator, AcrR family	Clostridium acetobutylicum
		15893748		Transcriptional regulator, AcrR family	Clostridium acetobutylicum
615	MPATYVYK	Q8YW94	Q8YW94	Transcriptional regulator	Nostoc sp. PCC 7120
		17229213		transcriptional regulator	Nostoc sp. PCC 7120
616	MPGTIYIK	Q9CLS4	Q9CLS4	Hypothetical protein PM1135	Pasteurella multocida
		15603000		unknown	Pasteurella multocida
617	MPASYVYT	Q93JG8	Q93JG8	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
		BAC73472	BAC73472	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29832303		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
618	MPATYRYV	O54171	O54171	Hypothetical protein SCO5942	Streptomyces coelicolor
619	VPATYQYV	Q9F367	Q9F367	Putative tetR-family transcriptional regulatory protein	Streptomyces coelicolor
		BAC71530	BAC71530	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29830361		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
620	APSTYRYM	Q9F3P2	Q9F3P2	Putative tetR-family transcriptional regulator	Streptomyces coelicolor

621	VPRTFNYK	Q8KNI9	Q8KNI9	CalR1	Micromonospora echinospora
622	IPATFNYK	Q9K7M9 15615894	Q9K7M9 BH3332	Hypothetical protein BH3332 BH3332~unknown conserved protein in others	Bacillus halodurans Bacillus halodurans
623	IPSTFRYK	AAO65794 BAC71001 29829832	AAO65794 BAC71001	Putative monensin cluster negative regulator Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces cinnamonensis Streptomyces avermitilis Streptomyces avermitilis MA-4680
624	VPSTLRYR	Q9EWH2	Q9EWH2	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
625	IPAMIRYK	Q98FB5 13473299	Q98FB5	Transcriptional regulator (TetR-family) transcriptional regulator (TetR-family)	Mesorhizobium loti Mesorhizobium loti
626	IPSMMRYK	BAC68623 29827454	BAC68623	Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces avermitilis Streptomyces avermitilis MA-4680
627	LPATLLYK	Q8KYR0	Q8KYR0	Transcriptional regulator, TetR family, (pXO1-34)	Bacillus anthracis
628	VPANYRFK	Q986M1 Q92WK6 Q8UJU2 13476076 16264081 17938967 27377535	Q986M1 Q92WK6 Q8UJU2	Transcriptional regulator Putative transcriptional regulator protein Transcriptional regulator, TetR family transcriptional regulator putative transcriptional regulator protein transcriptional regulator, TetR family transcriptional regulatory protein	Mesorhizobium loti Sinorhizobium meliloti Agrobacterium tumefaciens str. C58 Mesorhizobium loti Sinorhizobium meliloti Agrobacterium tumefaciens str. C58 (U. Washington) Bradyrhizobium japonicum USDA 110
629	VPANYRFR	27382135		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
630	IPANFKHK	Q9F8V9 15889808 17936424	Q9F8V9	TetR family bacterial regulatory protein (Transcriptional regulator, TetR family) AGR_C_4617p transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 Agrobacterium tumefaciens str. C58 (Cereon) Agrobacterium tumefaciens str. C58 (U. Washington)
631	VPMRYGFK	Q9F3D3	Q9F3D3	Putative tetR-family regulatory protein	Streptomyces coelicolor
632	VPANYNYI	Q9WX68	Q9WX68	Hypothetical 29.9 kDa protein (ORF265 protein)	Gluconacetobacter xylinus
633	IPQSYNYL	28377930		transcription regulator (putative)	Lactobacillus plantarum WCFS1
634	ISATYNYL	O69405	O69405	Putative transcriptional regulator	Eubacterium acidaminophilum
635	ITATYNYL	Q8RI33 19705108	Q8RI33	Transcriptional regulator, TetR family Transcriptional regulator, TetR family	Fusobacterium nucleatum subsp. nucleatum Fusobacterium nucleatum subsp. nucleatum ATCC 25586
636	IPSSYHFV	Q9I6J7 15595491 26986985 28872423	Q9I6J7	Probable transcriptional regulator probable transcriptional regulator transcriptional regulator, TetR family transcriptional regulator, TetR family	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1 Pseudomonas putida KT2440 Pseudomonas syringae pv. tomato str. DC3000
637	MTMSYYK	Q9AMH9	Q9AMH9	Putative repressor simReg2 (Sim16)	Streptomyces antibioticus
638	MTMSYNYK	Q9S2B2 21220212	Q9S2B2 SCO1718	Putative transcriptional regulator putative regulator	Streptomyces coelicolor Streptomyces coelicolor A3(2)
639	MTMSYRYK	Q9F2Y9	Q9F2Y9	Putative tetR family transcriptional regulatory protein	Streptomyces coelicolor
640	IPMSYSYK	Q98GZ7	Q98GZ7	Transcriptional regulator	Mesorhizobium loti

641	IPVSYNYK	13472716 Q8DTH8	Q8DTH8	transcriptional regulator Putative transcriptional regulator (TetR family)	Mesorhizobium loti Streptococcus mutans
642	IPKSYHYK	24379772 Q9RZ11	Q9RZ11	putative transcriptional regulator (TetR family) Transcriptional regulator, TETR family	Streptococcus mutans UA159 Deinococcus radiodurans
643	VPMAYYYK	15807811 28901154	DRA0142	transcriptional regulator, TetR family putative transcriptional regulator	Deinococcus radiodurans Vibrio parahaemolyticus RIMD 2210633
644	IQVTYNYK	Q9KE56	Q9KE56	Transcriptional regulator (TetR family)	Bacillus halodurans
645	IQVSYNYK	15613565 Q92AU2 Q8Y6H0 16803756	BH1002 Q92AU2 Q8Y6H0	transcriptional regulator (TetR family) Hypothetical protein lin1827 Hypothetical protein lmo1716 similar to putative transcription regulators	Bacillus halodurans Listeria innocua Listeria monocytogenes Listeria monocytogenes EGD-e Listeria innocua
646	IQVTYNHK	16800894 AAP27405 AAP10525	AAP27405 AAP10525	similar to putative transcription regulators Transcriptional regulator, TetR family Transcriptional regulator, TetR family	Bacillus anthracis str. Ames Bacillus cereus ATCC 14579
647	VAASYNYK	30263542 30021693 O86342 15607414	30263542 30021693 O86342	transcriptional regulator, TetR family Transcriptional regulator, TetR family Hypothetical protein Rv0273c hypothetical protein Rv0273c	Bacillus anthracis str. Ames Bacillus cereus ATCC 14579 Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv
648	VHTSYVYK	15839654 Q9KBT8 Q9K6J6	Q9KBT8 Q9K6J6	conserved hypothetical protein Hypothetical protein BH1836 Transcriptional regulator (TetR/AcrR family)	Mycobacterium tuberculosis CDC1551 Bacillus halodurans Bacillus halodurans
649	LPTSyrHV	15614399 15616295 P27307 Q8XGD6	BH1836 BH3733 YIJC_ECOLI Q8XGD6	BH1836-unknown conserved pro- tein transcriptional regulator (TetR/AcrR family) Hypothetical transcriptional regula- tor yijC Putative transcriptional repressor (TetR/AcrR family) (Possible TetR- family transcriptional regulatory protein)	Bacillus halodurans Bacillus halodurans Escherichia coli Shigella flexneri Salmonella typhimurium
		Q8X725 Q8ZA98	Q8X725 Q8ZA98	Orf, hypothetical protein TetR-family transcriptional regula- tory protein (Hypothetical protein)	Salmonella typhi Escherichia coli O157:H7 Yersinia pestis
		Q8FB90 AAP18728	Q8FB90 AAP18728	Hypothetical protein yijC Hypothetical protein yijC	Escherichia coli O6 Shigella flexneri 2a str. 2457T
		AAO71006	AAO71006	Possible TetR-family transcriptional regulatory protein	Salmonella typhi
		16131801 15804561	yijC yijC	orf, hypothetical protein orf, hypothetical protein	Escherichia coli K12 Escherichia coli O157:H7 EDL933
		15834148 16124045		hypothetical protein TetR-family transcriptional regula- tory protein	Escherichia coli O157:H7 Yersinia pestis CO92
		16767391	yijC	putative transcriptional repressor (TetR/AcrR family)	Salmonella typhimurium LT2
		16762316	yijC	possible TetR-family transcriptional regulatory protein	Salmonella enterica subsp. enterica serovar Typhi
		22124239 24115257 26250740 29143804	yijC yijC yijC	hypothetical protein orf, conserved hypothetical protein Hypothetical protein yijC	Yersinia pestis KIM Shigella flexneri 2a str. 301 Escherichia coli CFT073
		30064746	yijC	possible TetR-family transcriptional regulatory protein hypothetical protein	Salmonella enterica subsp. enterica serovar Typhi Ty2 Shigella flexneri 2a str.

650	LPTSyrHM	P44757	YIJC_HAEIN	Hypothetical transcriptional regulator HI0570	2457T Haemophilus influenzae
		Q9KVVJ2	Q9KVVJ2	Transcriptional regulator, TetR family	Vibrio cholerae
		Q8EK97	Q8EK97	Transcriptional regulator, TetR family	Shewanella oneidensis
		Q8DD45	Q8DD45	Transcriptional regulator	Vibrio vulnificus
		16272513		hypothetical protein	Haemophilus influenzae Rd
		15640183		transcriptional regulator, TetR family	Vibrio cholerae
		24371798		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
		27364585		Transcriptional regulator	Vibrio vulnificus CMCP6
		28899715		transcriptional regulator, TetR family	Vibrio parahaemolyticus RIMD 2210633
651	LPTAYRHT	O86578	O86578	Putative transcriptional regulator	Streptomyces coelicolor
		BAC70470	BAC70470	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29829301		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
652	LPAGYRHM	Q9HUS3	Q9HUS3	Hypothetical protein PA4890	Pseudomonas aeruginosa
		15600083		conserved hypothetical protein	Pseudomonas aeruginosa PAO1
653	LPTGYRHM	26991539		transcriptional regulator, TetR family	Pseudomonas putida KT2440
		28872022		transcriptional regulator, TetR family	Pseudomonas syringae pv. tomato str. DC3000
654	LATSyrHM	Q9CL95	Q9CL95	Hypothetical protein PM1345	Pasteurella multocida
		15603210		unknown	Pasteurella multocida
655	IPTSyQHK	Q8G5U4	Q8G5U4	Possible TetR-type transcriptional regulator	Bifidobacterium longum
		23465482		possible TetR-type transcriptional regulator	Bifidobacterium longum NCC2705
656	LPNTYRHI	Q9I3H4	Q9I3H4	Hypothetical protein PA1539	Pseudomonas aeruginosa
		15596736		hypothetical protein	Pseudomonas aeruginosa PAO1
657	LPNAYRHR	Q8XS82	Q8XS82	Putative transcription regulator protein	Ralstonia solanacearum
		AAM78000	AAM78000	Transcriptional regulator	Streptomyces carzinostaticus subsp. neocarzinostaticus
		17548820	RSp0599	PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	Ralstonia solanacearum
658	MRTSYRHV	Q937A9	Q937A9	Putative transcriptional regulator	Pseudomonas sp. ADP
659	LPNSYRQM	Q8PFL9	Q8PFL9	Transcriptional regulator tetR family	Xanthomonas axonopodis pv. citri
		Q8P436	Q8P436	Transcriptional regulator tetR family	Xanthomonas campestris pv. campestris
		21233303		transcriptional regulator tetR family	Xanthomonas campestris pv. campestris str. ATCC 33913
		21244680		transcriptional regulator tetR family	Xanthomonas axonopodis pv. citri str. 306
660	MPPTYHHL	Q51597	Q51597	CaM repressor	Pseudomonas putida
661	LPAAYHHK	Q9A690	Q9A690	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16126443		transcriptional regulator, TetR family	Caulobacter crescentus CB15
662	MPMSYHHK	BAC68286	BAC68286	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		BAC68287	BAC68287	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827117		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
		29827118		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
663	MPMAYKHK	BAC68198	BAC68198	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827029		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
664	MPSSYRHR	Q9RWE9	Q9RWE9	Transcriptional regulator, TETR family	Deinococcus radiodurans
		15805746	DR0720	transcriptional regulator, TetR family	Deinococcus radiodurans
665	MPSSYSHR	Q9EWK8	Q9EWK8	Putative transcriptional regulator	Streptomyces coelicolor
666	MPRAYNHR	Q8NMW8	Q8NMW8	Transcriptional regulator	Corynebacterium glutamicum

		Q8FN10	Q8FN10	Putative transcription regulator	<i>Corynebacterium efficiens</i>
		23308939		transcriptional regulator	<i>Corynebacterium glutamicum</i> ATCC 13032
		25028894		putative transcription regulator	<i>Corynebacterium efficiens</i> YS-314
667	AVGAFRHM	Q10829	YT12_MYCTU	Hypothetical transcriptional regulator Rv2912c	<i>Mycobacterium tuberculosis</i>
		15610049		hypothetical protein Rv2912c	<i>Mycobacterium tuberculosis</i> H37Rv
		15842455		transcriptional regulator, TetR family	<i>Mycobacterium tuberculosis</i> CDC1551
668	VPGAFRHR	Q8GNI4 27381603	Q8GNI4	Putative transcriptional regulator transcriptional regulatory protein	<i>Bradyrhizobium japonicum</i> <i>Bradyrhizobium japonicum</i> USDA 110
669	LKTNLYYK	P75899	YCDC_ECOLI	Hypothetical transcriptional regulator ycdC	<i>Escherichia coli</i> O6
		Q8ZQ38	Q8ZQ38	Putative transcriptional repressor (TetR/AcrR family)	<i>Shigella flexneri</i> <i>Salmonella typhimurium</i>
		Q8Z7N7 Q8X4Z7 AAP16527	Q8Z7N7 Q8X4Z7 AAP16527	Putative transcriptional regulator Putative tet operon regulator Putative tet operon regulator	<i>Salmonella typhi</i> <i>Escherichia coli</i> O157:H7 <i>Shigella flexneri</i> 2a str. 2457T
		AAO69421 16128979 15801002	AAO69421 ycdC ycdC	Putative transcriptional regulator putative tet operon regulator putative tet operon regulator	<i>Salmonella typhi</i> <i>Escherichia coli</i> K12 <i>Escherichia coli</i> O157:H7 EDL933
		15830513 16764479	ycdC	putative tet operon regulator putative transcriptional repressor (TetR/AcrR family)	<i>Escherichia coli</i> O157:H7 <i>Salmonella typhimurium</i> LT2
		16759999		putative transcriptional regulator	<i>Salmonella enterica</i> subsp. enterica serovar Typhi
		24112425 26247035	ycdC ycdC	putative tet operon regulator Hypothetical transcriptional regulator ycdC	<i>Shigella flexneri</i> 2a str. 301 <i>Escherichia coli</i> CFT073
		28868373		transcriptional regulator, TetR family	<i>Pseudomonas syringae</i> pv. tomato str. DC3000
		29142230		putative transcriptional regulator	<i>Salmonella enterica</i> subsp. enterica serovar Typhi Ty2
		30062550	ycdC	putative tet operon regulator	<i>Shigella flexneri</i> 2a str. 2457T
670	LKANLYYK	Q9I2N1 15597061	Q9I2N1	Probable transcriptional regulator probable transcriptional regulator	<i>Pseudomonas aeruginosa</i> <i>Pseudomonas aeruginosa</i> PAO1
671	IKPNLYYK	Q936K1	Q936K1	Hypothetical transcriptional regulator	<i>Rhizobium leguminosarum</i>
		Q98K52 Q8UCU9	Q98K52 Q8UCU9	Probable transcriptional regulator Transcriptional regulator, TetR family	<i>Mesorhizobium loti</i> <i>Agrobacterium tumefaciens</i> str. C58
		Q92MZ5	Q92MZ5	Putative transcription regulator protein	<i>Sinorhizobium meliloti</i>
		Q8G2N8	Q8G2N8	Transcriptional regulator, TetR family	<i>Brucella melitensis</i> biovar Suis
		13471610 15966205		probable transcriptional regulator PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	<i>Mesorhizobium loti</i> <i>Sinorhizobium meliloti</i>
		15889655		AGR_C_4325p	<i>Agrobacterium tumefaciens</i> str. C58 (Cereon)
		17936266		transcriptional regulator, TetR family	<i>Agrobacterium tumefaciens</i> str. C58 (U. Washington)
		23501187		transcriptional regulator, TetR family	<i>Brucella suis</i> 1330
672	VKPNLYYK	Q9A4N6	Q9A4N6	Transcriptional regulator, TetR family	<i>Caulobacter vibrioides</i>
		16127026		transcriptional regulator, TetR family	<i>Caulobacter crescentus</i> CB15
673	TKPNYYYYK	Q9I3K8 15596701	Q9I3K8	Probable transcriptional regulator probable transcriptional regulator	<i>Pseudomonas aeruginosa</i> <i>Pseudomonas aeruginosa</i> PAO1
		26990986		transcriptional regulator, TetR family	<i>Pseudomonas putida</i> KT2440
		28870853		transcriptional regulator, TetR family	<i>Pseudomonas syringae</i> pv. tomato str. DC3000
674	YKPSYYYYK	AAP09132	AAP09132	Transcriptional regulator, TetR family	<i>Bacillus cereus</i> ATCC 14579

		30020300		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
675	MKANHYK	Q9I6W5 15595365	Q9I6W5	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
		26987332		transcriptional regulator, TetR family	Pseudomonas putida KT2440
		28867999		transcriptional regulator, TetR family	Pseudomonas syringae pv. tomato str. DC3000
676	MRSNHYYK	Q9I681 15595633	Q9I681	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
677	MSGHLYK	O86510 BAC70424	O86510 BAC70424	Putative transcriptional regulator Putative TetR-family transcriptional regulator	Streptomyces coelicolor Streptomyces avermitilis
		29829255		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
678	MTGATYYK	26989543		transcriptional regulator, TetR family	Pseudomonas putida KT2440
679	VRPTYFYK	Q9CDD3 Q49962	Q9CDD3 Q49962	Putative transcriptional regulator U1756B (Putative TETR-family transcriptional regulator)	Mycobacterium leprae Mycobacterium leprae
		P96222	P96222	Hypothetical 23.7 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15610991		hypothetical protein Rv3855	Mycobacterium tuberculosis H37Rv
		15826909		putative transcriptional regulator	Mycobacterium leprae
		15827524		putative TetR-family transcriptional regulator	Mycobacterium leprae
		15843486		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
680	VQPSYLYK	P96701	P96701	YDGC protein	Bacillus subtilis
		16077625	ydgC	similar to hypothetical proteins	Bacillus subtilis subsp. subtilis str. 168
681	MKPMYLYK	O69500	O69500	Putative transcriptional regulator (Possible TETR-family transcriptional regulator)	Mycobacterium leprae
		O53661	O53661	Putative regulatory protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15607379		hypothetical protein Rv0238	Mycobacterium tuberculosis H37Rv
		15828381		possible TetR-family transcriptional regulator	Mycobacterium leprae
		15839619		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
682	MKPLYLYK	Q9L0U5	Q9L0U5	Putative tetR-family regulatory protein	Streptomyces coelicolor
		BAC72528	BAC72528	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29831359		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
683	MKPAYYYK	Q8XXA3	Q8XXA3	Putative transcription regulator transcription regulator protein	Ralstonia solanacearum
		Q8TLT0	Q8TLT0	Transcriptional regulator, TetR family	Methanosarcina acetivorans
		17546932	RSc2213	PUTATIVE TRANSCRIPTION REGULATORY PROTEIN	Ralstonia solanacearum
		20091767		transcriptional regulator, TetR family	Methanosarcina acetivorans C2A
684	MRPAHYK	29346816		transcriptional regulator, TetR family	Bacteroides thetaiotaomicron VPI-5482
685	MKPAYRYK	Q8TLS7	Q8TLS7	Transcriptional regulator, TetR family	Methanosarcina acetivorans
		20091770		transcriptional regulator, TetR family	Methanosarcina acetivorans C2A
686	MRPAYLQK	Q98FC5 13473289	Q98FC5	Transcriptional regulator transcriptional regulator	Mesorhizobium loti Mesorhizobium loti
687	MRPAYLIK	Q8U6D4	Q8U6D4	Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
		15890117		AGR_L_32p	Agrobacterium tumefaciens str. C58 (Cereon)
		17938562		transcriptional regulator, TetR family	Agrobacterium tumefaciens

688	MRPGYFLK	Q9EX61	Q9EX61	family Putative tetR-family transcriptional regulator		str. C58 (U. Washington) <i>Streptomyces coelicolor</i>
689	MRPTYTYT	Q9RZK1	Q9RZK1	Transcriptional regulator, family	TETR	<i>Deinococcus radiodurans</i>
		10957425		transcriptional regulator, family	TetR	<i>Deinococcus radiodurans</i>
690	MRPNYTYT	Q9RUK5	Q9RUK5	Transcriptional repressor, family	TETR	<i>Deinococcus radiodurans</i>
		Q9RRT9	Q9RRT9	Transcriptional repressor, family	TETR	<i>Deinococcus radiodurans</i>
		15806396	DR1379	transcriptional repressor, family	TetR	<i>Deinococcus radiodurans</i>
		15807386	DR2396	transcriptional repressor, family	TetR	<i>Deinococcus radiodurans</i>
691	PRPNYTYT	Q9RUK2	Q9RUK2	Transcriptional regulator, family	TETR	<i>Deinococcus radiodurans</i>
		15806401	DR1384	transcriptional regulator, family	TetR	<i>Deinococcus radiodurans</i>
692	VRPTYSYT	Q9RY76	Q9RY76	Transcriptional regulator, family	TETR	<i>Deinococcus radiodurans</i>
		15805115	DR0074	transcriptional regulator, family	TetR	<i>Deinococcus radiodurans</i>
693	MRGGYLYT	O31560 16077904	O31560 yfiR	YFIR protein similar to hypothetical proteins		<i>Bacillus subtilis</i> <i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168
694	VRSGYFYK	P95100	P95100	Hypothetical 23.8 kDa protein (Transcriptional regulator, family)	TetR	<i>Mycobacterium tuberculosis</i>
		15610195		hypothetical protein Rv3058c		<i>Mycobacterium tuberculosis</i> H37Rv
		15842625		transcriptional regulator, family	TetR	<i>Mycobacterium tuberculosis</i> CDC1551
695	LAPTYAEK	P75952	YCFQ_ECOLI	Hypothetical transcriptional regulator ycfQ		<i>Escherichia coli</i>
		Q8ZQ04	Q8ZQ04	Putative transcriptional repressor (TetR/AcrR family)		<i>Salmonella typhimurium</i>
		Q8Z714	Q8Z714	Putative TetR-family regulatory protein		<i>Salmonella typhi</i>
		Q8X8G3 Q8FIN0	Q8X8G3 Q8FIN0	YcfQ protein Hypothetical transcriptional regulator ycfQ		<i>Escherichia coli</i> O157:H7 <i>Escherichia coli</i> O6
		AAN42733 AAP16621	AAN42733 AAP16621	YcfQ gene product Hypothetical protein ycfQ		<i>Shigella flexneri</i> <i>Shigella flexneri</i> 2a str. 2457T
		AAO69332	AAO69332	Putative TetR-family regulatory protein		<i>Salmonella typhi</i>
		16129074	ycfQ	orf, hypothetical protein		<i>Escherichia coli</i> K12
		15801228	ycfQ	-		<i>Escherichia coli</i> O157:H7 EDL933
		15830743 16764568	ycfQ	hypothetical protein putative transcriptional repressor (TetR/AcrR family)		<i>Escherichia coli</i> O157:H7 <i>Salmonella typhimurium</i> LT2
		16760088		putative TetR-family regulatory protein		<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi
		24112516 26247255	ycfQ ycfQ	ycfQ gene product Hypothetical transcriptional regulator ycfQ		<i>Shigella flexneri</i> 2a str. 301 <i>Escherichia coli</i> CFT073
		29142141		putative TetR-family regulatory protein		<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi Ty2
		30062644	ycfQ	hypothetical protein		<i>Shigella flexneri</i> 2a str. 2457T
696	VPPSYAAK	P39334	YJGJ_ECOLI	Hypothetical transcriptional regulator yjgJ		<i>Escherichia coli</i>
		Q8ZP73	Q8ZP73	Putative bacterial regulatory helix-turn-helix proteins, araC family		<i>Salmonella typhimurium</i>
		Q8Z7A0 Q8PNB2	Q8Z7A0 Q8PNB2	Putative transcriptional regulator Hypothetical protein XAC1161		<i>Salmonella typhi</i> <i>Xanthomonas axonopodis</i> pv. <i>citri</i>
		Q8PBR1	Q8PBR1	Hypothetical protein XCC1058		<i>Xanthomonas campestris</i> pv. <i>campestris</i>
		AAN45658 AAO69209 16765017	AAN45658 AAO69209	Orf, conserved hypothetical protein Putative transcriptional regulator putative bacterial regulatory helix-turn-helix proteins, araC family		<i>Shigella flexneri</i> <i>Salmonella typhi</i> <i>Salmonella typhimurium</i> LT2

		16760213		putative transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi
		21230516		conserved hypothetical protein	Xanthomonas campestris pv. campestris str. ATCC 33913
		21241915		conserved hypothetical protein	Xanthomonas axonopodis pv. citri str. 306
		24115441	yjgJ	orf, conserved hypothetical protein	Shigella flexneri 2a str. 301
		29142018		putative transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi Ty2
697	LPPSYAAK	Q9HZR6 Q8U5U6	Q9HZR6 Q8U5U6	Probable transcriptional regulator AGR_pAT_422p	Pseudomonas aeruginosa Agrobacterium tumefaciens str. C58
		15598127		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		16119523		AGR_pAT_422p	Agrobacterium tumefaciens str. C58 (Cereon)
698	IPPSYACK	27378984		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
699	IPPSYATK	Q8UBE1	Q8UBE1	Transcriptional regulator, TetR family AGR_L_3469p	Agrobacterium tumefaciens str. C58 Agrobacterium tumefaciens str. C58 (Cereon)
		17936786		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
700	LPPSYAVK	27381370		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
701	ISPSYAAK	Q98ML2 13470746	Q98ML2	Probable transcriptional regulator probable transcriptional regulator	Mesorhizobium loti Mesorhizobium loti
702	LAPSYAAK	Q92WA7	Q92WA7	Putative transcriptional regulator protein	Sinorhizobium meliloti
		16264185		putative transcriptional regulator protein	Sinorhizobium meliloti
703	VAPSYAAK	Q9K485	Q9K485	Putative tetR-family protein	Streptomyces coelicolor
704	LPPQYNAK	Q98K79 13471583	Q98K79	Probable transcriptional regulator probable transcriptional regulator	Mesorhizobium loti Mesorhizobium loti
705	LRTSYAAK	Q9KXT8	Q9KXT8	Putative transcriptional regulator	Streptomyces coelicolor
706	LKPSYAAK	Q8EBL7	Q8EBL7	Transcriptional regulator, TetR family	Shewanella oneidensis
		24375000		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
707	LRPSYGAK	27382285		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
708	IRSSYNTK	Q9RBX8	Q9RBX8	Hypothetical 24.1 kDa protein	Vogesella indigofera
709	IRFSYNAK	Q9KM03 15601345	Q9KM03	Hypothetical protein VCA0586 conserved hypothetical protein	Vibrio cholerae Vibrio cholerae
710	IRFSYNSK	Q8EGM2	Q8EGM2	Transcriptional regulator, TetR family	Shewanella oneidensis
		24373148		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
711	IRFSYNTK	Q8D6S7 27366885 28900907	Q8D6S7	Transcriptional regulator Transcriptional regulator putative TetR-family	Vibrio vulnificus Vibrio vulnificus CMCP6 Vibrio parahaemolyticus RIMD 2210633
712	LRQSYDTK	O31510 16077753	O31510 yeeK	YEEK protein similar to hypothetical proteins	Bacillus subtilis Bacillus subtilis subsp. subtilis str. 168
713	VRNSYKTK	Q9RIX0	Q9RIX0	Putative membrane protein	Streptomyces coelicolor
714	MRASYATK	Q9L2D0	Q9L2D0	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
715	IRPSYRNK	Q9PQK1 13357850	Q9PQK1	Hypothetical protein UU290 unique hypothetical	Ureaplasma parvum Ureaplasma urealyticum
716	VKPSYRSK	Q98ER4 13473500	Q98ER4	Transcriptional regulatory protein transcriptional regulatory protein	Mesorhizobium loti Mesorhizobium loti
717	IRAA YRNK	Q97RQ5	Q97RQ5	Transcriptional regulator, TetR family	Streptococcus pneumoniae
		Q8DQI5	Q8DQI5	Conserved hypothetical protein	Streptococcus pneumoniae R6
		15900638		transcriptional regulator, TetR family	Streptococcus pneumoniae TIGR4
		15902697		Conserved hypothetical protein	Streptococcus pneumoniae R6
718	IRNAYRNK	Q99YH6	Q99YH6	Putative transcription regulator, TetR family	Streptococcus pyogenes MGAS8232 Streptococcus pyogenes MGAS315

		15675556		putative transcription regulator	Streptococcus pyogenes M1 GAS
		19746607	acrR	putative transcription regulator	Streptococcus pyogenes MGAS8232
		21911015		putative transcription regulator, TetR family	Streptococcus pyogenes MGAS315
		28895300		putative transcription regulator	Streptococcus pyogenes SSI-1
719	IRRTYRNK	Q97L43	Q97L43	Transcriptional regulator, AcrR family	Clostridium acetobutylicum
		15894010		Transcriptional regulator, AcrR family	Clostridium acetobutylicum
720	VRRTYRNK	Q97L42	Q97L42	Transcriptional regulator, AcrR family	Clostridium acetobutylicum
		15894011		Transcriptional regulator, AcrR family	Clostridium acetobutylicum
721	LRGTYSK	Q9HXS0	Q9HXS0	Probable transcriptional regulator	Pseudomonas aeruginosa
		15598916		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
722	LRSSYSK	Q9L119	Q9L119	Putative transcriptional regulatory protein	Streptomyces coelicolor
723	IRGAYSNL	Q9K4L1	Q9K4L1	Putative TetR-family regulator	Streptomyces coelicolor
724	IRGAYSNK	CAD18970	CAD18970	Putative lactone-dependent transcriptional regulator	Streptomyces cattleya
		27381649		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
725	VRGSYAAK	Q9ETD1	Q9ETD1	Hypothetical 21.2 kDa protein (Fragment)	Streptomyces coelicolor A3(2)
		Q99QK8	Q99QK8	Putative TetR-family regulator	Streptomyces coelicolor
		21234041	SCP1.42c	putative TetR-family regulator	Streptomyces coelicolor A3(2)
		21234309	SCP1.312	putative TetR-family regulator	Streptomyces coelicolor A3(2)
726	VQGSYAAK	Q9K3Q4	Q9K3Q4	Putative tetR-family regulatory protein	Streptomyces coelicolor
727	IRGSYGAK	BAC68415	BAC68415	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827246		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
728	LRGSYGAK	BAC72886	BAC72886	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29831717		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
729	LRGSYNAK	27382409		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
730	IAGSYKAK	Q92Z05	Q92Z05	Hypothetical protein RA0707	Sinorhizobium meliloti
		16263160		Conserved hypothetical protein	Sinorhizobium meliloti
731	ITGSYKAK	27377688		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
732	VQGSYKAK	Q8UDJ0	Q8UDJ0	Transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58
		15889416		AGR_C_3874p	Agrobacterium tumefaciens str. C58 (Cereon)
		17936019		transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 (U. Washington)
733	LKGSYKAK	27383296		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
734	VNGSFHAK	27379185		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
735	MPGSYCAK	Q9KKV1	Q9KKV1	Transcriptional regulator, TetR family	Vibrio cholerae
		Q8EBV2	Q8EBV2	Transcriptional regulator, TetR family	Shewanella oneidensis
		15601752		transcriptional regulator, TetR family	Vibrio cholerae
		24374904		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
736	VHGGHKQK	Q8GGS9	Q8GGS9	Putative tetR-family transcriptional regulator	Streptomyces atroolivaceus
737	VHGGYKQK	BAC68778	BAC68778	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		BAC69767	BAC69767	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827609		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680

		29828598		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
738	VHGAYKQK	Q9RI62	Q9RI62	Putative tetR-family transcriptional regulator	<i>Streptomyces coelicolor</i>
739	LQGAYKQK	Q98HK7	Q98HK7	Transcriptional regulator	<i>Mesorhizobium loti</i>
740	LQGGYKQK	Q92Y50	Q92Y50	Putative transcriptional regulator	<i>Sinorhizobium meliloti</i>
		Q8U882	Q8U882	Transcriptional regulator, TetR family	<i>Agrobacterium tumefaciens</i> str. C58
		16263491		Putative transcriptional regulator	<i>Sinorhizobium meliloti</i>
		15890765		AGR_L_1283p	<i>Agrobacterium tumefaciens</i> str. C58 (Cereon)
		17937906		transcriptional regulator, TetR family	<i>Agrobacterium tumefaciens</i> str. C58 (U. Washington)
741	LTGAYSQK	Q9I077	Q9I077	Probable transcriptional regulator	<i>Pseudomonas aeruginosa</i>
		15597962		probable transcriptional regulator	<i>Pseudomonas aeruginosa</i> PAO1
742	LTGAYKRRK	Q8DTZ5	Q8DTZ5	Putative transcription regulator	<i>Streptococcus mutans</i>
		24379595		putative transcription regulator	<i>Streptococcus mutans</i> UA159
743	IVGAYSRK	27379010		transcriptional regulatory protein	<i>Bradyrhizobium japonicum</i> USDA 110
744	SVGSYHTK	Q9CIV3	Q9CIV3	Transcriptional regulator	<i>Lactococcus lactis</i> subsp. <i>lactis</i>
		15672235	ycfA	transcription regulator	<i>Lactococcus lactis</i> subsp. <i>lactis</i>
745	ITGAFASK	Q9AA74	Q9AA74	Transcriptional regulator, TetR family	<i>Caulobacter vibrioides</i>
		16124985		transcriptional regulator, TetR family	<i>Caulobacter crescentus</i> CB15
746	IVGSFTTK	Q9AA73	Q9AA73	Transcriptional regulator, TetR family	<i>Caulobacter vibrioides</i>
		16124986		transcriptional regulator, TetR family	<i>Caulobacter crescentus</i> CB15
747	AQALFYHV	Q93HL5	Q93HL5	Regulator protein	<i>Streptomyces avermitilis</i>
		BAC70556	BAC70556	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29829387		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
748	IKSLFHKK	Q8VV87	Q8VV87	LexA repressor-like protein	<i>Terrabacter</i> sp.
749	IKSLFYHK	O28458	O28458	Transcriptional regulatory protein, TETR family	<i>Archaeoglobus fulgidus</i>
		11499405		transcriptional regulatory protein, TetR family	<i>Archaeoglobus fulgidus</i> DSM 4304
750	MNALVFHK	Q98NA2	Q98NA2	Hypothetical protein mlr0230	<i>Mesorhizobium loti</i>
		13470504		unknown protein	<i>Mesorhizobium loti</i>
751	INALLYHK	Q9X7X0	Q9X7X0	Putative transcriptional regulator	<i>Streptomyces coelicolor</i>
		BAC69348	BAC69348	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29828179		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
752	LKAALYHK	BAC68973	BAC68973	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29827804		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
753	LAALIHKK	O86312	O86312	Hypothetical 23.2 kDa protein (Transcriptional regulator, TetR family)	<i>Mycobacterium tuberculosis</i>
		15608359		hypothetical protein Rv1219c	<i>Mycobacterium tuberculosis</i> H37Rv
		15840663		transcriptional regulator, TetR family	<i>Mycobacterium tuberculosis</i> CDC1551
754	LAALKYHK	27382916		transcriptional regulatory protein	<i>Bradyrhizobium japonicum</i> USDA 110
755	LHSLRHKK	Q9KLH7	Q9KLH7	Transcriptional regulator, TetR family	<i>Vibrio cholerae</i>
		Q8D7G1	Q8D7G1	Transcriptional regulator	<i>Vibrio vulnificus</i>
		15601522		transcriptional regulator, TetR family	<i>Vibrio cholerae</i>
		27366648		Transcriptional regulator	<i>Vibrio vulnificus</i> CMCP6
		28900661		transcriptional regulator, TetR family	<i>Vibrio parahaemolyticus</i> RIMD 2210633
756	LHANIHHI	Q9A8K8	Q9A8K8	Transcriptional regulator, TetR family	<i>Caulobacter vibrioides</i>

				family			
		16125594		transcriptional regulator,	TetR	Caulobacter	crescentus
757	LHAAAHHL	Q8PNX2	Q8PNX2	family Transcriptional regulator		Xanthomonas	axonopodis
		21241687		transcriptional regulator		Xanthomonas	axonopodis
758	LHAATHHL	27380900		transcriptional regulatory protein		Bradyrhizobium	japonicum
759	LHAAAHHK	Q9AK28	Q9AK28	Putative tetR-family transcriptional regulator		Streptomyces	coelicolor
760	LHAAAHRH	BAC71258	BAC71258	Putative TetR-family transcriptional regulator		Streptomyces	avermitilis
		29830089		putative TetR-family transcriptional regulator		Streptomyces	avermitilis
761	YKPSHHHK	Q8VML7	Q8VML7	Putative transcriptional regulator		Pseudomonas	putida
		27379117		transcriptional regulatory protein		Bradyrhizobium	japonicum
762	YKASHHHK	Q98DK2	Q98DK2	Transcriptional regulator (TetR-family)		Mesorhizobium	loti
		Q8FDY0	Q8FDY0	Hypothetical protein		Escherichia coli O6	
		13473915		transcriptional regulator (tetR-family)		Mesorhizobium	loti
		26249435		Hypothetical protein		Escherichia coli CFT073	
		28871402		transcriptional regulator, TetR family	TetR	Pseudomonas syringae	pv. tomato str. DC3000
		28871407		transcriptional regulator, TetR family	TetR	Pseudomonas syringae	pv. tomato str. DC3000
763	YNASHHHK	Q8UIL4	Q8UIL4	Transcriptional regulator, TetR family	TetR	Agrobacterium	tumefaciens str. C58
		15887628		AGR_C_478p		Agrobacterium	tumefaciens str. C58 (Cereon)
		17934195		transcriptional regulator, TetR family	TetR	Agrobacterium	tumefaciens str. C58 (U. Washington)
764	YTASHYHK	AAO90352	AAO90352	Transcriptional regulator, TetR family	TetR	Coxiella burnetii	
		29654146		transcriptional regulator, TetR family	TetR	Coxiella burnetii	RSA 493
765	FAASHYHK	27383095		transcriptional regulatory protein		Bradyrhizobium	japonicum
766	FGGAHHHK	Q983U1	Q983U1	Transcription regulator		Mesorhizobium	loti
		13476759		transcription regulator		Mesorhizobium	loti
767	YKTAHYHK	Q97TL3	Q97TL3	HTH transcriptional regulator		Clostridium	acetobutylicum
		15004791		TetR/AcrR family		Clostridium	acetobutylicum
768	YNSSHYYK	Q8TPY2	Q8TPY2	HTH transcriptional regulator		Clostridium	acetobutylicum
		20090622		TetR/AcrR family	TetR	Methanosarcina	acetivorans
				transcriptional regulator, TetR family	TetR	Methanosarcina	acetivorans
769	YNAAHNYK	Q8PZR0	Q8PZR0	Transcriptional regulator		Methanosarcina	mazei
		21226534		transcriptional regulator		Methanosarcina	mazei
770	YRTGSHHK	Q9ANS7	Q9ANS7	LuxT		Methanosarcina	mazei
		Q9KL32	Q9KL32	Transcriptional regulator, TetR family	TetR	Vibrio	harveyi
		Q8EAZ4	Q8EAZ4	Transcriptional regulator, TetR family	TetR	Vibrio	cholerae
		Q8D3S9	Q8D3S9	Transcriptional regulator LuxT		Shewanella	oneidensis
		15601671		transcriptional regulator, TetR family	TetR	Vibrio	vulnificus
		24375239		transcriptional regulator, TetR family	TetR	Vibrio	cholerae
		27367950		Transcriptional regulator LuxT		Shewanella	oneidensis MR-1
		28900275		LuxT		Vibrio	vulnificus
771	YKSGIYHR	Q8NMG3	Q8NMG3	Transcriptional regulator		Vibrio	parahaemolyticus
		23308956		transcriptional regulator		RIMD 2210633	
						Corynebacterium	glutamicum
772	FRPANHYK	P96821	P96821	Transcriptional regulator		Corynebacterium	glutamicum
		15607286		Hypothetical 30.6 kDa protein (Transcriptional regulator, TetR family)	TetR	ATCC 13032	
		15839525		hypothetical protein Rv0144		Mycobacterium	tuberculosis
				transcriptional regulator, TetR	TetR	H37Rv	
						Mycobacterium	tuberculosis

773	YRSCNGHK	Q8D480 27367793	Q8D480	family Transcriptional regulator Transcriptional regulator	CDC1551 Vibrio vulnificus Vibrio vulnificus CMCP6
774	LKGGLHHK	Q8G9V2 Q92N23 15966169	Q8G9V2 Q92N23	Hypothetical protein Hypothetical protein R02416 HYPOTHETICAL PROTEIN	Escherichia coli Sinorhizobium meliloti Sinorhizobium meliloti
775	LKGGLHYK	Q55363 16331647	Q55363	Hypothetical protein slr0895 unknown protein	Synechocystis sp. PCC 6803 Synechocystis sp. PCC 6803
776	LKGGLYHK	Q8CUW2 AAP08046 23098450 30019214	Q8CUW2 AAP08046	Transcriptional regulator Transcriptional regulator, family transcriptional regulator	TetR Bacillus cereus ATCC 14579 Oceanobacillus iheyensis HTE831 Bacillus cereus ATCC 14579
777	LKGGFYHK	Q8CQC0 27467136	Q8CQC0	Transcriptional regulator transcriptional regulator	Staphylococcus epidermidis Staphylococcus epidermidis ATCC 12228
778	VKGGYHNK	Q9K8P5 15615520	Q9K8P5 BH2958	Transcriptional regulator (TetR/AcrR family) transcriptional regulator (TetR/AcrR family)	Bacillus halodurans Bacillus halodurans
779	VKGGFHHK	Q8UD08 15889597 17936203 27379338	Q8UD08	Transcriptional regulator, family AGR_C_4219p transcriptional regulator, family transcriptional regulatory protein	TetR Agrobacterium tumefaciens str. C58 Agrobacterium tumefaciens str. C58 (Cereon) Agrobacterium tumefaciens str. C58 (U. Washington) Bradyrhizobium japonicum USDA 110
780	LKGGLYNK	Q98FR8 13473146	Q98FR8	Hypothetical protein mll3646 hypothetical protein	Mesorhizobium loti Mesorhizobium loti
781	LKGGLYSK	Q8YHB1 Q8G0K2 17987174 23501968	Q8YHB1 Q8G0K2	Transcriptional regulator, family TRANSCRIPTIONAL REGULA- TOR, TETR FAMILY transcriptional regulator, family	TETR Brucella melitensis biovar Suis Brucella melitensis 16M Brucella suis 1330
782	FKGGMHQK	Q8XWB4 17547279	Q8XWB4 RSc2560	Putative transcription regulator protein PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	Ralstonia solanacearum Ralstonia solanacearum
783	LKGGMHQK	27376324		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
784	MTVGLHHK	28870342		transcriptional regulator, family	TetR Pseudomonas syringae pv. tomato str. DC3000
785	LQVGLHHK	28870850		transcriptional regulator, family	TetR Pseudomonas syringae pv. tomato str. DC3000
786	VVAGLHHK	28870483		transcriptional regulator, family	TetR Pseudomonas syringae pv. tomato str. DC3000
787	VKSGFAHK	Q8UK32 16119520 17938876	Q8UK32	Transcriptional regulator, family AGR_pAT_419p transcriptional regulator, family	TetR Agrobacterium tumefaciens str. C58 Agrobacterium tumefaciens str. C58 (Cereon) Agrobacterium tumefaciens str. C58 (U. Washington)
788	IKSGFAHK	Q8D4S4 27367595 28899908	Q8D4S4	Transcriptional regulator Transcriptional regulator putative transcription regulator protein	Vibrio vulnificus Vibrio vulnificus CMCP6 Vibrio parahaemolyticus RIMD 2210633
789	IKSGFAHR	Q8Y264 17545191	Q8Y264 RSc0472	Putative transcription regulator protein PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	Ralstonia solanacearum
790	IKSGHKHK	Q8CJS4	Q8CJS4	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
791	IKAGIGHK	BAC68928 29827759	BAC68928	Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces avermitilis Streptomyces avermitilis MA-4680
792	MAMSYHHV	Q984Q9 13476539	Q984Q9	Transcription regulator transcription regulator	Mesorhizobium loti Mesorhizobium loti

793	VAGQHHA	Q8ZPE2 Q8Z730 AAO69122 16764919 16760304 29141931	Q8ZPE2 Q8Z730 AAO69122	Putative transcriptional regulator, TetR family Putative regulatory protein Putative regulatory protein putative transcriptional regulator, TetR family putative regulatory protein putative regulatory protein	Salmonella typhimurium Salmonella typhi Salmonella typhi Salmonella typhimurium LT2 Salmonella enterica subsp. enterica serovar Typhi Salmonella enterica subsp. enterica serovar Typhi Ty2 Streptomyces roseosporus Streptomyces coelicolor Streptomyces avermitilis
794	FRSSYEYR	P72569 O69909 BAC70340 29829171	P72569 O69909 BAC70340	WhiR protein Putative transcriptional regulator Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
795	FRSVYRYV	Q9CHV8 15672592	Q9CHV8 yfgG	Transcriptional regulator transcription regulator	Lactococcus lactis subsp. lactis Lactococcus lactis subsp. lactis
796	FRSVYRYT	28378164		transcription regulator	Lactobacillus plantarum WCFS1
797	LRANYKYK	Q97TH4 15004830	Q97TH4	HTH transcriptional regulator TetR family HTH transcriptional regulator TetR family	Clostridium acetobutylicum Clostridium acetobutylicum
798	ARTNYRYK	Q9RVE8 15806101	Q9RVE8 DR1081	Transcriptional regulator, TETR family transcriptional regulator, TetR family	Deinococcus radiodurans Deinococcus radiodurans
799	MRKSYRYK	26989670		transcriptional regulator, TetR family	Pseudomonas putida KT2440
800	VRSSYRYK	28379358		transcription regulator	Lactobacillus plantarum WCFS1
801	IKSNYRYS	28871411		transcriptional regulator, TetR family	Pseudomonas syringae pv. tomato str. DC3000
802	IRGSYQYL	Q9CHR1 15672643	Q9CHR1 ygfC	Transcriptional regulator transcription regulator	Lactococcus lactis subsp. lactis Lactococcus lactis subsp. lactis
803	ARNSYRYI	O54102	O54102	SC10A5.11 protein	Streptomyces coelicolor
804	PKMLYRYL	Q9A509 16126897	Q9A509	Transcriptional regulator, TetR family transcriptional regulator, TetR family	Caulobacter vibrioides Caulobacter crescentus CB15
805	IKQLYRYL	Q8UFT3 17935214	Q8UFT3	Transcriptional regulator, TetR family transcriptional regulator, TetR family	Agrobacterium tumefaciens str. C58 Agrobacterium tumefaciens str. C58 (U. Washington)
806	VKKTYYRYL	28379556		transcription regulator (putative)	Lactobacillus plantarum WCFS1
807	LRPLYRYG	Q9XAE5 BAC70666 29829497	Q9XAE5 BAC70666	Hypothetical protein SCO5296 Putative TetR-family transcriptional regulator putative TetR-family transcriptional regulator	Streptomyces coelicolor Streptomyces avermitilis Streptomyces avermitilis MA-4680
808	IIGLYKYG	Q8TN01 20091324	Q8TN01	Transcriptional regulator, TetR family transcriptional regulator, TetR family	Methanosarcina acetivorans Methanosarcina acetivorans C2A
809	VRANCRFR	Q92Y53 16263488	Q92Y53	Putative transcriptional activator Putative transcriptional activator	Sinorhizobium meliloti Sinorhizobium meliloti
810	LKSNLRYR	Q9S1Q0	Q9S1Q0	Putative transcriptional regulator	Streptomyces coelicolor
811	LKSALRYR	Q9FBX0	Q9FBX0	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
812	VKQGLRYK	Q8G3K3 23466300	Q8G3K3	Possible TetR-type transcriptional regulator possible TetR-type transcriptional regulator	Bifidobacterium longum Bifidobacterium longum NCC2705
813	VRMAYRNI	Q9CFF7 15673503	Q9CFF7 ypgC	Transcription regulator transcription regulator	Lactococcus lactis subsp. lactis Lactococcus lactis subsp. lactis
814	VRMAYRNL	Q97DH3 15896741	Q97DH3	Transcriptional regulator, AcrR family Transcriptional regulator, AcrR family	Clostridium acetobutylicum Clostridium acetobutylicum

		28379510		family transcription regulator (putative)	Lactobacillus plantarum WCFS1
815	VRAAYRRL	Q8NNG0	Q8NNG0	Transcriptional regulator	Corynebacterium glutamicum
		19553442		transcriptional regulator	Corynebacterium glutamicum ATCC 13032
816	IRTYYRNF	28378377		transcription regulator (putative)	Lactobacillus plantarum WCFS1
817	ADAAFKHM	O67930	O67930	Hypothetical protein AQ_2185	Aquifex aeolicus
		15607117		hypothetical protein	Aquifex aeolicus VF5
818	VDASYTHL	Q9ZBU4	Q9ZBU4	Putative transcriptional regulator	Streptomyces coelicolor
819	VDASYSHV	BAC69809	BAC69809	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29828640		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
820	PTASYSHA	Q9F391	Q9F391	Putative transcriptional regulator	Streptomyces coelicolor
		BAC67791	BAC67791	Putative transcriptional regulator	Streptomyces avermitilis
		29826622		putative transcriptional regulator	Streptomyces avermitilis MA-4680
821	TTMAYTHM	P96941	P96941	Hypothetical 24.7 kDa protein (Transcriptional repressor, putative)	Mycobacterium tuberculosis
		15607793		hypothetical protein Rv0653c	Mycobacterium tuberculosis H37Rv
		15840056		transcriptional repressor, putative	Mycobacterium tuberculosis CDC1551
822	HVATFNRY	P21308	LUXR_VIBHA	Regulatory protein luxR	Vibrio harveyi
		Q8VQC6	Q8VQC6	VanT	Listonella anguillarum
		Q8KX64	Q8KX64	Luminescence regulator LitR	Vibrio fischeri
		O50285	O50285	OpaR	Vibrio parahaemolyticus
		Q9L8G8	Q9L8G8	SmcR (VvpR) (SmcR-like protein)	Vibrio vulnificus
		O30343	O30343	Hemagglutinin/protease regulatory protein HAPR	Vibrio cholerae
		Q8DC19	Q8DC19	SmcR	Vibrio vulnificus
		27364997		SmcR	Vibrio vulnificus CMCP6
		28899290		OpaR	Vibrio parahaemolyticus RIMD 2210633
823	VPRTHNYR	Q8NRW8	Q8NRW8	Transcriptional regulator	Corynebacterium glutamicum
		23308813		transcriptional regulator	Corynebacterium glutamicum ATCC 13032
824	VPRTNMYR	Q9AK09	Q9AK09	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
825	IARTHNYR	Q8FQY2	Q8FQY2	Hypothetical protein	Corynebacterium efficiens
		25027541		hypothetical protein	Corynebacterium efficiens YS-314
826	VVRTFNRY	Q9RD60	Q9RD60	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
827	VVNTYNYK	Q9K4C7	Q9K4C7	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
828	IVNTYNYK	BAC70582	BAC70582	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29829413		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
829	IKKTTNYK	Q9KZX6	Q9KZX6	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
		BAC68464	BAC68464	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29827295		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
830	LPSTINIY	Q8EKA2	Q8EKA2	Transcriptional regulator, TetR family	Shewanella oneidensis
		24371793		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
831	TIAATYYK	P41037	YBIH_ECOLI	Hypothetical transcriptional regulator ybiH	Escherichia coli
		Q93AA6	Q93AA6	Putative transcription regulator YbiH	Shigella flexneri
		Q8ZQN9	Q8ZQN9	Putative transcriptional repressor (TetR/AcrR family)	Yersinia pestis
		Q8Z878	Q8Z878	Hypothetical protein STY0854	Salmonella typhi
		Q8X7Y8	Q8X7Y8	Putative transcriptional regulator	Escherichia coli O157:H7
		Q8FJN5	Q8FJN5	Hypothetical transcriptional regulator ybiH	Escherichia coli O6
		AAP16258	AAP16258	Putative transcriptional regulator	Shigella flexneri 2a str. 2457T

	AAO69689	AAO69689	Hypothetical tetR-family transcriptional regulator	Salmonella typhi	
	16128764	ybiH	putative transcriptional regulator	Escherichia coli K12	
	15800547	ybiH	putative transcriptional regulator	Escherichia coli O157:H7 EDL933	
	15830128		putative transcriptional regulator	Escherichia coli O157:H7	
	16764182	ybiH	putative transcriptional repressor (TetR/AcrR family)	Salmonella typhimurium LT2	
	16759740	ybiH	hypothetical tetR-family transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi	
	24112164	ybiH	putative transcriptional regulator	Shigella flexneri 2a str. 301	
	26246769	ybiH	Hypothetical transcriptional regulator ybiH	Escherichia coli CFT073	
	29142487	ybiH	hypothetical tetR-family transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi Ty2	
	30062281	ybiH	putative transcriptional regulator	Shigella flexneri 2a str. 2457T	
832	TLQAPYYK	Q92V43	Q92V43	Putative transcriptional regulator, TetR family protein	Sinorhizobium meliloti
		16264622		putative transcriptional regulator, TetR family protein	Sinorhizobium meliloti
833	TISLRYYK	Q9KP84	Q9KP84	Transcriptional regulator, TetR family	Vibrio cholerae
		15642485		transcriptional regulator, TetR family	Vibrio cholerae
834	TVAMRYYK	Q8D4H7	Q8D4H7	Transcriptional regulator	Vibrio vulnificus
		27367693		Transcriptional regulator	Vibrio vulnificus CMCP6
835	TAAPRYK	Q8EGA2	Q8EGA2	Transcriptional regulator, TetR family	Shewanella oneidensis
		24373271		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
836	VVAHSYYK	O34970	YTTP_BACSU	Probable transcriptional regulator yttP (Stress response protein yttP)	Bacillus subtilis
		16080015	yttP		Bacillus subtilis subsp. subtilis str. 168
837	VVANSYYK	AAP11559	AAP11559	Transcriptional regulator IcaR	Bacillus cereus ATCC 14579
		30022727		Transcriptional regulator IcaR	Bacillus cereus ATCC 14579
838	VVALSYYK	Q9K800	Q9K800	Hypothetical protein BH3207	Bacillus halodurans
		15615769	BH3207	BH3207~unknown conserved protein	Bacillus halodurans
839	VVLSYYK	Q8EPB1	Q8EPB1	Hypothetical conserved protein	Oceanobacillus iheyensis
		23099656		hypothetical protein	Oceanobacillus iheyensis HTE831
840	IVAMSYYK	AAP26384	AAP26384	Transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		AAP10456	AAP10456	Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
		AAO74585	AAO74585	Hemolysin II regulatory protein	Bacillus cereus
		30262521		transcriptional regulator, TetR family	Bacillus anthracis str. Ames
		30021624		Transcriptional regulator, TetR family	Bacillus cereus ATCC 14579
841	MLSNQYYE	P22645	YDH1_XANAU	Hypothetical 22.8 kDa protein in dhIA 5' region	Xanthobacter autotrophicus
842	MLNLQYYK	Q9KHJ6	Q9KHJ6	Putative regulatory protein EncS	Streptomyces maritimus
843	MLSQSYK	Q97IH3	Q97IH3	Transcriptional regulator, AcrR family	Clostridium acetobutylicum
		15894945		Transcriptional regulator, AcrR family	Clostridium acetobutylicum
844	LLSQHYK	Q8R8Z3	Q8R8Z3	Transcriptional regulator	Thermoanaerobacter tengcongensis
		20808255	AcrR3	Transcriptional regulator	Thermoanaerobacter tengcongensis
845	VTAAHYHR	Q9XDV7	Q9XDV7	DNA-binding protein	Streptomyces griseus
846	VSAAHYHR	Q8CJN5	Q8CJN5	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
847	TTALHYK	Q9AB97	Q9AB97	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16124589		transcriptional regulator, TetR family	Caulobacter crescentus CB15
848	ITAAHYK	Q9HRF3	Q9HRF3	Vng0726c	Halobacterium sp. NRC-1
		15789901	VNG0726C	Vng0726c	Halobacterium sp. NRC-1
849	MAPMHYHR	Q8Y319	Q8Y319	Probable transcription regulator	Ralstonia solanacearum

		17544881	RSc0162	protein PROBABLE TRANSCRIPTION REGULATOR PROTEIN	Ralstonia solanacearum
850	LLAANYHK	Q8KT77 Q9EX90 Q9HZJ9 Q8Y1Z9	Q8KT77 Q9EX90 Q9HZJ9 Q8Y1Z9	PsrA protein PsrA protein Probable transcriptional regulator Putative transcription regulator protein	Pseudomonas chlororaphis Pseudomonas putida Pseudomonas aeruginosa Ralstonia solanacearum
		15598202		probable transcriptional regulator	Pseudomonas aeruginosa PAO1
		26988868		transcriptional regulator, TetR family	Pseudomonas putida KT2440
		17545257	RSc0538	PUTATIVE TRANSCRIPTION REGULATOR PROTEIN	Ralstonia solanacearum
		27380567		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
		28870669	psrA	transcriptional regulator PsrA	Pseudomonas syringae pv. tomato str. DC3000
851	MLAANYHK	Q8G9D4	Q8G9D4	Transcriptional regulator, TetR family (Fragment)	Azoarcus sp. EbN1
852	LIAANYHK	Q8PKZ2	Q8PKZ2	Transcriptional regulator tetR family	Xanthomonas axonopodis pv. citri
		Q8P985	Q8P985	Transcriptional regulator tetR family	Xanthomonas campestris pv. campestris
		21231428	psrA	transcriptional regulator tetR family	Xanthomonas campestris pv. campestris str. ATCC 33913
		21242758		transcriptional regulator tetR family	Xanthomonas axonopodis pv. citri str. 306
853	LLASNYHK	Q56726 Q9KRA1	Q56726 Q9KRA1	ORF (Fragment) Transcriptional regulator, TetR family	Vibrio parahaemolyticus Vibrio cholerae
		Q8EE94	Q8EE94	Transcriptional regulator, TetR family	Shewanella oneidensis
		Q8D8Y3 EAA02414	Q8D8Y3 EAA02414	Transcriptional regulator EbiP2140 (Fragment)	Vibrio vulnificus Anopheles gambiae str. PEST
		15641745		transcriptional regulator, TetR family	Vibrio cholerae
		24374037		transcriptional regulator, TetR family	Shewanella oneidensis MR-1
		27366114 28898003		Transcriptional regulator Na ⁺ /H ⁺ -antiporter protein	Vibrio vulnificus CMCP6 Vibrio parahaemolyticus RIMD 2210633
854	VLASNYHK	Q9A6L3	Q9A6L3	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16126319		transcriptional regulator, TetR family	Caulobacter crescentus CB15
855	SNASNYHK	Q9RF00	Q9RF00	PhlH	Pseudomonas fluorescens
856	SIASNYHK	O07252	O07252	Hypothetical 21.4 kDa protein (Transcriptional regulator, TetR family)	Mycobacterium tuberculosis
		15607469		hypothetical protein Rv0328	Mycobacterium tuberculosis H37Rv
		15839715		transcriptional regulator, TetR family	Mycobacterium tuberculosis CDC1551
857	NLASNYHR	BAC55323	BAC55323	Putative members of the TetR/AcrR family protein	Pseudomonas stutzeri
858	ALASGYHK	Q9Z562	Q9Z562	Putative transcriptional regulator (EbrC repressor)	Streptomyces coelicolor
		BAC70438	BAC70438	Putative TetR-family transcriptional regulator	Streptomyces lividans Streptomyces avermitilis
		29829269		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
859	TLASHYVK	Q9APW2 26989525	Q9APW2	Hypothetical 21.1 kDa protein transcriptional regulator, TetR family	Pseudomonas aeruginosa Pseudomonas putida KT2440
860	LLAAHVCK	Q8GAF9	Q8GAF9	Putative transcriptional regulator	Arthrobacter nicotinovorans
861	NRGALYHR	Q9A363	Q9A363	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16127573		transcriptional regulator, TetR family	Caulobacter crescentus CB15
862	TIGSAYHK	Q98DQ8 Q8YBT2	Q98DQ8 Q8YBT2	Transcriptional regulator Transcriptional regulator, TETR family	Mesorhizobium loti Brucella melitensis
		Q8FWJ1	Q8FWJ1	Transcriptional regulator, TetR	Brucella melitensis biovar

		13473858 17989149		family transcriptional regulator TRANSCRIPTIONAL REGULATOR, TETR FAMILY	Suis Mesorhizobium loti Brucella melitensis 16M
		23500211		transcriptional regulator, TetR family	Brucella suis 1330
863	VPGSTNHK	Q98C04 13474463	Q98C04	Transcriptional regulator	Mesorhizobium loti
864	VQGSTNHK	27376197		transcriptional regulator transcriptional regulatory protein	Mesorhizobium loti Bradyrhizobium japonicum USDA 110
865	VPGSTYHK	27378949		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
866	THGAHYCK	Q8KUH9	Q8KUH9	Transcriptional regulator	Actinosynnema pretiosum subsp. auranticum
867	IKGGQYCK	Q9I461 15596480	Q9I461	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
868	VSRYQQR	Q8GNL7	Q8GNL7	Putative transcriptional regulator	Mycobacterium abscessus
869	MDRYEDR	Q8GNL5	Q8GNL5	Putative transcriptional regulator	Mycobacterium abscessus
870	VERYYESR	O53836 BAC67860	O53836 BAC67860	Hypothetical protein Rv0825c Putative TetR-family transcriptional regulator	Mycobacterium tuberculosis Streptomyces avermitilis
		15607965		hypothetical protein Rv0825c	Mycobacterium tuberculosis H37Rv
		15840238		hypothetical protein	Mycobacterium tuberculosis CDC1551
		29826691		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
871	IVRYYESK	O53681 15607416	O53681 fadD27	Hypothetical protein fadD27 fadD27	Mycobacterium tuberculosis Mycobacterium tuberculosis H37Rv
		15839656		hypothetical protein	Mycobacterium tuberculosis CDC1551
872	VTRQYEEL	Q9RJE7	Q9RJE7	Putative tetR-family transcriptional regulator	Streptomyces coelicolor
		BAC75221	BAC75221	Putative TetR-family transcriptional regulator	Streptomyces avermitilis
		29834052		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
873	LYSNYDGK	Q9PID4 15791735	Q9PID4 Cj0368c	Transcriptional regulatory protein transcriptional regulatory protein	Campylobacter jejuni Campylobacter jejuni subsp. jejuni NCTC 11168
874	IRSRYRLS	Q98816 13475611	Q98816	Transcriptional regulator	Mesorhizobium loti
875	MLKRYQLK	O85695	O85695	transcriptional regulator Putative transcriptional regulator (Putative regulatory protein)	Mesorhizobium loti Streptomyces lividans
		BAC75182	BAC75182	Putative TetR-family transcriptional regulator	Streptomyces coelicolor Streptomyces avermitilis
		29834013		putative TetR-family transcriptional regulator	Streptomyces avermitilis MA-4680
876	AVGSYNLL	Q98ID3 13472230	Q98ID3	Probable transcriptional regulator probable transcriptional regulator	Mesorhizobium loti Mesorhizobium loti
877	VVGSYNEF	27377436		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
878	TNGAYNLV	27382725		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
879	SKGVHRHF	Q9L0M1	Q9L0M1	Hypothetical protein SCO4639	Streptomyces coelicolor
880	SKGLHAHF	Q93J82	Q93J82	Putative TetR-family transcriptional regulator	Streptomyces coelicolor
881	SPQLHYYM	Q9ZFC4	Q9ZFC4	Probable regulatory protein (Fragment)	Mycobacterium sp. FM10
		Q9ZFC1 Q93LY6	Q9ZFC1 Q93LY6	Probable regulatory protein Putative repressor PgaY	Mycobacterium smegmatis Streptomyces sp. PGA64
882	IPDTGHYK	Q93LY6	Q93LY6	Transposon Tn5405 and insertion sequence IS1182 (From Staphylococcus aureus) ORFA and transposase	Staphylococcus aureus
883	VPGQRYYH	Q53772	Q53772		
884	IPSQTYK	27379191		transcriptional regulatory protein	Bradyrhizobium japonicum USDA 110
885	VPGLRHYV	Q9A3S9	Q9A3S9	Transcriptional regulator, TetR family	Caulobacter vibrioides
		16127353		transcriptional regulator, TetR family	Caulobacter crescentus CB15
886	ALGARHYQ	O34381	O34381	Transcriptional regulator	Bacillus subtilis

		16078770	pkxA	transcriptional regulator	<i>Bacillus subtilis</i> subsp. subtilis str. 168
887	LPPLKYYK	BAC68842	BAC68842	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29827673		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
888	LRAAIQRR	Q9EVJ6	Q9EVJ6	Repressor protein MphR(A)	<i>Escherichia coli</i>
889	FRGLTQRK	Q9A5F7	Q9A5F7	Transcriptional regulator, TetR family	<i>Caulobacter vibrioides</i>
		16126732		transcriptional regulator, TetR family	<i>Caulobacter crescentus</i> CB15
890	FCGSLQRK	Q8PNA0	Q8PNA0	Transcriptional regulator uid family	<i>Xanthomonas axonopodis</i> pv. citri
		21241927	uidR	transcriptional regulator uid family	<i>Xanthomonas axonopodis</i> pv. citri str. 306
891	FCGTLQRK	Q8PBP5	Q8PBP5	Transcriptional regulator uid family	<i>Xanthomonas campestris</i> pv. campestris
		21230532	uidR	transcriptional regulator uid family	<i>Xanthomonas campestris</i> pv. campestris str. ATCC 33913
892	FGATVQRK	Q98GD9	Q98GD9	Transcriptional regulator, repressor protein MphR(A)	<i>Mesorhizobium loti</i>
		13472924		transcriptional regulator, repressor protein MphR(A)	<i>Mesorhizobium loti</i>
893	FAATVQRR	Q8U770	Q8U770	Transcriptional regulator, TetR family	<i>Agrobacterium tumefaciens</i> str. C58
		15890404		AGR_L_578p	<i>Agrobacterium tumefaciens</i> str. C58 (Cereon)
		17938273		transcriptional regulator, TetR family	<i>Agrobacterium tumefaciens</i> str. C58 (U. Washington)
894	LLDERRFK	P36656	YJDC_ECOLI	Protein yjdC	<i>Escherichia coli</i> O6
		Q8XDS1	Q8XDS1	Orf, hypothetical protein (YjdC)	<i>Shigella flexneri</i>
		AAP19493	AAP19493	Hypothetical protein yjdC	<i>Escherichia coli</i> O157:H7
		16131960	yjdC	orf, hypothetical protein	<i>Shigella flexneri</i>
		15804727	yjdC	orf, hypothetical protein	<i>Shigella flexneri</i> 2a str. 2457T
		15834370		hypothetical protein	<i>Escherichia coli</i> O157:H7
		24115490	yjdC	orf, conserved hypothetical protein	<i>Shigella flexneri</i> 2a str. 301
		26251025	yjdC	Protein yjdC	<i>Escherichia coli</i> CFT073
		30065510	yjdC	hypothetical protein	<i>Shigella flexneri</i> 2a str. 2457T
895	LLDERLYK	CAD84211	CAD84211	Bacterial regulatory proteins, TetR family	<i>Nitrosomonas europaea</i>
		30248317		Bacterial regulatory proteins, TetR family	<i>Nitrosomonas europaea</i> ATCC 19718
896	LLDTQRFK	Q8ZKC4	Q8ZKC4	Putative bacterial regulatory protein, merR family	<i>Salmonella typhimurium</i>
		Q8Z1A9	Q8Z1A9	Putative transcriptional regulator	<i>Salmonella typhi</i>
		AAO71824	AAO71824	Putative transcriptional regulator	<i>Salmonella typhi</i>
		16767571	yjdC	putative bacterial regulatory protein, merR family	<i>Salmonella typhimurium</i> LT2
		16763143	yjdC	putative transcriptional regulator	<i>Salmonella enterica</i> subsp. enterica serovar Typhi
		29144622	yjdC	putative transcriptional regulator	<i>Salmonella enterica</i> subsp. enterica serovar Typhi Ty2
897	LVSDTRFR	Q8ZIZ3	Q8ZIZ3	Putative TetR-family regulatory protein	<i>Yersinia pestis</i>
		Q8D1E8	Q8D1E8	Hypothetical protein	<i>Yersinia pestis</i>
		16120676		putative TetR-family regulatory protein	<i>Yersinia pestis</i> CO92
		22124512		hypothetical protein	<i>Yersinia pestis</i> KIM
898	LYSARRYH	O53737	O53737	Putative transcriptional regulator	<i>Mycobacterium tuberculosis</i>
		15607593		hypothetical protein Rv0452	<i>Mycobacterium tuberculosis</i> H37Rv
		15839840		conserved hypothetical protein	<i>Mycobacterium tuberculosis</i> CDC1551
899	LHGARRYH	BAC69492	BAC69492	Putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i>
		29828323		putative TetR-family transcriptional regulator	<i>Streptomyces avermitilis</i> MA-4680
900	VMAAGYRR	Q9X940	Q9X940	Putative regulatory protein	<i>Streptomyces coelicolor</i>
901	FMSVLNNI	29347373		putative transcriptional regulator	<i>Bacteroides thetaiotaomicron</i> VPI-5482
902	HLGSTYYI	P75811	YBJK_ECOLI	Hypothetical transcriptional regula-	<i>Escherichia coli</i>

		Q8X6S8	Q8X6S8	tor ybjK Putative DEOR-type transcriptional regulator	Escherichia coli O157:H7
		Q8FJK0 AAN42432 AAP16305	Q8FJK0 AAN42432 AAP16305	Hypothetical protein ybjK Orf, conserved hypothetical protein	Escherichia coli O6 Shigella flexneri
		16128814	b0846	Putative DEOR-type transcriptional regulator	Shigella flexneri 2a str. 2457T
		15800598		putative DEOR-type transcriptional regulator	Escherichia coli K12
		15830180		putative DEOR-type transcriptional regulator	Escherichia coli O157:H7 EDL933
		24112215 26246821 30062328	ybjK ybjK	orf, conserved hypothetical protein Hypothetical protein ybjK	Shigella flexneri 2a str. 301 Escherichia coli CFT073
903	HLGSTYYM	Q8GGH4 Q8ZQK5	Q8GGH4 Q8ZQK5	Putative DEOR type regulator Paral putative regulator (TetR/Acr family)	Shigella flexneri 2a str. 2457T Enterobacter cloacae Salmonella typhimurium
		Q8Z852	Q8Z852	Putative tetR-family transcriptional regulator	Salmonella typhi
		AAO69639	AAO69639	Putative TetR-family transcriptional regulator	Salmonella typhi
		16764231		paral putative regulator (TetR/Acr family)	Salmonella typhimurium LT2
		16759779		putative tetR-family transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi
		29142448		putative TetR-family transcriptional regulator	Salmonella enterica subsp. enterica serovar Typhi Ty2
904	HVGSTYYL	P95092 15610203	P95092	Hypothetical protein Rv3066 hypothetical protein Rv3066	Mycobacterium tuberculosis H37Rv
		15842635		conserved hypothetical protein	Mycobacterium tuberculosis CDC1551
905	HLGSTYHI	Q93J02	Q93J02	Putative tetR-family transcriptional regulatory protein	Streptomyces coelicolor
906	HLGSTYHL	Q9X9V5	Q9X9V5	Hypothetical 21.1 kDa protein (EbrA repressor)	Streptomyces coelicolor
		BAC74047 29832878	BAC74047	Hypothetical protein hypothetical protein	Streptomyces lividans Streptomyces avermitilis Streptomyces avermitilis MA-4680
		21220404	SCO1917	conserved hypothetical protein	Streptomyces coelicolor A3(2)
907	HLASTYYL	P71830	P71830	Hypothetical transcriptional regulator CY369.20 hypothetical protein Rv0775	Mycobacterium tuberculosis H37Rv
		15607915		conserved hypothetical protein	Mycobacterium tuberculosis CDC1551
908	HLSATYYI	Q9HXU1 15598895	Q9HXU1	Probable transcriptional regulator probable transcriptional regulator	Pseudomonas aeruginosa Pseudomonas aeruginosa PAO1
		26988229		transcriptional regulator, TetR family	Pseudomonas putida KT2440
		28868708		transcriptional regulator, TetR family	Pseudomonas syringae pv. tomato str. DC3000
909	FLGSTYYM	Q9F2S8	Q9F2S8	Putative tetR-family transcriptional regulatory protein	Streptomyces coelicolor
910	FAGTSNYR	Q8KUI4	Q8KUI4	Transcriptional regulator	Actinosynnema pretiosum subsp. auranticum
911	FPGTSNYR	Q9KXT6	Q9KXT6	Hypothetical protein SCO4305	Streptomyces coelicolor
912	HPGTKNYR	BAC21010	BAC21010	Hypothetical protein	Streptomyces griseus
913	AADTALHH	Q93M20	Q93M20	Putative transcriptional regulator	Streptomyces aureofaciens
914	ISAPYVNV	Q9KFZ9 15612880	Q9KFZ9 BH0317	Hypothetical protein BH0317 BH0317~unknown	Bacillus halodurans
915	AVAPYSVL	Q8RI13	Q8RI13	Transcriptional regulator, TetR family	Fusobacterium nucleatum subsp. nucleatum
		EAA24512	EAA24512	Transcriptional regulator, TetR family	Fusobacterium nucleatum subsp. vincentii ATCC 49256
		19705128		Transcriptional regulator, TetR family	Fusobacterium nucleatum subsp. nucleatum ATCC 25586
916	APAPYKSM	Q8RET6	Q8RET6	Transcriptional regulator, TetR	Fusobacterium nucleatum

		19704339		family Transcriptional regulator, TetR family		subsp. nucleatum Fusobacterium nucleatum subsp. nucleatum ATCC 25586
917	TQPYFRLK	Q9L0M0	Q9L0M0	Putative tetR-family transcriptional regulator		Streptomyces coelicolor
		BAC72615	BAC72615	Putative TetR-family transcriptional regulator		Streptomyces avermitilis
		29831446		putative TetR-family transcriptional regulator		Streptomyces avermitilis MA-4680
918	TQPYFRFK	Q8EL51	Q8EL51	Transcriptional regulator (TetR-family)		Oceanobacillus iheyensis
		23100835		transcriptional regulator		Oceanobacillus iheyensis HTE831
919	ATQPYLEM	AAO81319	AAO81319	Transcriptional regulator, TetR family		Enterococcus faecalis
		AAO81935	AAO81935	Transcriptional regulator, TetR family		Enterococcus faecalis
		29376095		transcriptional regulator, TetR family		Enterococcus faecalis V583
		29376711		transcriptional regulator, TetR family		Enterococcus faecalis V583
920	IKPNLY..	Q8YF80	Q8YF80	Transcriptional regulator, TETR family		Brucella melitensis
		17987925		TRANSCRIPTIONAL REGULATOR, TETR FAMILY		Brucella melitensis 16M

Capítulo 3:

Detección de los factores sigma presentes en el genoma de *Pseudomonas putida* KT2440.

7.1. Resumen:

Pseudomonas putida KT2440 es un colonizador eficiente de una amplia variedad de hábitats, incluyendo nichos acuáticos, edáficos y la rizosfera de plantas de interés agrícola. De acuerdo con esta capacidad y con la necesidad de adaptarse a múltiples condiciones ambientales, *P. putida* ha desarrollado mecanismos sofisticados de regulación transcripcional.

Se ha analizado a nivel genómico el repertorio de factores sigma en *P. putida* KT2440 habiéndose identificado 24 de ellos, 19 de los cuales pertenecen a la subfamilia de “factores sigma de función extracitoplasmática” (ECF). Trece de los ECF presentaron “alta” similitud con el factor sigma FecI de *Escherichia coli*, que está implicado en la regulación de la respuesta celular a la adquisición de hierro. En 11 casos, se detectó un gen similar a *fecR* adyacente al gen del tipo *fecI*, y en 10 casos, se detectó, vecino a la agrupación *fecI/fecR*, un gen que codificaría un receptor para sideróforos de hierro. Esto podría explicar la capacidad de *P. putida* KT2440 para crecer en condiciones de baja disponibilidad de hierro. Cinco de las diez agrupaciones de genes *fecI/fecR*/receptor de hierro de *P. putida* presentaron un cluster homólogo en el genoma del patógeno humano *Pseudomonas aeruginosa*.

7.2. Introducción:

Las bacterias del género *Pseudomonas* colonizan una variedad de hábitats y degradan un conjunto amplio de compuestos orgánicos e inorgánicos tanto naturales como xenobióticos, y de síntesis industrial (Steffan *et al.*, 1994; Ramos *et al.*, 1997; Marqués *et al.*, 1999a).

Las secuencias de ADN de los genomas de *Pseudomonas aeruginosa* PAO1 y *Pseudomonas putida* KT2440 se han publicado recientemente (Stover *et al.*, 2000; Nelson *et al.*, 2002), y se espera que su análisis proporcione datos que ayuden a descifrar sus estilos de vida. Estas bacterias exhiben un elevado número de marcos de lectura abierta (ORF, del inglés *open reading frames*) que codifican reguladores transcripcionales, así en *P. aeruginosa* y en *P. putida* se han identificado alrededor de 450 reguladores frente a sólo 250 reguladores transcripcionales en *Escherichia coli* y los sólo 4 del parásito intracelular *Chlamydia muridarum*. Las proteínas reguladoras identificadas en el género *Pseudomonas* pertenecen a múltiples familias de reguladores transcripcionales que controlan de manera específica la expresión de otros genes.

Como se ha descrito en la Introducción General, la transcripción en procariotas está mediada por la ARN polimerasa, a la que los factores sigma confieren la especificidad en el reconocimiento de promotores (Ishihama, 2000).

En *P. putida* los estudios sobre factores sigmas son relativamente escasos. Los primeros estudios identificaron el factor sigma primario σ^{70} , como implicado en la transcripción de los genes esenciales y de funciones fisiológicas básicas, así como su implicación en el control de genes regulados positiva o negativamente (Mermod *et al.*, 1984; Dominguez y Márques, 2004).

El primer factor sigma alternativo caracterizado en *P. putida* fue σ^{38} , también conocido como RpoS. El gen *rpoS* de *P. putida* se clonó de un banco de genes de *P. putida* KT2440 tras complementar una cepa de *E. coli* deficiente en *rpoS* (Ramos-González y Molin, 1998). Este estudio reveló que el gen *rpoS* se expresaba a un nivel relativamente constante a lo largo de la curva de crecimiento de *P. putida*, y que mutantes *rpoS* de *P. putida* mostraban una reducción en la supervivencia bajo condiciones de deficiencia de carbono o choques de etanol. Un análisis del transcriptoma del mutante RpoS de *P. putida* en geles bidimensionales reveló que RpoS controlaba la expresión de al menos 50 productos génicos bajo condiciones de deficiencia de carbono.

Manzanera *et al.* (2001) identificaron RpoH (ó σ^{32}) en *P. putida*. La proteína RpoH se encontró primero en *E. coli* como el factor sigma alternativo que controlaba el regulón de choque térmico de este microorganismo (Grossman *et al.*, 1987). En *P. putida*, el gen *rpoH* se expresa desde diferentes promotores bajo distintas condiciones ambientales (Manzanera *et al.*, 2001). Sin embargo, no se ha logrado obtener mutantes en este gen en *P. putida* y, por lo tanto, el papel específico de los genes cuya expresión está mediada por la RNA polimerasa con RpoH en este microorganismo permanece aún por elucidar. Marqués *et al.* (1999b) publicaron que la transcripción desde el promotor Pm, de la ruta de fisión *meta* del plásmido TOL, era inicialmente mediada por RpoH en respuesta a la activación del sistema de respuesta a choque térmico por 3-metilbenzoato, el efector de la proteína XylS que controla positivamente la expresión desde Pm.

En 1989, σ^{54} (también llamada RpoN) se clonó a partir de una genoteca de *P. putida* mediante hibridación cruzada con una sonda del correspondiente gen de *E. coli* (Inouye *et al.*, 1989). El gen que codifica σ^{54} , *rpoN*, está presente en una sola copia en el genoma de *P. putida* en una agrupación con los marcos de lectura abierta *orf102*, *orf28* y los genes *ptsN* y *ptsO* (Cases y Lorenzo, 2000). Köhler *et al.* (1989) revelaron que una inserción del transposón Tn5 en el gen *rpoN* resultaba en mutantes que fueron incapaces de usar algunas fuentes de C y N. También se ha descrito que el factor sigma

RpoN se requiere para la expresión de tanto genes cromosómicos como plasmídicos tales como los del metabolismo del *m*-xileno, del plásmido TOL pWW0 (Revisado por Ramos *et al.*, 1997; Marqués *et al.*, 1999a; y Ruíz *et al.*, 2004). En *P. putida*, uno de los promotores dependientes de σ^{54} mejor caracterizados, es Pu, aunque su expresión se modula también por fuente de carbono (Duetz *et al.*, 1994; Holtel *et al.*, 1994; Cases and Lorenzo, 2000) o un crecimiento rápido en medio rico, previne la expresión desde Pu (Cases *et al.*, 1999; Carmona *et al.*, 2000).

El factor σ^{28} también pertenece a la familia σ^{70} (también conocido como FliA o SigD), y se identificó en *P. putida* PRS2000 dentro de una región de genes de quimiotaxis (Ditty *et al.*, 1998). Su organización es similar en *P. putida* KT2440 y *P. aeruginosa* PA01. En *P. aeruginosa*, FliA participa en el control de un sistema de cascada de la síntesis del flagelo. El nivel superior de control transcripcional en la síntesis del flagelo está mediada por RpoN y el regulador transcripcional positivo FleQ (Totten *et al.*, 1990; Arora *et al.*, 1997). Este sistema controla al sistema de dos componentes FleSR, el cual controla la síntesis de flagelina. FliA de *P. aeruginosa* se regula por el factor antisigma FlgM (Frisk *et al.*, 2002). En el genoma de *P. putida* KT2440 se ha identificado FlgM una proteína muy similar a su correspondiente factor antisigma. Recientemente Segura y colaboradores han construido un mutante deficiente en la síntesis de FliA que no presenta motilidad ni respuesta quimiotáctica (Ana Segura, datos sin publicar).

7.3. Resultados y discusión

7.3.1. Identificación de los factores sigma del genoma de *P. putida* KT2440.

Para identificar todos los factores sigma presentes en el genoma de *P. putida* KT2440, se realizó una búsqueda de los ORF traducidos que presentasen similitud con los 168 factores sigma de bacterias disponibles en la fecha del análisis en la base de datos SWISS-PROT (Julio, 2002).

La búsqueda mediante BLAST de cada una de las 168 secuencias contra el genoma de *P. putida* KT2440 (Julio, 2002), proporcionó un conjunto de 613 puntuaciones superiores al umbral que se estableció como significativo $E < 10^{-4}$. Estos resultados se analizaron con herramientas específicas que desarrollamos para esta tarea, lo que nos permitió identificar 24 ORFs que podrían codificar potenciales factores sigmas en *P. putida* KT2440.

Una búsqueda mediante BLAST de las 24 secuencias de *P. aeruginosa* anotadas como factores sigma hipotéticos, contra el genoma de *P. putida* KT2440, no proporcionó ningún factor sigma adicional en este último microorganismo.

Las tablas 33 y 34 muestran los identificadores y anotaciones de los 24 factores sigma hipotéticos detectados en el genoma de *P. putida* y *P. aeruginosa*, respectivamente.

Hay que señalar que de los 49 genomas anotados y disponibles en la fecha del análisis, *P. putida* y *P. aeruginosa* (cada uno con 24 factores sigma hipotéticos) junto con *Mesorhizobium loti* (23 factores sigma hipotéticos), eran los 3 microorganismos, secuenciados, con mayor número de factores sigma (Tabla 35). En una actualización a día de hoy sólo *Streptomyces coelicolor* presenta más factores sigma (63) que los microorganismos arriba seleccionados (Agosto, 2005).

Tabla 35: Factores sigma en genomas los bacterianos anotados.

Número de factores sigma en los genomas bacterianos anotados			
<i>Mycoplasma genitalium</i>	1	<i>Lactococcus lactis</i>	5
<i>Mycoplasma pneumonia</i>	1	<i>Listeria innocua</i>	5
<i>Mycoplasma pulmonis</i>	1	<i>Listeria monocytogenes</i>	5
<i>Streptococcus pneumonia</i>	1	<i>Treponema palidum</i>	5
<i>Ureaplasma urealyticum</i>	1	<i>Salmonella enterica typhi</i>	5
<i>Buchnera sp APS</i>	2	<i>Salmonella typhimurium</i>	6
<i>Deinococcus radiodurans</i>	2	<i>Agrobacterium tumefaciens</i>	7
<i>Rickettsia prowazekii</i>	2	<i>Corynebacterium glutamicum</i>	7
<i>Staphylococcus aureus</i>	2	<i>Escherichia coli</i>	7
<i>Streptococcus pyogenes</i>	2	<i>Yersinia pestis</i>	7
<i>Borrelia burgdorferi</i>	3	<i>Shinorhizobium meliloti</i>	8
<i>Campylobacter jejuni</i>	3	<i>Synechocystis</i>	8
<i>Chlamydia muridarum</i>	3	<i>Vibrio cholerae</i>	8
<i>Chlamydia trachomatis</i>	3	<i>Nostoc</i>	9
<i>Chlamydia pneuminae</i>	3	<i>Clostridium perfringens</i>	11
<i>Helicobacter pylori</i>	3	<i>Ralstonia solanacearum</i>	13
<i>Rickettsia conorii</i>	3	<i>Mycobacterium tuberculosis</i>	14
<i>Aquifex aeolicus</i>	4	<i>Clostridium acetobutylicum</i>	17
<i>Fusobacterium nucleatum</i>	4	<i>Caulobacter crescentus</i>	17
<i>Haemophilus influenzae</i>	4	<i>Bacillus subtilis</i>	18
<i>Mycobacterium leprae</i>	4	<i>Bacillus halodurans</i>	19
<i>Neisseria meningitidis</i>	4	<i>Mesorhizobium loti</i>	23
<i>Pasteurella multocida</i>	4	<i>Pseudomonas aeruginosa</i>	24
<i>Thermotoga maritima</i>	4	<i>Pseudomonas putida</i>	24
<i>Xilella fastidiosa</i>	4		

En los genomas anotados disponibles en el NCBI <http://ncbi.nlm.nih.gov/PMGifs/Genomes/micr.jhtml>, (septiembre 2002) incluimos todos los marcos de lectura abierta que aparecían anotados como factores sigma.

Tabla 33: Factores sigma codificados por el genoma de *P. putida* KT2440.

Nombre del factor σ	Id ORF	Mejor BLAST SWPROT/TrEMBL	E	Organismo	Mejor BLAST con anotación funcional	E	Organismo		
Pp-RpoH or Pp- σ 32	PP5108	σ 32 or RpoH	Q9KJX7	1×10^{-160}	<i>Pseudomonas putida</i>	σ 32 or RpoH	Q9KJX7	1×10^{-160}	<i>Pseudomonas putida</i>
Pp-RpoS or Pp- σ 38	PP1623	σ 38 or RpoS	Q9R9V7	0.0	<i>Pseudomonas putida</i>	σ 38 or RpoS	Q9R9V7	0.0	<i>Pseudomonas putida</i>
Pp-FliA or Pp- σ 28	PP4341	σ 28 or FliA	O52258	1×10^{-133}	<i>Pseudomonas putida</i>	σ 28 or FliA	O52258	1×10^{-133}	<i>Pseudomonas putida</i>
Pp-RpoN or Pp- σ 54	PP0952	σ 54 or RpoN	P15591	0.0	<i>Pseudomonas putida</i>	σ 54 or RpoN	P15591	0.0	<i>Pseudomonas putida</i>
Pp-RpoD or Pp- σ 70	PP0387	σ 70 or RpoD	O54540	0.0	<i>Pseudomonas putida</i>	σ -70 or RpoD	O54540	0.0	<i>Pseudomonas putida</i>
Pp-ECF-1	PP2088	SigX (Pot. RNAP σ)	Q9R6S3	1×10^{-100}	<i>Pseudomonas fluorescens</i>	SigX (Potencial RNAP σ)	Q9R6S3	1×10^{-100}	<i>Pseudomonas fluorescens</i>
Pp-ECF-2	PP4611	Potencial RNAP σ	Q9HXB4	2×10^{-54}	<i>Pseudomonas aeruginosa</i>	PupI (Potencial RNAP σ)	Q52208	3×10^{-46}	<i>P. putida</i> WCS358
Pp-ECF-3	PP1427	AlgU (Potenc. RNAP σ)	Q93D72	1×10^{-105}	<i>Pseudomonas fluorescens</i>	AlgU (Potencial RNAP σ)	Q93D72	1×10^{-105}	<i>Pseudomonas fluorescens</i>
Pp-ECF-4	PP4244	PfI (Potencial RNAP σ)	Q52157	5×10^{-86}	<i>Pseudomonas putida</i>	PfI (Potencial RNAP σ)	Q52157	5×10^{-86}	<i>P. putida</i> WCS358
Pp-ECF-5	PP4608	Potencial RNAP σ	Q9I6Y3	1×10^{-53}	<i>Pseudomonas aeruginosa</i>	BupI (Potencial ECF σ)	Q9L430	7×10^{-16}	<i>Bordetella bronchiseptica</i>
Pp-ECF-6	PP4208	Potencial RNAP σ	Q9I193	1×10^{-54}	<i>Pseudomonas aeruginosa</i>	BupI (Potencial ECF σ)	Q9L430	5×10^{-17}	<i>Bordetella bronchiseptica</i>
Pp-ECF-7	PP0352	Potencial RNAP σ	Q9I646	5×10^{-72}	<i>Pseudomonas aeruginosa</i>	FiuI (Potencial RNAP σ)	O68593	1×10^{-66}	<i>Pseudomonas aeruginosa</i>
Pp-ECF-8	PP0162	Potencial RNAP σ	Q9I114	5×10^{-64}	<i>Pseudomonas aeruginosa</i>	PupI (Potencial RNAP σ)	Q52208	3×10^{-39}	<i>P. putida</i> WCS358
Pp-ECF-9	PP3086	Potencial RNAP σ	Q9I2J2	2×10^{-44}	<i>Pseudomonas aeruginosa</i>	PupI (Potencial RNAP σ)	Q52208	3×10^{-33}	<i>P. putida</i> WCS358
Pp-ECF-10	PP4553	Potencial RNAP σ	Q9I3Z3	1×10^{-148}	<i>Pseudomonas aeruginosa</i>				
Pp-ECF-11	PP2888	PrtI (Potenc. RNAP σ)	Q9KIM9	1×10^{-41}	<i>Pseudomonas fluorescens</i>	PrtI (Potencial RNAP σ)	Q9KIM9	1×10^{-41}	<i>Pseudomonas fluorescens</i>
Pp-ECF-12	PP3006	ECF family σ factor	Q8UI26	7×10^{-38}	<i>Agrobacterium tumefaciens</i>	Prob σ -24 or σ -E homolog	Q8XV48	1×10^{-23}	<i>Ralstonia solanacearum</i>
Pp-ECF-13	PP0994	Potencial RNAP σ	Q9A3F1	1×10^{-30}	<i>Caulobacter crescentus</i>	SigD (Familia ECF)	Q8UD59	6×10^{-28}	<i>Agrobacterium tumefaciens</i>
Pp-ECF-14	PP1008	Potencial RNAP σ	Q9I6Y3	3×10^{-19}	<i>Pseudomonas aeruginosa</i>	PvdS	Q8RQA9	6×10^{-11}	<i>Pseudomonas fluorescens</i>
Pp-ECF-15	PP0667	Potencial RNAP σ	Q9I2J2	2×10^{-34}	<i>Pseudomonas aeruginosa</i>	PupI (Potencial RNAP σ)	Q52208	4×10^{-30}	<i>P. putida</i> WCS358
Pp-ECF-16	PP0704	Potencial RNAP σ	Q9I114	2×10^{-43}	<i>Pseudomonas aeruginosa</i>	PupI (Potencial RNAP σ)	Q52208	4×10^{-41}	<i>P. putida</i> WCS358
Pp-ECF-17	PP3577	Potencial RNAP σ	Q9I2J2	1×10^{-43}	<i>Pseudomonas aeruginosa</i>	Fecl	P23484	6×10^{-41}	<i>Escherichia coli</i>
Pp-ECF-18	PP0865	Potencial RNAP σ	Q9I193	9×10^{-25}	<i>Pseudomonas aeruginosa</i>	PrhI σ factor	Q93JM8	1×10^{-16}	<i>Ralstonia solanacearum</i>
Pp-ECF-19	PP2192	Potencial RNAP σ	Q9I193	9×10^{-24}	<i>Pseudomonas aeruginosa</i>	Fecl	P23484	4×10^{-15}	<i>Escherichia coli</i>

Tabla 34: Factores sigma codificados por el genoma de *P. aeruginosa* PA01.

Nombre del factor σ	Id ORF	Mejor BLAST SWP/TrEMBL	E	Organismo	Mejor BLAST con anotación funcional	E	Organismo		
Pa-RpoH or Pa- σ 32	PA0376	σ 32 or RpoH	P42378	1×10^{-159}	<i>P. aeruginosa</i>	σ 32 or RpoH	P42378	1×10^{-159}	<i>P. aeruginosa</i>
Pa-RpoS or Pa- σ 38	PA3622	σ 38 or RpoS	P45684	0.0	<i>P. aeruginosa</i>	σ 38 or RpoS	P45684	0.0	<i>P. aeruginosa</i>
Pa-FliA or Pa- σ 28	PA1455	σ 28 or FliA	P29248	1×10^{-136}	<i>P. aeruginosa</i>	σ 28 or FliA	P29248	1×10^{-136}	<i>P. aeruginosa</i>
Pa-RpoN or Pa- σ 54	PA4462	σ 54 or RpoN	P49988	0.0	<i>P. aeruginosa</i>	σ 54 or RpoN	P49988	0.0	<i>P. aeruginosa</i>
Pa-RpoD or Pa- σ 70	PA0576	σ 70 or RpoD	P26480	0.0	<i>P. aeruginosa</i>	σ 70 or RpoD	P26480	0.0	<i>P. aeruginosa</i>
Pa-ECF-1	PA1776	Potencial RNAP σ	Q9I2W5	1×10^{-91}	<i>P. aeruginosa</i>	SigX (Potencial RNAP σ)	Q9ZIG9	1×10^{-90}	<i>P. aeruginosa</i>
Pa-ECF-2	PA3899	Potencial RNAP σ	Q9HXB4	8×10^{-92}	<i>P. aeruginosa</i>	PupI	Q52208	1×10^{-45}	<i>Pseudomonas putida</i>
Pa-ECF-3	PA0762	AlgU or σ 30 or σ H	Q06198	1×10^{-105}	<i>P. aeruginosa</i>	AlgU or σ 30 or σ H	Q06198	1×10^{-105}	<i>P. aeruginosa</i>
Pa-ECF-4	PA2426	PvdS (Potencial RNAP σ)	P95425	1×10^{-102}	<i>P. aeruginosa</i>	PvdS (Potencial RNAP σ)	P95425	1×10^{-102}	<i>Pseudomonas aeruginosa</i>
Pa-ECF-5	PA0149	Potencial RNAP σ	Q9I6Y3	1×10^{-99}	<i>P. aeruginosa</i>	RpoE	Q8VTB1	2×10^{-13}	<i>Bacteroides fragilis</i>
Pa-ECF-6	PA2387	Potencial RNAP σ	Q9I193	1×10^{-85}	<i>P. aeruginosa</i>	BupI (Potencial ECF σ)	Q9L430	6×10^{-21}	<i>Bordetella bronchiseptica</i>
Pa-ECF-7	PA0472	Potencial RNAP σ	Q9I646	1×10^{-93}	<i>P. aeruginosa</i>	FiuI (Potencial RNAP σ)	O68593	3×10^{-88}	<i>Pseudomonas aeruginosa</i>
Pa-ECF-8	PA2468	Potencial RNAP σ	Q9I114	6×10^{-95}	<i>P. aeruginosa</i>	FecI	P23484	5×10^{-39}	<i>Escherichia coli</i>
Pa-ECF-9	PA1912	Potencial RNAP σ	Q9I2J2	4×10^{-92}	<i>P. aeruginosa</i>	FecI	Q8P4W3	8×10^{-40}	<i>Xanthomonas campestris</i>
Pa-ECF-10	PA1351	Potencial RNAP σ	Q9I3Z3	0.0	<i>P. aeruginosa</i>	SigI (Potencial RNAP σ)	O50445	2×10^{-4}	<i>Mycobacterium tuberculosis</i>
Pa-ECF-11	PA0675	Potencial RNAP σ	Q9I5P9	4×10^{-97}	<i>P. aeruginosa</i>	PigD (Potencial RNAP σ)	O69005	4×10^{-95}	<i>P. aeruginosa</i>
Pa-ECF-12	PA1300	Potencial RNAP σ	Q9I444	3×10^{-97}	<i>P. aeruginosa</i>	FecI	P23484	4×10^{-18}	<i>Escherichia coli</i>
Pa-ECF-13	PA1363	Potencial ECF σ	Q9I3Y1	1×10^{-140}	<i>P. aeruginosa</i>	PigD (Potencial RNAP σ)	O69005	3×10^{-14}	<i>P. aeruginosa</i>
Pa-ECF-14	PA2050	Potencial RNAP σ	Q9I265	5×10^{-91}	<i>P. aeruginosa</i>	PrhI, sistema de transducción de señal	Q93JM8	1×10^{-38}	<i>Ralstonia solanacearum</i>
Pa-ECF-15	PA2093	Potencial ECF σ	Q9I222	9×10^{-94}	<i>P. aeruginosa</i>	PrhI, sistema de transducción de señal	Q93JM8	4×10^{-37}	<i>Ralstonia solanacearum</i>
Pa-ECF-16	PA2896	Potencial RNAP σ	Q9HZV1	1×10^{-107}	<i>P. aeruginosa</i>	σ E	Q98CJ9	4×10^{-17}	<i>Rhizobium loti</i>
Pa-ECF-17	PA3285	Potencial RNAP σ	Q9HYV8	1×10^{-113}	<i>P. aeruginosa</i>	σ E	P34086	3×10^{-18}	<i>Escherichia coli</i>
Pa-ECF-18	PA3410	Potencial RNAP σ (HasI)	Q9HYJ5	3×10^{-92}	<i>P. aeruginosa</i>	RpoE	Q8VTB1	1×10^{-15}	<i>Bacteroides fragilis</i>
Pa-ECF-19	PA4896	Potencial ECF σ	Q9HUR7	8×10^{-99}	<i>P. aeruginosa</i>	PupI	Q52208	2×10^{-30}	<i>Pseudomonas putida</i>

Los números de identificación de los marcos de lectura abierta (*orf*) están disponibles en <http://www.tigr.org> (segunda columna). Las tres columnas centrales muestran el mejor resultado de BLAST con las proteínas depositadas en SWISSPROT/TrEMBL junto con el número de acceso, el valor de E y el organismo del que procede. Las columnas de la derecha muestran el mejor resultado de BLAST con anotaciones funcionales.

7.3.2. Diecinueve factores sigma de *P. putida* KT2440 pertenecen a la subfamilia ECF.

Los análisis realizados identificaron 19 nuevos factores sigma en *P. putida* KT2440 (Tabla 33) además de los 5 descritos en la *Introducción*. Todos los nuevos factores sigma pertenecían a la subfamilia ECF dentro de la familia σ^{70} . Estos factores sigma se denominaron genéricamente Pp-ECF-1 hasta Pp-ECF-19. Trece de estos factores sigma ECF tienen la mejor puntuación en BLAST con proteínas de *P. aeruginosa*, tres de ellas con proteínas de *P. fluorescens*, una con una proteína de *Agrobacterium tumefaciens*, otra con una proteína de *Caulobacter crescentus* y la mejor puntuación para Pp-ECF-4 fue PfrI de la cepa WCS358 de *P. putida*.

Pp-ECF-1 mostró alta identidad con SigX de *P. fluorescens*. El gen que codifica Pp-ECF-1 se localizó en 5' con respecto al gen que codifica la proteína de membrana externa OprF. En *P. fluorescens*, la disrupción de *sigX* reduce significativamente la expresión de OprF y además tiene un marcado efecto sobre la tasa de crecimiento en medio rico, indicando que también interviene en el control de otros genes (Brinkman *et al.*, 1999).

La proteína Pp-ECF-3 presenta un alto grado de similitud con AlgU de *P. aeruginosa* ($E = 10^{-105}$) (Tabla 33). AlgU fue inicialmente identificada por DeVries y Ohman (1994) como implicada en el fenotipo mucoso de las cepas de *P. aeruginosa* aisladas de pacientes con fibrosis quística. Más tarde, se relacionó a este factor sigma con la tolerancia a desecación y estrés osmótico en la cepa de control biológico *P. fluorescens* CHA0 (Schnider-Keel *et al.* 2002). Resultados similares se han descrito en *P. syringae* (Keith y Bender, 1999). Garret *et al.* (1999) también describieron que AlgU controlaba negativamente la expresión de la biosíntesis del flagelo en *P. aeruginosa*. En relación con este aspecto Segura *et al.* (2001) y Kieboom *et al.* (2001) describieron mutantes con la capacidad de síntesis del flagelo afectada que eran hipersensibles a exposición a disolventes, pero se desconoce si AlgU está implicado en esta respuesta, pues no se han construido y caracterizado mutantes de AlgU en *P. putida*. No obstante en un reciente análisis proteómico de respuesta a estrés por tolueno, Segura *et al.* (2005), han identificado la expresión de proteínas de respuesta a estrés térmico, por lo que es posible que AlgU, directa o indirectamente, esté implicado en dicho proceso.

Pp-ECF-11 es similar al factor sigma PrtI de *P. fluorescens* (Tabla 33), codificado por el operón dicistrónico *prtIR*. Este factor sigma parece controlar la secreción de una proteasa extracelular en la cepa psicrófila (tolerante al frío) de *P. fluorescens* LS107d2. PrtR es un activador transmembrana de PrtI, y ambos son necesarios para la expresión de *aprX* a 29°C, el gen que codifica la proteasa antes mencionada (Burguer *et al.*, 2000). Aunque la organización genética de un operón similar al *prtIR* está presente en *P. putida* KT2440, no hemos detectado ORFs que codifiquen proteasas con secuencias similares a AprX. Además hasta la fecha no hay evidencias de secreción de proteasas al medio extracelular en *P. putida*.

Pp-ECF-12 muestra similitud al factor sigma-E (sigma 24), que normalmente controla los genes que participan en la resistencia a temperaturas elevadas y estrés oxidativo (Mecas *et al.*, 1993; Maeda *et al.*, 2000) Recientemente Estrella Duque ha demostrado que este factor sigma participa directamente en la tolerancia a tolueno y se ha denominado RpoT (E. Duque, comunicación personal).

Pp-ECF-13 muestra relación con SigD (sigma 28) que regula la síntesis de flagelos en *B. subtilis*.

7.3.3. Análisis de los loci del genoma de *P. putida* KT2440 que codifican supuestos factores sigma-ECF con similitud de secuencia a *fecI* de *E. coli*.

El análisis del genoma de *P. putida* KT2440 indicaba que había 13 factores sigma ECF que mostraban similitud con el gen *fecI* de *E. coli*: Pp-ECF-2, -4, -5, -6, -7, -8, -9, -14, -15, -16, -17, -18 y Pp-ECF-19.

La proteína FecI en *E. coli* es un factor sigma que regula la captación de citrato de hierro (Enz *et al.*, 2000). Las proteínas codificadas por los genes del operón *fecABCDE*, transportan el citrato de hierro (III) extracelular al citoplasma. Los genes *fecI/fecR*, localizados en el extremo 5' de *fecA*, activan la transcripción del operón. FecA es el receptor que interacciona con el citrato férrico extracelular transmitiendo una señal al sensor transmembrana FecR. Los dominios citoplasmáticos de FecR interaccionan con el factor sigma FecI, que a su vez estimula la expresión del operón *fecABCDE*. En *E. coli* el operón *fecIR* está regulado negativamente por la proteína Fur, que sensa el nivel de hierro intracelular (Mahren *et al.*, 2002).

En el alineamiento múltiple (CLUSTALW) de todos los factores sigma detectados en *P. putida*, el dendrograma reveló que todos los factores sigma-ECF similares a FecI, se reunían en un solo grupo, dividido en dos ramas (Figura 23). En la Tabla 36 se

muestra la cuantificación de la similitud entre todas las proteínas de esta agrupación y la proteína FecI de *E. coli* y diferencia también los 2 mismos subgrupos que el dendrograma. Así el primer subgrupo, el de secuencias con valores de similitud con FecI mayores (entre 39 y 48%), corresponde a la primera rama, mientras que el segundo subgrupo, el de las secuencias con valores de similitud con FecI menores (entre 16 y 26%), corresponde a la segunda rama del dendrograma (Figura 23).

El grado de similitud entre las proteínas de *P. putida* similares a FecI indica que la divergencia no fue muy reciente y sugiere un posible origen evolutivo que podría implicar transferencia genética horizontal desde diferentes microorganismos con una rápida evolución dirigida por restricciones funcionales derivadas de requerimientos adaptativos ambientales.

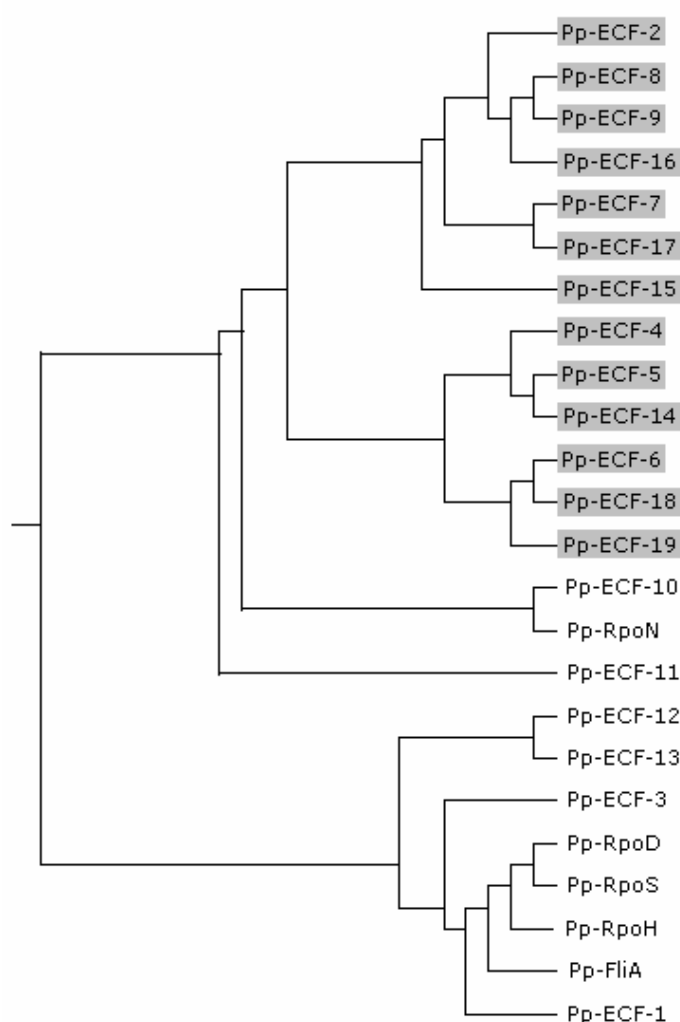


Figura 23: Dendrograma de los 24 factores sigma detectados en el genoma de *Pseudomonas putida* KT2440. El dendrograma muestra las relaciones filogenéticas entre los 24 factores sigma hipotéticos del genoma de *P. putida* KT2440 correspondientes al alineamiento múltiple de sus secuencias usando el programa CLUSTALW disponible en <http://clustalw.genome.ad.jp/>

También hemos analizado la organización de las regiones adyacentes a todos los genes de *P. putida* similares a *fecI* y se ha encontrado que en 11 de los casos, el gen similar a *fecI* está ligado a un gen similar a *fecR* (al que denominamos con las iniciales TMS de *TransMembrane Sensor*) (Figura 24), mientras que en dos casos (Pp-ECF-4 y Pp-ECF-6), el gen similar a *fecI* aparece sin un gen adyacente similar a *fecR*.

Pp-ECF-4 es casi idéntico a la proteína PflI de *P. putida* WCS358, un activador que regula la expresión de la biosíntesis de pseudobactina 358 en condiciones de limitación de hierro (Venturi *et al.*, 1995). La pseudobactina 358 es un sideróforo fluorescente amarillo-verde producido por esta cepa y puede ser considerado el equivalente funcional a la pioverdina de *P. aeruginosa*. El gen que codifica a Pp-ECF-4 está localizado junto a un gen que codifica una pioverdina sintetasa y una proteína sin función conocida, probablemente implicada en la biosíntesis del sideróforo (Figura 24). Respecto de Pp-ECF-6 carecemos de datos acerca de la función de los genes adyacentes.

Es de interés señalar que todas las agrupaciones del tipo *fecIR* de *P. putida* están seguidas por genes que codifican un receptor de ferrisideróforos o un receptor de hemo, excepto en el caso del gen que codifica Pp-ECF-16; este gen parece estar asociado a un sistema de eflujo de función desconocida (Figura 24).

El gen que codifica Pp-ECF-2 está ligado a 2 genes que codifican proteínas muy similares a FecR y FecA de *E. coli* (Figura 24). Esta conservación en secuencia y en organización genética sugeriría una función similar para esas proteínas en ambas bacterias. Este hallazgo junto con los resultados de la comparación de secuencias, sugieren que Pp-ECF-2 es muy probablemente el equivalente funcional de FecI de *E. coli*.

El gen que codifica Pp-ECF-14 está asociado con un gen similar a FecR seguido por un gen que codifica un supuesto receptor de hemo (Figura 24) muy similar a la proteína PfhR de *P. fluorescens* (Tabla 37). Su ortólogo en *P. aeruginosa* es el receptor PhuR dentro del locus *phu* (del inglés *Pseudomonas haem uptake*) que contiene el operón *phuSTUVW* que codifica un típico sistema transportador ABC (Ochsner *et al.*, 2000).

Agrupado con el gen que codifica a Pp-ECF-15 (mejor BLAST *hit* PupI, ver Tabla 33) y Pp-TMS-15 hay un gen cuyo producto es muy similar a la proteína PupB, que es el receptor de la pseudobactina férrica BN7/BN8 descrito en *P. putida* WCS358 (Tabla 37). La transcripción del gen *pupB* está regulada por el factor sigma ECF PupI y su correspondiente sensor transmembrana, PupR (Koster *et al.*, 1994, Visca *et al.*, 2002).

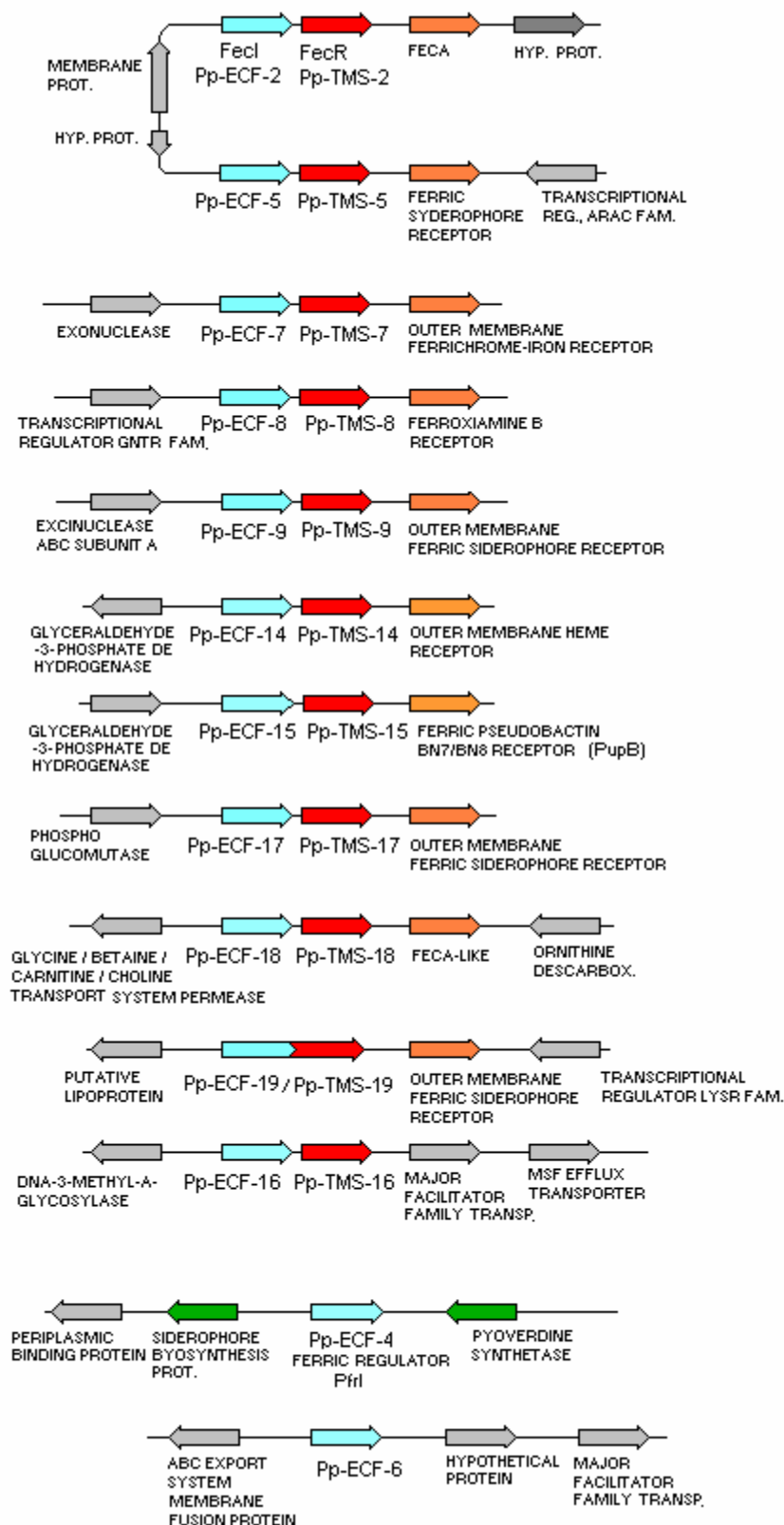


Figura 24: Genes similares a *fecl* y sus regiones adyacentes en *P. putida* KT2440. Hay 13 factores sigma ECF hipotéticos similares a *Fecl* de *E. coli* en *P. putida* KT2440. Los genes que codifican los factores sigma ECF similares a *Fecl* se representan con flechas azules y los genes de los sensores transmembrana similares a *FecR* se representan por flechas rojas. Los genes que codifican receptores de ferrosideróforos o receptores de hemo, se representan por flechas naranjas. Los genes no relacionados con la adquisición de hierro se representan con flechas grises.

Los análisis genéticos y bioquímicos de los miembros de la familia sigma-70 definen 4 regiones funcionales divididas en subregiones (Lonetto *et al.*, 1992). El alineamiento múltiple de los factores sigma ECF con σ^{70} mostró en ellos la ausencia de la mayor parte de la región 1 (característica de los factores σ ECF) y una conservación elevada de residuos relevantes en la región 2 y 4 del sigma-70 primario RpoD (Leoni *et al.*, 2000).

El alineamiento múltiple de los 13 factores sigma ECF de la subfamilia de FecI de *P. putida* KT2440, con FecI de *E. coli* (Figura 25) reveló una conservación mayor en los dominios 2.2 y 4.2. El dominio 2.2 de FecI participa con los dominios 2.1, 2.4 y 4.1 en las interacciones con la ARN polimerasa. El dominio 4.2 incluye un dominio hélice-giro-hélice que probablemente se une a la región -35 de los promotores. En el dominio 2.1 (Figura 25), las secuencias Pp-ECF-2, -7, -8, -9, -15, -16 y -17, las que forman la primera rama del dendrograma, las más similares a FecI, (Figura 23 y Tabla 36), muestran una tirosina en posición 25 y dos triptófanos en las posiciones 31 y 35, conservadas.

Como la región 2.1 de la proteína FecI de *E. coli* interacciona con los residuos 13-50 de FecR (Stiefel *et al.*, 2001), alineamos los primeros 60 residuos de las 11 proteínas de *P. putida* similares a FecR (Pp-TMS) asociadas a proteínas parecidas a FecI. Encontramos que con excepción de Pp-TMS-16, todas las proteínas similares a FecR que acompañan a proteínas similares de FecI del subgrupo de la primera rama del dendrograma compartían 3 triptófanos conservados que no estaban presentes en las otras proteínas similares a FecR de *P. putida* (Figura 26). Curiosamente los estudios de mutagénesis al azar de FecR en *E. coli* identificaron específicamente a estos triptófanos como residuos importantes en la actividad de FecR (Stiefel *et al.*, 2001).

El alineamiento de la proteína FecA de *E. coli* con otros receptores de sideróforos bacterianos reveló que su fragmento N-terminal, responsable de la interacción con FecR e implicada en la transducción de señal, no está presente en algunos de ellos (Kim *et al.*, 1997). Para probar la posible actividad reguladora de los receptores de membrana de *P. putida*, ligados a las proteínas del tipo FecIR, investigamos si contenían esta extensión N-terminal. Para ello seleccionamos un conjunto de secuencias (Pie de Figura 27) por similitud con RumA (EMBL *accession number* X90691), el receptor de la rizosferrina férrica de *Morganella morganii* (Kühn *et al.*, 1996), que muestra elevada similitud con FecA, pero que carece del fragmento N-terminal propio de FecA. El alineamiento de estas secuencias con los diez receptores férricos de *P. putida* KT2440 asociados a las agrupaciones similares a FecIR (Figura 27) mostró que todos los receptores férricos de *P. putida* tienen la extensión N-terminal, excepto Pp-ECF-14 que curiosamente, es el único receptor de hemo del grupo (Figura 2 y Tabla 37).

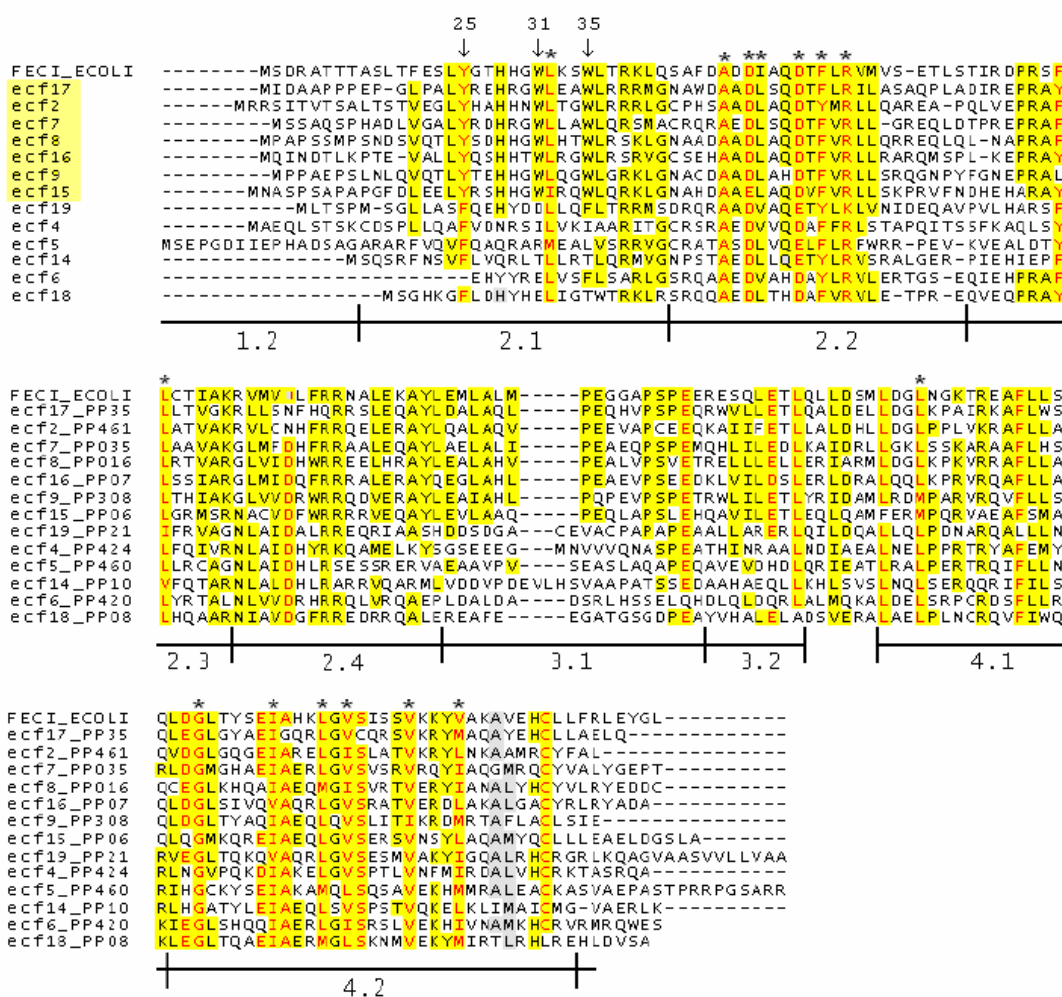


Figura 25: Alineamiento múltiple de los factores sigma ECF similares a FecI de *P. putida* KT2440 y FecI de *E. coli*. El alineamiento se realizó con el programa CLUSTALW disponible en <http://clustalw.genome.ad.jp/>. Los residuos idénticos y siguientes grupos de residuos se consideraron idénticos: DE, NQ, RK, ST, FYW e ILVM, como en Leoni *et al.* (2002). Los residuos que coincidían en el 100% de las secuencias alineadas se señalan con un asterisco y los residuos que coinciden en >50% de las secuencias alineadas se muestran resaltados en amarillo. Los dominios definidos para la familia sigma-70 están indicados de acuerdo con Leoni *et al.* (2002).

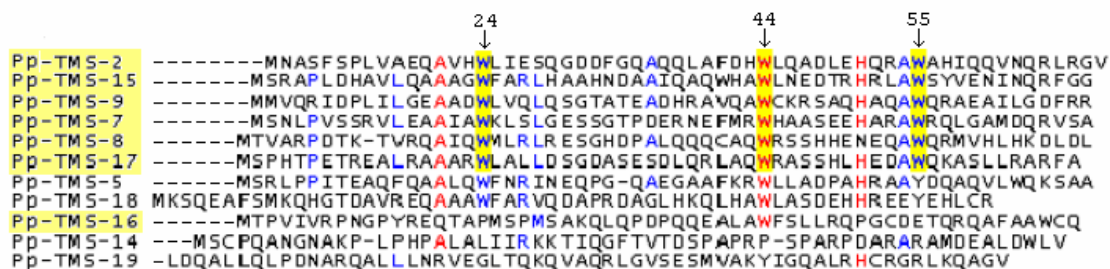


Figura 26: Alineamiento múltiple de los 60 primeros residuos de las 11 proteínas similares a FecR ligadas a los factores sigma similares a FecI en *P. putida* KT2440. Las proteínas similares a FecR Pp-TMS-2, -15, -9, -7, -8 y -17 comparten tres triptófanos (W), que se muestran resaltados.

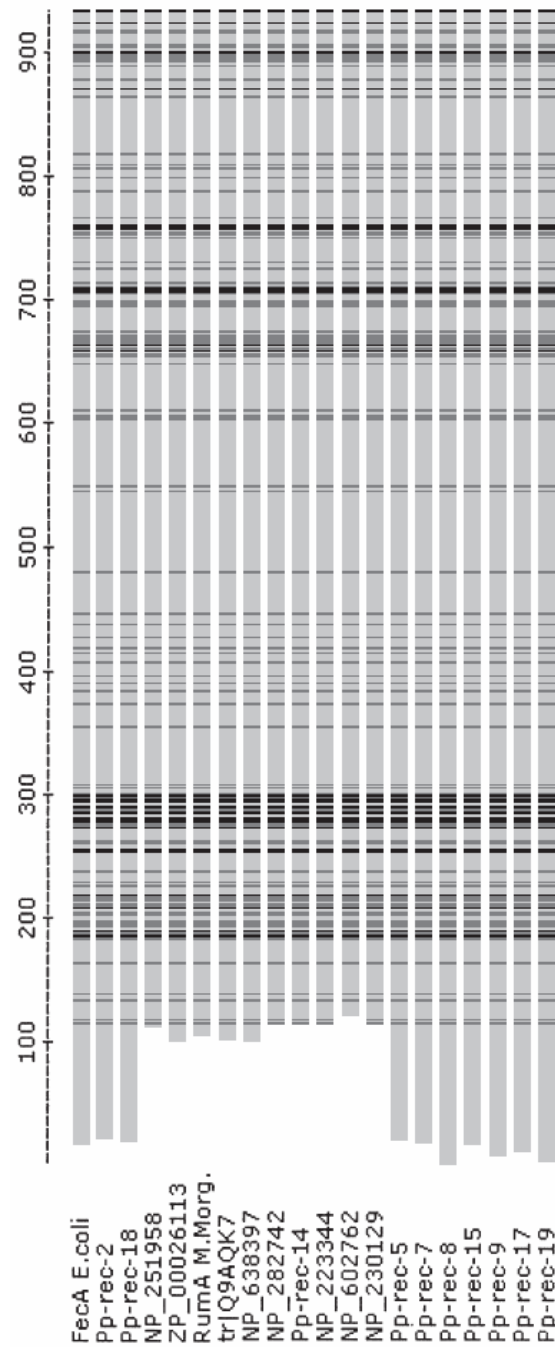


Figura 27: Representación esquemática del alineamiento múltiple de la región N-terminal de los receptores similares a FecR de *P. putida* KT2440 ligados a grupos *fecIR* con un conjunto de receptores sin extensión N-terminal responsable de la interacción con FecR. El alineamiento múltiple se efectuó con los 10 receptores férricos de *P. putida* KT2440 asociados a los grupos reguladores *fecIR* con un conjunto de secuencias similares a RumaA (Número de acceso EMBL X90691). Las secuencias extraídas de SWISS-PROT similares a RumaA fueron: NP_251958, un probable receptor dependiente de TonB de *Pseudomonas aeruginosa* PA01; ZP_00026113, una proteína hipotética de *Ralstonia metallidurans*; trIQ9AQK7, un receptor de vibrioferrina férrica de *Vibrio parahaemolyticus*; NP_638397, un transportador de hierro dependiente de citrato de *Xanthomonas campestris*; NP_282742, un receptor de membrana externa del sistema de captación de hemo de *Campylobacter jejuni*; NP_223344, una proteína transportadora de citrato de hierro (III) de *Helicobacter pylori* J99; NP_602762, un receptor de hemo de *Fusobacterium nucleatum*; y NP_230129, una proteína de virulencia de membrana externa regulada por hierro de *Vibrio cholera*. Las líneas verticales grises y negras indican los residuos en los que el grado de conservación en el alineamiento es superior al 40% y al 80% respectivamente.

Tabla 36: Valores de similitud entre los factores sigma ECF de *P. putida* y la proteína FecI de *E. coli*.

	Pp ECF-2	Pp ECF-4	Pp ECF-5	Pp ECF-6	Pp ECF-7	Pp ECF-8	Pp ECF-9	Pp ECF-14	Pp ECF-15	Pp ECF-16	Pp ECF-17	Pp ECF-18	Pp ECF-19	<i>E. coli</i> FecI
Pp ECF-2	100%	16 %	16 %	22 %	41 %	47 %	40 %	17 %	36 %	47 %	45 %	18 %	23 %	43 %
Pp ECF-4		100 %	20 %	17 %	21 %	16 %	13 %	20 %	18 %	20 %	14 %	19 %	24 %	16 %
Pp ECF-5			100 %	21 %	19 %	18 %	18 %	29 %	17 %	21 %	17 %	22 %	23 %	17 %
Pp ECF-6				100 %	28 %	28 %	21 %	22 %	22 %	24 %	22 %	33 %	29 %	26 %
Pp ECF-7					100 %	43 %	36 %	16 %	34 %	44 %	42 %	25 %	26 %	43 %
Pp ECF-8						100 %	49 %	17 %	43 %	52 %	44 %	23 %	21 %	43 %
Pp ECF-9							100 %	18 %	40 %	45 %	43 %	22 %	19 %	40 %
Pp ECF-14								100 %	17 %	21 %	12 %	17 %	22 %	16 %
Pp ECF-15									100 %	37 %	42 %	21 %	22 %	37 %
Pp ECF-16										100 %	41 %	23 %	25 %	41 %
Pp ECF-17											100 %	30 %	24 %	48 %
Pp ECF-18												100 %	30 %	23 %
Pp ECF-19													100 %	25 %
<i>E.coli</i> FecI														100%

Hemos definido un parámetro al que denominamos **valor de similitud** para cuantificar la similitud entre un par de secuencias A y B. Este parámetro se basa en tres valores: (i) puntuación BLAST obtenida al comparar las secuencias A y B; (ii) puntuación BLAST obtenida al comparar la secuencia A consigo misma; (iii) puntuación BLAST obtenida al comparar la secuencia B consigo misma. La puntuación más baja que resulta de comparar las secuencias A y B consigo mismas representa la "puntuación máxima" posible. En la mayoría de los casos, la proteína más corta de cada par determina el valor de la puntuación máxima. El valor de similitud se define como la razón entre la puntuación obtenida al comparar las dos secuencias y la "puntuación máxima". Cada celda de la tabla indica valores de similitud expresados en porcentaje. Se ha usado el programa BLAST descargado del NCBI (<http://www.ncbi.nlm.nih.gov/Sitemap/index.html>).

Tabla 37: Receptores asociados con los agrupamientos de proteínas similares a FecIR en *P. putida*.

Receptor	Id ORF	Mejor resultado BLAST SWISS-PROT/TrEMBL	E	Organism	Mejor resultado BLAST con anotación funcional	E	Organism		
Pp-rec-2	PP4613	FecA	Q9HXB2	0.0	<i>P. aeruginosa</i>	FecA	P13036	0.0	<i>Escherichia coli</i>
Pp-rec-5	PP4606	Prob. rec. TonB-dependiente	Q9I6Y1	0.0	<i>P. aeruginosa</i>	Rec. ferricromo-hierro (FhuA)	P72609	3E-78	<i>Synechocystis sp.</i>
Pp-rec-7	PP0350	Prob. rec. ferrosideróforo hidroxamato	Q9I648	0.0	<i>P. aeruginosa</i>	Rec. ferrosideróforo hidroxamato (FiuA)	O68595	0.0	<i>P. aeruginosa</i>
Pp-rec-8		Prob. rec. TonB-dependiente	Q9I116	0.0	<i>P. aeruginosa</i>	Rec. ferrioxamina B	Q56145	0.0	<i>S. typhimurium</i>
Pp-rec-9	PP3084	Rec. ferricromo fcuA	Q05202	1E-165	<i>Y. enterocolitica</i>	Rec. ferricromo fcuA	Q05202	1E-165	<i>Y. enterocolitica</i>
Pp-rec-14	PP1006	Rec. hemo membrana ext. PfhR	Q9X5P3	0.0	<i>P. fluorescens</i>	Rec. hemo membrana externa PfhR	Q9X5P3	0.0	<i>P. fluorescens</i>
Pp-rec-15	PP0669	Prob. rec. TonB-dependiente	Q9HWL3	1E-168	<i>P. aeruginosa</i>	Rec. pseudobactina-férrica BN7/BN8 (PupB)	P38047	1E-165	<i>P. putida</i>
Pp-rec-17	PP3575	Prob. rec. TonB-dependiente	Q9I2J4	1E-162	<i>P. aeruginosa</i>	Pro. rec. siderofofo férrico (BfrB)	P94265	1E-136	<i>B. bronchiseptica</i>
Pp-rec-18	PP0867	Prob. rec. TonB-dependiente	Q9HYX3	0.0	<i>P. aeruginosa</i>	FecA	P13036	1E-180	<i>Escherichia coli</i>
Pp-rec-19	PP2193	Prob. rec. TonB-dependiente	Q9HVC0	0.0	<i>P. aeruginosa</i>	Rec. aerobactina férrica IutA	Q934J5	1E-144	<i>V. parahaemolyticus</i>

Los números de identificación de los marcos de lectura abierta (*orf*) están disponibles en <http://www.tigr.org> (segunda columna). Las tres columnas centrales muestran el mejor resultado de BLAST con las proteínas depositadas en SWISSPROT/TrEMBL junto con el número de acceso, el valor de E y el organismo del que procede. Las columnas de la derecha muestran el mejor resultado de BLAST con anotaciones funcionales.

Tabla 38: Comparación de los factores sigma ECF de *P. putida* KT2440 y *P. aeruginosa* PA01.

	Pa- ECF- 1	Pa- ECF- 2	Pa- ECF- 3	Pa- ECF- 4	Pa- ECF- 5	Pa- ECF- 6	Pa- ECF- 7	Pa- ECF- 8	Pa- ECF- 9	Pa- ECF- 10	Pa- ECF- 11	Pa- ECF- 12	Pa- ECF- 13	Pa- ECF- 14	Pa- ECF- 15	Pa- ECF- 16	Pa- ECF- 17	Pa- ECF- 18	Pa- ECF- 19
Pp-ECF-1	94 %	12 %	15 %	7 %	10 %	-	-	9 %	-	8 %	-	-	7 %	11 %	-	20 %	16 %	10 %	6 %
Pp-ECF-2	10 %	66 %	7 %	17 %	15 %	24 %	42 %	45 %	51 %	9 %	14 %	25 %	17 %	41 %	35 %	13 %	40 %	22 %	38 %
Pp-ECF-3	19 %	10 %	93 %	7 %	11 %	15 %	5 %	7 %	10 %	9 %	6 %	9%	-	8 %	6 %	19 %	21 %	14 %	7 %
Pp-ECF-4	7 %	15 %	9 %	86 %	18 %	19 %	22 %	19 %	16 %	5 %	15 %	23 %	17 %	13 %	16 %	12 %	6 %	20 %	16 %
Pp-ECF-5	11 %	19 %	13 %	20 %	58 %	24 %	21 %	19 %	23 %	8 %	18 %	21 %	17 %	18 %	16 %	15 %	-	25 %	14 %
Pp-ECF-6	12 %	24 %	13 %	18 %	20 %	68 %	27 %	27 %	22 %	12 %	21 %	33 %	23 %	4 %	21 %	-	8 %	27 %	19 %
Pp-ECF-7	-	45 %	7 %	22 %	16 %	24 %	79 %	45 %	46 %	12 %	12 %	23 %	16 %	42 %	37 %	12 %	-	20 %	33 %
Pp-ECF-8	-	50 %	7 %	17 %	15 %	30 %	44 %	70 %	50 %	7 %	14 %	30 %	19 %	37 %	34 %	13 %	7 %	21 %	35%
Pp-ECF-9	-	47 %	10 %	14 %	19 %	24 %	37 %	45 %	53 %	8 %	15 %	26 %	18 %	36 %	32 %	15 %	-	19 %	38 %
Pp-ECF10	-	6 %	8 %	6 %	6 %	10 %	12 %	-	-	67 %	6 %	7 %	5 %	7 %	7 %	8 %	5 %	-	11 %
Pp-ECF11	10 %	13 %	12 %	7 %	9 %	-	10 %	-	15 %	8 %	-	10 %	8 %	12 %	10 %	11 %	7 %	-	14 %
Pp-ECF12	13 %	12 %	14 %	-	9 %	11 %	11 %	10 %	11 %	10 %	-	7 %	-	10 %	9 %	12 %	10 %	10 %	10 %
Pp-ECF13	8 %	13 %	15 %	9 %	13 %	14 %	11 %	9 %	11 %	9 %	7 %	13 %	7 %	9 %	6 %	11 %	14 %	13 %	8 %
Pp-ECF14	8 %	16 %	12 %	20 %	29 %	22 %	17 %	17 %	14 %	5 %	22 %	24 %	13 %	16 %	9 %	15 %	10 %	28 %	11 %
Pp-ECF15	6 %	35 %	7 %	21 %	14 %	23 %	36 %	41 %	43 %	8 %	10 %	23 %	15 %	29 %	25 %	15 %	-	22 %	32 %
Pp-ECF16	9 %	50 %	9 %	21 %	21 %	27 %	46 %	52 %	51 %	12 %	17 %	29 %	18 %	38 %	37 %	16 %	-	25 %	37 %
Pp-ECF17	9 %	50 %	7 %	14 %	17 %	20 %	41 %	46 %	52 %	9 %	13 %	21 %	14 %	35 %	35 %	11 %	7 %	22 %	39 %
Pp-ECF18	9 %	24 %	12 %	20 %	25 %	36 %	27 %	24 %	24 %	13 %	20 %	28 %	23 %	20 %	19 %	12 %	8 %	26 %	18 %
Pp-ECF19	-	21 %	14 %	23 %	23 %	35 %	24 %	25 %	26 %	12 %	25 %	29 %	23 %	18 %	19 %	13 %	9 %	30 %	18 %

Tabla 39: Comparación de los sensores transmembrana (TMS) asociados con los factores sigma ECF de *P.putida* KT2440 y *P. aeruginosa* PA01.

	Pa-TMS02	Pa-TMS05	Pa-TMS6	Pa-TMS07	Pa-TMS08	Pa-TMS09	Pa-TMS11	Pa-TMS12	Pa-TMS13	Pa-TMS14	Pa-TMS15	Pa-TMS18	Pa-TMS19
Pp-TMS02	43 %	17 %	14 %	26 %	33 %	31 %	12 %	14 %	11 %	23 %	17 %	16 %	25 %
Pp-TMS05	15 %	41 %	21 %	18 %	11 %	14 %	16 %	17 %	19 %	15 %	12 %	24 %	18 %
Pp-TMS07	32 %	17 %	13 %	47 %	30 %	27 %	11 %	15 %	15 %	25 %	20 %	18 %	25 %
Pp-TMS08	29 %	12 %	10 %	21 %	44 %	30 %	9 %	13 %	14 %	19 %	15 %	16 %	20 %
Pp-TMS09	36 %	18 %	13 %	27 %	34 %	38 %	17 %	18 %	15 %	26 %	19 %	19 %	34 %
Pp-TMS14	17 %	20 %	13 %	14 %	14 %	16 %	14 %	12 %	12 %	13 %	11 %	14 %	13 %
Pp-TMS15	32 %	16 %	13 %	23 %	26 %	25 %	11 %	17 %	16 %	24 %	18 %	18 %	27 %
Pp-TMS16	18 %	18 %	12 %	12 %	12 %	17 %	13 %	15 %	15 %	17 %	16 %	13 %	16 %
Pp-TMS17	28 %	16 %	10 %	19 %	27 %	29 %	16 %	15 %	15 %	21 %	20 %	13 %	24 %
Pp-TMS18	23 %	21 %	25 %	20 %	22 %	19 %	14 %	27 %	26 %	21 %	13 %	22 %	18 %
Pp-TMS19	16 %	18 %	13 %	15 %	13 %	18 %	12 %	16 %	19 %	16 %	11 %	21 %	15 %

(Valores de similitud calculados tal como se describe al pie de la Tabla 36).

Estos resultados, tomados en conjunto, sugieren que cada receptor de *P. putida* dotado de extensión N-terminal probablemente interacciona con su correspondiente proteína similar a FecR para transmitir una señal a su sistema específico de captación de hierro férrico.

7.3.4. Comparación de factores sigma ECF de *P. putida* y *P. aeruginosa*.

Para revelar las características diferenciales en el repertorio de los factores sigma entre la bacteria del suelo *P. putida* y el patógeno humano *P. aeruginosa*, se realizó un análisis comparativo (para homogenizar la redacción, usamos el mismo tipo de nomenclatura sistemática empleado con *P. putida*).

En ambas especies se han caracterizado los ortólogos funcionales para los factores sigma σ^{32} , σ^{38} , FliA, σ^{54} y σ^{70} . Para establecer relaciones de similitud entre los factores sigma ECF de *P. aeruginosa* y *P. putida*, hemos analizado los resultados de la comparación con BLAST entre todos los factores sigma ECF. Para evaluar la similitud hemos empleado un valor (sv; ver pie de nota en Tabla 36) que considera la similitud a lo largo de las secuencias completas. Los factores sigma ECF de 1 a 10 de *P. putida* proporcionaron proteínas relacionadas en *P. aeruginosa* con valores de similitud muy elevados (> 50%) (Tabla 38).

ECF-1 en ambas bacterias está asociado a OprF, y el elevado valor de similitud (94%) indica una posible función común. Pa-ECF-3 (AlgU) proporcionó un valor de similitud (93%) con el Pp-ECF-3. El factor sigma de control de síntesis de pioverdina (PvdS) de *P. aeruginosa* produjo un valor de similitud (86%) con Pp-ECF-4. PvdS está además implicado en el control de la expresión de al menos 2 factores de virulencia: exotoxina A y una exoproteasa (Lamont *et al.*, 2002; Shen *et al.*, 2002). En *P. putida* BTP1, se ha propuesto que el sistema de adquisición de hierro mediado por la pioverdina, contribuye a la capacidad ecológica de la cepa y su capacidad de colonizar y persistir en la rizosfera (Ongena *et al.*, 2002).

Cinco de los factores sigma ECF de *P. aeruginosa* (Pa-ECF-2, -5, -7, -8 y -9) fueron significativamente similares a proteínas del tipo FecI de *P. putida*. En este grupo encontramos que las proteínas similares a FecR y receptoras de hierro, codificadas por los genes adyacentes eran también bastante similares (Tabla 39 Tabla 40), constituyendo 5 agrupaciones cuyas secuencias y organización genética estuvieron conservadas en ambas especies. El hecho de que *P. putida* y *P. aeruginosa* compartan 5 sistemas de adquisición de hierro sugiere un origen evolutivo común y puede correlacionar con su capacidad para compartir ciertos nichos.

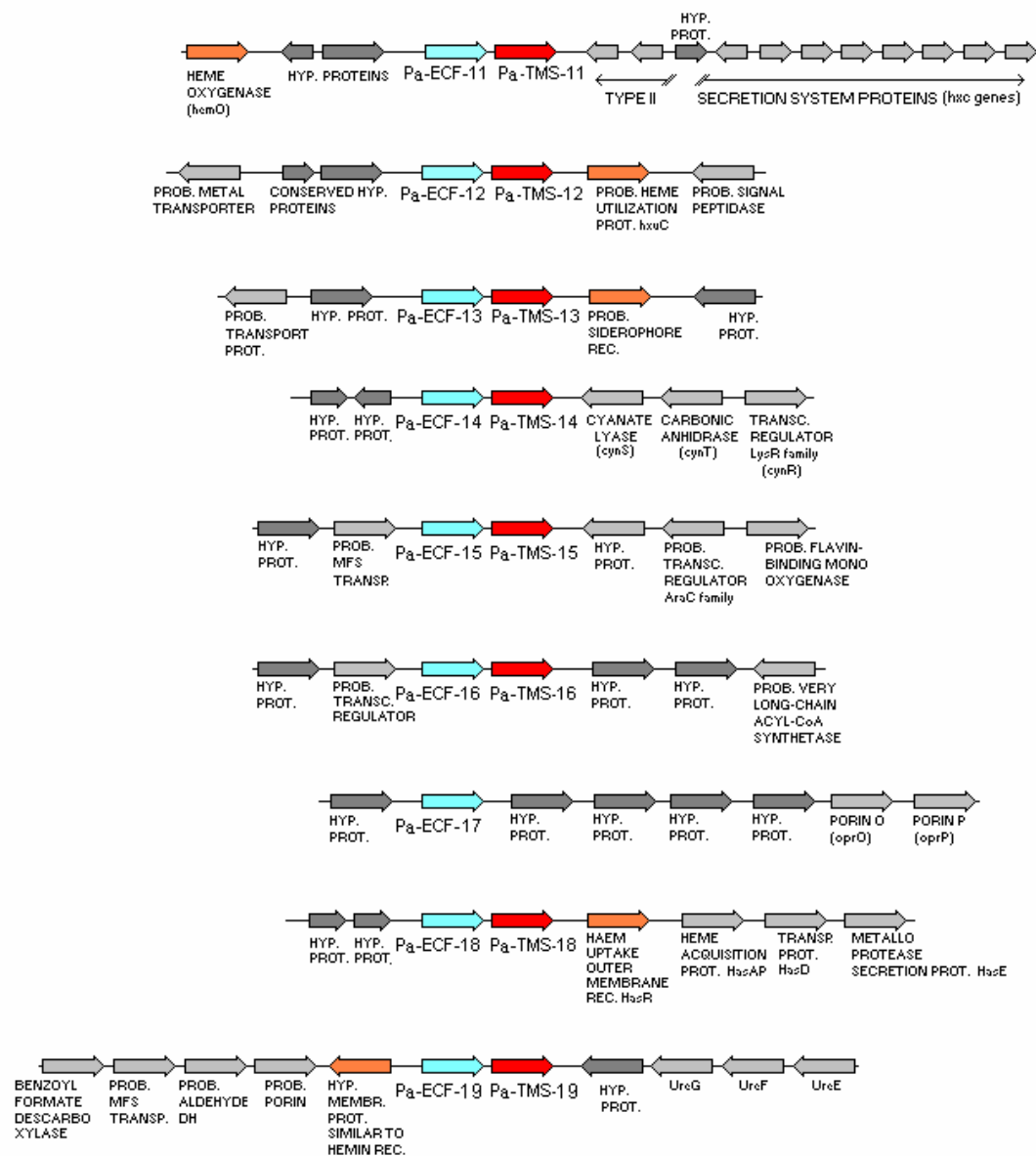


Figura 28: Genes de *Pseudomonas aeruginosa* que codifican factores sigma ECF y sus regiones adyacentes. Organización genética correspondiente a los factores sigma ECF de *P. aeruginosa* sin un similar en *P. putida*. Los genes correspondientes a factores sigma ECF se representan con flechas azules, los genes de sensores transmembrana se representan con flechas rojas, y los genes que codifican receptores de sideróforos férricos o receptores de hemo, aparecen como flechas naranjas. Los genes que codifican proteínas sin relación con la adquisición de hierro se muestran como flechas grises.

Por otro lado hemos encontrado 9 factores sigma ECF in *P. aeruginosa* que no tenían sistemas similares en *P. putida* (Pa-ECF-11-19) (Tabla 38 y Figura 28). Estos factores sigma diferentes de *P. aeruginosa* pudieran estar implicados en la patogenicidad de esta última cepa. La capacidad de usar hemo es característica de varios patóge-

nos humanos. Hemos encontrado que 3 de los factores sigma específicos de *P. aeruginosa* están asociados a genes que codifican proteínas receptoras de hemo (Pa-ECF-12, Pa-ECF-18 y Pa-ECF-19) (ver Figura 28). Pa-ECF-18, denominado HasR, se ha caracterizado bioquímicamente y genéticamente (Ochsner *et al.*, 2000), y el locus *has* codifica un sistema de captación de hemo constituido por un receptor de membrana externa (HasR), una proteína de unión a hemo extracelular y un aparato de secreción de tipo I. Otro factor sigma ECF interesante, sin equivalente en *P. putida* es el Pa-ECF-11, que está ligado a 3 genes llamados *pig*, regulados por Fur y a los genes *hxc* que codifican un sistema de secreción de tipo II (Ochsner y Vasil, 1996).

Tabla 40: Comparación de los receptores de hierro (rec) asociados a los factores sigma ECF de *P. putida* KT2440 y *P. aeruginosa* PA01.

	Pa-rec02	Pa-rec05	Pa-rec07	Pa-rec08	Pa-rec09	Pa-rec12	Pa-rec13	Pa-rec18
Pp-rec02	67 %	9 %	8 %	9 %	8 %	3 %	5 %	4 %
Pp-rec05	8 %	64 %	15 %	16 %	10 %	3 %	6 %	4 %
Pp-rec07	9 %	14 %	69 %	27 %	9 %	3 %	5 %	3 %
Pp-rec08	5 %	15 %	33 %	62 %	10 %	4 %	5 %	4 %
Pp-rec09	4 %	11 %	9 %	8 %	34 %	4 %	3 %	3 %
Pp-rec14	6 %	4 %	6 %	6 %	3 %	8 %	6 %	9 %
Pp-rec15	8 %	10 %	11 %	12 %	8 %	3 %	3 %	3 %
Pp-rec17	8 %	8 %	9 %	10 %	36 %	4 %	4 %	3 %
Pp-rec18	40 %	8 %	9 %	6 %	6 %	6 %	6 %	4 %
Pp-rec19	7 %	6 %	5 %	6 %	4 %	5 %	18 %	6 %

(Los valores de similitud se calcularon como se describe al pie de la Tabla 36. Se marcan los valores de similitud correspondientes a las cinco agrupaciones similares a *fecIRA* cuya secuencia y orden de genes está conservado entre ambas especies).

Conclusiones

Los resultados obtenidos y expuestos en esta memoria permitieron establecer las siguientes conclusiones:

- I. La base de datos BacTregulators está ideada para detectar los reguladores en los genomas procariotas y clasificarlos en familias y considera la integración de la información que sobre cada regulador se ha recopilado de diferentes bases de datos y de la bibliografía. Se basa en unos criterios bien definidos que enfatizan la importancia de los datos genómicos en la definición de una proteína, lo que evita la redundancia.
- II. La posibilidad de obtener un informe que integra los datos de las entradas relacionadas con la entrada de la petición, por similitud de secuencia y nivel taxonómico (seleccionados por el usuario), es útil no sólo para compilar el conocimiento acerca de las propiedades biológicas de un regulador conocido, sino también para propagarlo a entradas sin función conocida (si la función es similar en todas las proteínas caracterizadas del grupo, es probable que las proteínas no caracterizadas posean una función similar), contribuyendo además a su anotación más segura en los genomas.
- III. La información en BacTregulators está altamente fragmentada y estructurada lo que permite la aplicación de herramientas que interroguen a la base de datos sobre diversas cuestiones relacionadas con las funciones reguladas y su fuente de origen.
- IV. Se ha construido un perfil convencional (basado en una secuencia semilla que hemos seleccionado) para la familia de reguladores TetR. Este perfil es muy eficaz en la detección de los miembros de la familia y en su discriminación de otras proteínas con dominios hélice-giro-hélice similares a los de TetR.
- V. Los genes que codifican proteínas de la familia TetR se localizan tanto en elementos cromosómicos como extracromosómicos.
- VI. El alto grado de conservación de la secuencia primaria y secundaria en el fragmento empleado para construir el perfil tiene su reflejo en la práctica identidad de su estructura tridimensional, lo que nos ha permitido proponer un modo de interacción con el ADN común para los miembros de la familia.

- VII. Los miembros de la familia TetR participan en una amplia variedad de esquemas de regulación. La red puede ser tan simple como en TetR, o implicar cascadas reguladoras sucesivas y/o convergentes. En estos procesos, los miembros de la familia TetR controlan la expresión de genes cuyos productos están implicados en la resistencia a múltiples drogas y compuestos tóxicos, enzimas implicadas en diferentes rutas catabólicas, biosíntesis de antibióticos, estrés osmótico, colonización de tejidos en organismos hospedadores que puede derivar en establecimiento de simbiosis o en patogenicidad de bacterias tanto gram negativas como gram positivas.
- VIII. Los factores sigma ECF se caracterizan por su capacidad para coordinar la transcripción en respuesta a estímulos extra citoplasmáticos. El número elevado de factores sigma presentes en los genomas de *P. putida* y *P. aeruginosa* correlaciona con la capacidad ampliamente demostrada de ambos microorganismos para disparar respuestas adaptativas ante variaciones ambientales y para asegurar una adquisición eficiente de hierro.
- IX. La existencia tanto en *P. putida* KT2440 como *P. aeruginosa* PA01 de cinco agrupaciones formadas por un factor sigma ECF, un sensor transmembrana y un receptor para sideróforos, muestra que estas cepas están equipadas con un conjunto de sistemas de adquisición de hierro que sensan el sideróforo capaz de unir hierro en el medio extracelular. La detección de un sideróforo en particular activaría la transcripción del sistema de captura de hierro más eficiente y competitivo, lo que facilitaría la distribución ubicua de esta bacteria.
- X. Muchos de los factores σ de *P. putida* tienen un equivalente en el patógeno humano *P. aeruginosa*, pero éste posee factores sigma ECF relacionados con la adquisición de hemo y sistemas de secreción asociados a la virulencia. En *P. aeruginosa*, el hierro parece constituir una señal importante para la virulencia y todo un conjunto de factores sigma ECF participa en la regulación de determinantes de patogenicidad importantes. Estos sistemas constituyen dianas potenciales para el diseño de fármacos contra este patógeno oportunista.

Bibliografía

- Abo-Amer6565, A.E., J. Munn, K. Jackson, M. Aktas, P. Golby, D.J. Kelly, and S.C. Andrews. 2004. DNA interaction and phosphotransfer of the C4-dicarboxylate-responsive DcuS-DcuR two-component regulatory system from *Escherichia coli*. *J. Bacteriol.* 186:1879-1889.
- Ahlert, J., E. Shepard, N. Lomovskaya, E. Zazopoulos, A. Staffa, B.O. Bachmann, K. Huang, L. Fonstein, A. Czisny, R.E. Whitwam, C.M. Farnet, and J.S. Thorson. 2002. The calicheamicin gene cluster and its iterative type I enediyne PKS. *Science* 297:1173-1176.
- Aigle, B., A. Wietzorrek, E. Takano and M.J. Bibb. 2000. A single amino acid substitution in region 1.2 of the principal sigma factor of *Streptomyces coelicolor* A3(2) results in pleiotropic loss of antibiotic production. *Mol. Microbiol.* 37:995-1004.
- Alekshun, M.N., and S.B. Levy. 1999. Alteration of the repressor activity of MarR, the negative regulator of the *Escherichia coli* *marRAB* locus, by multiple chemicals in vitro. *J. Bacteriol.* 181:4669- 4672.
- Allignet, J., S. Aubert, K.G. Dyke, and N. El Solh. 2001. *Staphylococcus caprae* strains carry determinants known to be involved in pathogenicity: a gene encoding an autolysin-binding fibronectin and the *ica* operon involved in biofilm formation. *Infect. Immun.* 69:712-718.
- Allmeier, H., B. Cresnar, M. Greck, and R. Schmitt. 1992. Complete nucleotide sequence of Tn1721: gene organization and a novel gene product with features of a chemotaxis protein. *Gene* 111:11-120.
- Amabile-Cuevas, C.F., and B. Demple. 1991. Molecular characterization of the *soxRS* genes of *Escherichia coli*: two genes control a superoxide stress regulon. *Nucl. Acids Res.* 19:4479-4484.
- Andresen, P.A., I. Kaasen, O.B. Styrvoid, G. Boulnois, and A.R. Strøm. 1988. Molecular cloning, physical mapping and expression of the *bet* genes governing the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. *J. Gen. Microbiol.* 134:1737-1746.
- Aramaki, H., N. Yagi, and M. Suzuki. 1995. Residues important for the function of a multihelical DNA binding domain in the new transcription factor family of Cam and Tet repressors. *Protein Eng.* 8:1259-1266.
- Aramaki, H., Y. Sagara, H. Kabata, N. Shimamoto, and T. Horiuchi. 1995. Purification and characterization of a *cam* repressor *CamR*. for the cytochrome P-450cam hydroxylase operon on the *Pseudomonas putida* CAM plasmid. *J. Bacteriol.* 177:3120-3127.
- Aramaki, H., Y. Sagara, M. Hosoi, and T. Horiuchi. 1993. Evidence for autoregulation of *camR*, which encodes a repressor for the cytochrome P-450cam hydroxylase operon on the *Pseudomonas putida* CAM plasmid. *J. Bacteriol.* 175:7828-7833.
- Aravind, L. Anantharaman, V. Balaji, S. Babu, M.M. Iyer, L.M. 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol. Rev.* 29:231-62.
- Argos, P., M. Hanei, J.M. Wilson, and W.N. Kelley. 1983. A possible nucleotide-binding domain in the tertiary fold of phosphoribosyltransferases. *J. Biol. Chem.* 258:6450-6457.
- Ariza, R.R., S.P. Cohen, N. Bachawat, S.B. Levy, and B. Demple. 1994. Repressor mutations in the *marRAB* operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. *J. Bacteriol.* 176:143-148.
- Ariza, R.R., Z. Li, N. Ringstad, and B. Demple. 1995. Activation of multiple antibiotic resistance and binding of stress-inducible promoters by *Escherichia coli* *rob* protein. *J. Bacteriol.* 177:1655-1661.
- Arora, S.K., Ritchings, B.W., Almira, E.C., Lory, S., and Ramphal, R. 1997. A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene expression in *Pseudomonas aeruginosa* in a cascade manner. *J. Bacteriol.* 179: 5574-5581.
- Attwood T.K., M.D. Croning, D.R. Flower, A.P. Lewis, J.E. Mabey, P. Scordis, J.N. Selley, and W. Wright. 2000. Prints-S: "The database formerly known as Prints". *Nucl. Acids Res.* 28:225-227.
- Attwood, T.K., P. Bradley, D.R. Flower, A. Gaulton, N. Maudling, A.L. Mitchell, G. Moulton, A. Nordle, K. Paine, P. Taylor, A. Uddin, and C. Zygouri. 2003. PRINTS and its automatic supplement, prePRINTS. *Nucl. Acids Res.* 31:400-402.
- August, P.R., L. Tang, Y.J. Yoon, S. Ning, R. Muller, T.W. Yu, M. Taylor, D. Hoffmann, C.G. Kim, X. Zhang, C.R. Hutchinson, and H.G. Floss. 1998. Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* 5:69-79.

- Babu, M.M., and S.A. Teichmann. 2003. Evolution of transcription factors and the gene regulatory network in *Escherichia coli*. *Nucl. Acids Res.* 31:1234-1244.
- Bairoch, A., and K.E. Rudd. 1995. PROSITE. Accession no. PDOC00874.
- Balsalobre, L., M.J. Ferrándiz, J. Liñares, F. Tubau, and A.G. de la Campa. 2003. Viridans group Streptococci are donors in horizontal transfer of topoisomerase IV genes to *Streptococcus pneumoniae*. *Antimicrob. Agents Chemother.* 47:2072-2081.
- Barbosa, T.M., and S.B. Levy. 2000. Differential expression of over 60 chromosomal genes in *Escherichia coli* by constitutive expression of MarA. *J. Bacteriol.* 182:3467-3474.
- Baron, U., M. Gossen, and H. Bujard. 1997. Tetracycline-controlled transcription in eukaryotes: novel transactivators with graded transactivation potential. *Nucl. Acids Res.* 25:2723-2729.
- Bassler, B.L., M. Wright, and M.R. Silverman. 1994. Multiple signalling systems controlling expression of luminescence in *Vibrio harveyi*: sequence and function of genes encoding a second sensory pathway. *Mol. Microbiol.* 13:273-286.
- Bassler, B.L., M. Wright, and M.R. Silverman. 1994. Sequence and function of LuxO, a negative regulator of luminescence in *Vibrio harveyi*. *Mol. Microbiol.* 12:403-412.
- Bate, N., A.R. Butler, A.R. Gandecha, and E. Cundliffe. 1999. Multiple regulatory genes in the tylosin biosynthetic cluster of *Streptomyces fradiae*. *Chem. Biol.* 6:617-624.
- Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Ewinger, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall, and E.L. Sonnhammer. 2002. The Pfam protein families database. *Nucl. Acids Res.* 30:276-280.
- Bateman, A., E. Birney, R. Durbin, S.R. Eddy, K.L. Howe, and E.L. Sonnhammer. 2000. The Pfam protein families database. *Nucl. Acids Res.* 28:263-266.
- Baulard, A.R., Betts, J.C., Engohang-Ndong, J., Quan, S., McAdam, R.A., Brennan, P.J., Locht, C., Besra, G.S. 2000. Activation of the pro-drug ethionamide is regulated in mycobacteria. *J. Biol. Chem.* 275:28326-28331.
- Baumeister, R., G. Muller, B. Hecht, and W. Hillen. 1992. Functional roles of amino acid residues involved in forming the alpha-helix-turn-alpha-helix operator DNA binding motif of Tet repressor from Tn10. *Proteins* 14:168-177.
- Baumeister, R., V. Helbl, and W. Hillen. 1992. Contacts between Tet repressor and *tet* operator revealed by new recognition specificities of single amino acid replacement mutants. *J. Mol. Biol.* 226:1257-1270.
- Becskei, A., and L. Serrano. 2000. Engineering stability in gene networks by autoregulation. *Nature* 405:590-593.
- Beinlich, K.L., R. Chuanchuen, and H.P. Schweizer. 2001. Contribution of multidrug efflux pumps to multiple antibiotic resistance in veterinary clinical isolates of *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett.* 198:129-134.
- Belli, G., E. Gari, L. Piedrafita, M. Aldea, and E. Herrero. 1998. An activator/repressor dual system allows right tetracycline regulated gene expression in budding yeast. *Nucl. Acids Res.* 26:942-947.
- Beloin, C., S. Ayora, R. Exley, L. Hirschbein, N. Ogasawara, Y. Kasahara, J.C. Alonso, and F. Le Hégat. 1997. Characterization of an *Irp*-like *IrpC* gene from *Bacillus subtilis*. *Mol. Gen. Genet.* 256:63-71.
- Bennik, H.J.M., P.J. Pomposiello, D.F. Thorne, and B. Dimple. 2000. Defining a *rob* regulon in *Escherichia coli* by using transposon mutagenesis. *J. Bacteriol.* 182:3794-3801.
- Berens, C., Altschmied, L., and W. Hillen. 1992. The role of the N terminus in Tet repressor for *tet* operator binding determined by a mutational analysis. *J. Biol. Chem.* 267:1945-1952.
- Berens, C., and W. Hillen. 2003. Gene regulation by tetracycline. Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes. *Eur. J. Biochem.* 270:3109-3121.
- Berens, C., D. Schnappinger, and W. Hillen. 1997. The role of the variable region in Tet repressor for inducibility by tetracycline. *J. Biol. Chem.* 272:6936-6942.
- Bertoni, G. Fujita, N. Ishihama, A. de Lorenzo, V. 1998. Active recruitment of sigma54-RNA polymerase to the Pu promoter of *Pseudomonas putida*: role of IHF and alphaCTD. *EMBO J.* 17:5120-8.

- Biburger, M., C. Berens, T. Lederer, T. Krec, and W. Hillen. 1998. Intragenic suppressors of induction-deficient TetR mutants: localization and potential mechanism of action. *J. Bacteriol.* 180:737-741.
- Blanco, C., P. Ritzenthaler, and M. Mata-Gilsinger. 1986. Negative dominant mutations of the *uidR* gene in *Escherichia coli*: genetic proof for a cooperative regulation of *uidA* expression. *Genetics* 112:173-182.
- Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S. and Schneider, M. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.*, 31, 365–370.
- Boopathi, E., and Rao, K.S.A. 1999. Siderophore from *Pseudomonas putida* type A1: structural and biological characterization. *Biochim Biophys Acta* 1435: 30–40.
- Branden, C., and J. Tooze. 1991. Introduction to protein structure. Garland Publishing Co., New York, N.Y.
- Brandi, A., C.L. Pon, and C.O. Gualerzi. 1994. Interaction of the main cold shock protein CS7.4 CspA. of *Escherichia coli* with the promoter region of *hns*. *Biochimie* 76:1090-1098.
- Brennan, R.G., S.L. Roderick, Y. Takeda, and B.W. Matthews. 1990. Protein-DNA conformational changes in the crystal structure of a lambda-Cro-operator complex. *Proc. Natl. Acad. Sci. USA* 87:8165-8169.
- Brinkman, F.S., Schoofs, G., Hancock, R.E., and De Mot, R. 1999. Influence of a putative ECF sigma factor on expression of the major outer membrane protein, OprF, in *Pseudomonas aeruginosa* and *Pseudomonas fluorescens*. *J. Bacteriol.* 181: 4746–4754.
- Brown, M.H., and R.A. Skurray. 2001. Staphylococcal multidrug efflux protein QacA. *J. Mol. Biotech.* 3:163-170.
- Bucher, P., and A. Bairoch. 1994. A generalized profile syntax for biomolecular sequence motifs and its function in automatic sequence interpretation. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2:53-61.
- Bucher, P., K. Karplus, N. Moeri, and K. Hofmann. 1996. A flexible motif search technique based on generalized profiles. *Comput. Chem* 20:3-24.
- Buck, M., Gallegos, M.T., Studholme, D.J., Guo, Y., and Gralla, J.D. 2000. The bacterial enhancer-dependent sigma 54 transcription factor. *J. Bacteriol.* 182: 4129–4136.
- Budarina, Z.I., D.V. Nikitin, N. Zenkin, M. Zakharova, E. Semenova, M.G. Shlyapnikov, E.A. Rodikova, S. Masyukova, O. Ogarkov, G.E. Baida, A.S. Solonin, and K. Severinov. 2004. A new *Bacillus cereus*, DNA-binding protein, HlyIIR, negatively regulates expression of *B. cereus* haemolysin II. *Microbiology* 150:3691-3701.
- Burger, M., Woods, R.G., McCarthy, C., and Beacham, I.R. 2000. Temperature regulation of protease in *Pseudomonas fluorescens* LS107d2 by an ECF sigma factor and a transmembrane activator. *Microbiology* 12: 3149–3155.
- Burkovski, A. 2003. I do it my way: Regulation of ammonium uptake and ammonium assimilation in *Corynebacterium glutamicum*. *Arch. Microbiol.* 179:83- 88.
- Busenlehner, L.S., M.A. Pennella, and D.P. Giedroc. 2003. The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. *FEMS Microbiol. Rev.* 27:131-143.
- Caballero, J.L., F. Malpartida, and D.A. Hopwood. 1991. Transcriptional organization and regulation of an antibiotic export complex in the producing *Streptomyces* culture. *Mol. Gen. Genet.* 228:372-380.
- Cao, J.G., and E.A. Meighen. 1989. Purification and structural identification of an autoinducer for the luminescence system of *Vibrio harveyi*. *J. Biol. Chem.* 264:21670-2166.
- Carmona, M., Rodríguez, M.J., Martínez-Costa, O., and de Lorenzo, V. 2000. In vivo and in vitro effects of (p)ppGpp on the σ^{54} promoter Pu of the TOL plasmid of *Pseudomonas putida*. *J. Bacteriol.* 182:4711–4718.
- Cases, I., and de Lorenzo, V. 2000. Genetic evidence of distinct physiological regulation mechanisms in the σ^{54} Pu promoter of *Pseudomonas putida*. *J. Bacteriol.* 182: 956–960.
- Cases, I. de Lorenzo, V. 2005. Promoters in the environment: transcriptional regulation in its natural context. *Nat. Rev. Microbiol.* 3:105-18.

- Cases, I. Pérez-Martín, J. and de Lorenzo, V. 1999. The IANtr PtsN. protein of *Pseudomonas putida* mediates the C source inhibition of the σ^{54} -dependent Pu promoter of the TOL plasmid. *J. Biol. Chem.* 274: 15562–15568.
- Cases, I. de Lorenzo, V. and C.A. Ouzounis. 2003. Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol.* 11:248-253.
- Cases, I. Ussery, D.W. de Lorenzo, V. 2003b. The sigma54 regulon (sigmulon) of *Pseudomonas putida*. *Environ. Microbiol.* 5:1281-93.
- Chan, Y.Y., T.M.C. Tan, Y.M. Ong, and K.L. Chua. 2004. BpeAB-OprB, a multidrug efflux pump in *Burkholderia pseudomallei*. *Antimicrob. Agents Chemother.* 48:1128-1135.
- Chandler, M.S. 1992. The gene encoding cAMP receptor protein is required for competence development in *Haemophilus influenzae* Rd. *Proc. Natl. Acad. Sci. USA* 89:1626-1230.
- Chater, K.F. 1993. Genetics of differentiation in *Streptomyces*. *Annu. Rev. Microbiol.* 47:685-713.
- Chater, K.F., and C.J. Bruton. 1985. Resistance, regulatory and production genes for the antibiotic methylenomycin are clustered. *EMBO J.* 4:1893-1897.
- Chatterjee, J., C.M. Miyamoto and E.A. Meighen. 1996. Autoregulation of LuxR: the *Vibrio harveyi lux*-operon activator functions as a repressor. *Mol. Microbiol.* 20:415-425.
- Chen, X., S. Schauder, N. Potier, A. Van Dorsselaer, I. Pelczer, B.L. Bassler, and F.M. Hughson. 2002. Structural identification of a bacterial quorum-sensing signal containing boron. *Nature* 415:545-549.
- Cheng, J.C., C.P. Shao, and L.I. Hor. 1996. Cloning and nucleotide sequencing of the protease gene of *Vibrio vulnificus*. *Gene* 183:255-257.
- Cho, S.H., K. Naber, J. Hacker and W. Ziebuhr. 2002. Detection of the *icaADBC* gene cluster and biofilm formation in *Staphylococcus epidermidis* isolates from catheter-related urinary tract infections. *Int. J. Antimicrob. Agents* 19:570-575.
- Cho, Y.H., E.J. Kim, H.J. Chung, J.H. Choi, K.F. Chater, B.E. Ahn, J.H. Shin, and J.H. Roe. 2003. The *pqrAB* operon is responsible for paraquat resistance in *Streptomyces coelicolor*. *J. Bacteriol.* 185:6756-6763.
- Choi, E.N., M.C. Cho, Y. Kim, C.K. Kim, and K. Lee. 2003. Expansion of growth substrate range in *Pseudomonas putida* F1 by mutations in both *cymR* and *todS*, which recruit a ring-fission hydrolase CmtE and induce the *tod* catabolic operon, respectively. *Microbiology* 149:795-805.
- Choi, H.K., N.Y. Park, D.I. Kim, H.J. Chung, S. Ryu and S.H. Choi. 2002. Promoter analysis and regulatory characteristics of *vvhBA* encoding cytolytic hemolysin of *Vibrio vulnificus*. *J. Biol. Chem.* 277:47292-47299.
- Cohen, S.P., H. Hachler, and S.B. Levy. 1993. Genetic and functional analysis of the multiple antibiotic resistance *mar*. locus in *Escherichia coli*. *J. Bacteriol.* 175:1484–1492.
- Compan, I., and D. Touati. 1994. Anaerobic activation of *arcA* transcription in *Escherichia coli*: roles of Fnr and ArcA. *Mol. Microbiol.* 11:955–964.
- Conlon, K.M., H. Humphreys and J.P. O’Gara. 2002a. Regulation of *icaR* gene expression in *Staphylococcus epidermidis*. *FEMS Microbiol. Lett.* 216:171-177.
- Conlon, K.M., H. Humphreys and J.P. O’Gara. 2002b. *icaR* encodes a transcriptional repressor involved in environmental regulation of *ica* operon expression and biofilm formation in *Staphylococcus epidermidis*. *J. Bacteriol.* 184:4400-4408.
- Corpet F., J. Gouzy, and D. Kahn. 1999. Recent improvements of the Prodom database of protein domain families. *Nucl. Acids Res.* 27:263-267.
- Cousin Jr. S.L., W.L.H. Whittington, and M.C. Roberts. 2003. Acquired macrolide resistance genes and the 1 bp deletion in the *mtrR* promoter in *Neisseria gonorrhoeae*. *J. Antimicrob. Chemother.* 51:131-133.
- Cramton, S.E., C. Gerke, N.F. Schnell, W.W. Nichols and F. Gotz. 1999. The intercellular adhesion *ica*. locus is present in *Staphylococcus aureus* and is required for biofilm formation. *Infect. Immun.* 67:5427-5433.

- Croxatto, A., V.J. Chalker, J. Lauritz, J. Jass, A. Hardman, P. Williams, M. Camara and D.L. Milton. 2002. VanT, a homologue of *Vibrio harveyi* LuxR, regulates serine, metalloprotease, pigment, and biofilm production in *Vibrio anguillarum*. *J. Bacteriol.* 184:1617-1629.
- de Lorenzo, V., and H.P. Schweizer. 2004. Molecular tools for genetic analysis of Pseudomonads. In: J.L. Ramos ed.. *Pseudomonas Volume I*, J.L. Ramos Ed... Kluwer, London, U.K., pp 317-350.
- Delahay, R.M., B.D. Robertson, J.T. Balthazar, W.M. Shafer, and C.A. Ison. 1997. Involvement of the gonococcal MtrE protein in the resistance of *Neisseria gonorrhoeae* to toxic hydrophobic agents. *Microbiology* 143:2127-2133.
- Delany, I., R. Ieva, C. Alaimo, R. Rappuoli, and V. Scarlato. 2003. The iron-responsive regulator Fur is transcriptionally autoregulated and not essential in *Neisseria meningitidis*. *J. Bacteriol.* 185:6032-6041.
- Derzelle, S., E. Turlin, E. Duchaud, S. Pages, F. Kunst, A. Givaudan, and A. Danchin. 2004. The PhoP-phoQ two-component regulatory system of *Photorhabdus luminescens* is essential for virulence in insects. *J. Bacteriol.* 186:1270-1279.
- DeVries, C.A., and Ohman, D.E. 1994. Mucoid-to-nonmucoid conversion in alginate-producing *Pseudomonas aeruginosa* often results from spontaneous mutations in *algT*, encoding a putative alternate sigma factor, and shows evidence for autoregulation. *J. Bacteriol.* 176: 6677-6687.
- de Wetering, C. Heinzen, B. Jolivet, U. Sequin, D. Aubel, C.J. Thompson, and M. Fussenegger. 2003. *Streptomyces*-derived quorum-sensing systems engineered for adjustable transgene expression in mammalian cells and mice. *Nucl. Acids Res.* 31:71.
- Dingermann, T., U. Frank-Stoll, H. Wener, A. Wissmann, W. Hillen, M. Jacquet, and R. Marschalek. 1992. RNA polymerase III catalysed transcription can be regulated in *Saccharomyces cerevisiae* by the bacterial tetracycline repressor-operator system. *EMBO J.* 11:1487-1492.
- Ditty, J.L., Grimm, A.C., and Harwood, C.S. 1998. Identification of a chemotaxis gene region from *Pseudomonas putida*. *FEMS Microbiol Lett* 159: 267-273.
- Dogra, C., V. Raina, R. Pal, M. Suar, S. Lal, K.H. Gartemann, C. Holliger, J.R. van der Meer, and R. Lal. 2004. Organization of *lin* genes and *IS6100* among different strains of hexachlorocyclohexane-degrading *Sphingomonas paucimobilis*: evidence for horizontal gene transfer. *J. Bacteriol.* 186:2225-2235.
- Doull, J.L., A.K. Singh, M. Hoare, and S.W. Ayer. 1994. Conditions for the production of jadomycin B by *Streptomyces venezuelae* ISP5230: effects of heat shock, ethanol treatment and phage infection. *J. Ind. Microbiol.* 13:120-125.
- Dover, L.G., P.E. Corsino, I.R. Daniels, S.L. Cocklin, V. Tatituri, G.S. Besra, and K. Fütterer. 2004. Crystal structure of the TetR/CamR family repressor *Mycobacterium tuberculosis* EthR implicated in ethionamide resistance. *J. Mol. Biol.* 340:1095- 1105.
- Duetz, W.A., Marqués, S., de Jong, C., Ramos, J.L., and van Andel, J.G. 1994. Inducibility of the TOL catabolic pathway in *Pseudomonas putida* (pWWO) growing on succinate in continuous culture: evidence of carbon catabolite repression control. *J. Bacteriol.* 176: 2354-2361.
- Duetz, W.A., Marqués, S., Wind, B., Ramos, J.L., and van Andel, J.G. 1996. Catabolite repression of the toluene degradation pathway in *Pseudomonas putida* harboring pWWO under various conditions of nutrient limitation in chemostat culture. *Appl Environ Microbiol* 62: 601-606.
- Dugan, J., D.D. Rockey, L. Jones, and A.A. Andersen. 2004. Tetracycline resistance in *Chlamydia suis* mediated by genomic islands inserted into the chlamydial *inv*-like gene. *Antimicrob. Agents Chemother.* 48:3989-3995.
- Duque, E., A. Segura, G. Mosqueda, and J.L. Ramos. 2001. Global and cognate regulators control the expression of the organic solvent efflux pumps TtgABC and TtgDEF of *Pseudomonas putida*. *Mol. Microbiol.* 39:1100-1106.
- Eaton, R.W. 1996. *p*-Cumate catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA carrying the *cmt* operon. *J. Bacteriol.* 178:1351-1362.
- Eaton, R.W. 1997. *p*-Cymene catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA encoding conversion of *p*-cymene to *p*-cumate. *J. Bacteriol.* 179:3171-3180.
- Egan, S. 2002. Growing repertoire of AraC/XylS activators. *J. Bacteriol.* 184:5529-5532.

- El-Hajj, H.H., H. Zhang, and B. Weiss. 1988. Lethality of a *dut* (deoxyuridine triphosphatase) mutation in *Escherichia coli*. *J. Bacteriol.* 170:1069-1075.
- Engel, P., L.L. Scharfenstein, J.M. Dyer, and J.W. Cary. 2001. Disruption of a gene encoding a putative gamma-butyrolactone-binding protein in *Streptomyces tendae* affects nikkomycin production. *Appl. Microbiol. Biotechnol.* 56:414-419.
- English, N., C.N. Palmer, W.L. Alworth, L. Kang, V. Hughes, and C.R. Wolf. 1997. Fatty acid signals in *Bacillus megaterium* are attenuated by cytochrome P-450-mediated hydroxylation. *Biochem. J.* 327:363-368.
- English, N., V. Hughes, and C.R. Wolf. 1994. Common pathways of cytochrome P450 gene regulation by peroxisome proliferators and barbiturates in *Bacillus megaterium* ATCC14581. *J. Biol. Chem.* 269:26836-26841.
- English, N., V. Hughes, and C.R. Wolf. 1996. Induction of cytochrome P-450 BM-3 CYP 102. by non-steroidal anti-inflammatory drugs in *Bacillus megaterium*. *Biochem. J.* 316:279-283.
- Engohang-Ndong, J., Baillat, D., Aumercier, M., Bellefontaine, F., Besra, G.S., Locht, C., and Baulard, A.R. 2004. EthR, a repressor of the TetR/CamR family implicated in ethionamide resistance in mycobacteria, octamerizes cooperatively on its operator. *Mol. Microbiol.* 51:175-188.
- Enz, S., Mahren, S., Stroehner, U.H., and Braun, V. 2000. Surface signaling in ferric citrate transport gene induction: Interaction of the FecA, FecR and FecI regulatory proteins. *J. Bacteriol.* 182: 637-646.
- Eshoo, M.W. 1988. lac fusion analysis of the bet genes of *Escherichia coli*: regulation by osmolarity, temperature, oxygen, choline, and glycine betaine. *J. Bacteriol.* 170:5208-5215.
- Espinosa-Urgel, M., Kolter, R., and Ramos, J.L. 2002. Root colonization by *Pseudomonas putida*: love at first sight. *Microbiology* 148: 341-343.
- Everett, M., Y.F. Jin, V. Ricci, and L.J.V. Piddock. 1996. Contribution of individual mechanisms to fluoroquinolone resistance in 36 *Escherichia coli* strains isolated from humans and animals. *Antimicrob. Agents Chemother.* 40:2380-2386.
- Farrow, K.A., D. Lyras, and J.I. Rood. 2001. Genomic analysis of the erythromycin resistance element Tn5398 from *Clostridium difficile*. *Microbiology* 147:2717-2728.
- Faust, B., D. Hoffmeister, G. Weitnauer, L. Westrich, S. Haag, P. Schneider, H. Decker, E. Kunzel, J. Rohr, and A. Bechthold. 2000. Two new tailoring enzymes, a glycosyltransferase and an oxygenase, involved in biosynthesis of the angucycline antibiotic urdamycin A in *Streptomyces fradiae* Tu2717. *Microbiology* 146:147-154.
- Felscher, D.W., and J.M. Bishop. 1999. Reversible tumorigenesis by MYC in hematopoietic lineages. *Mol. Cell.* 2:451-461.
- Fernández-Moreno, M.A., J.L. Caballero, D.A. Hopwood, and F. Malpartida. 1991. The act cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *blaD* tRNA gene of *Streptomyces*. *Cell* 66:769-780.
- Fidopiastis, P.M., C.M. Miyamoto, M.G. Jobling, E.A. Meighen and E.G. Ruby. 2002. LitR, a new transcriptional activator in *Vibrio fischeri*, regulates luminescence and symbiotic light organ colonization. *Mol. Microbiol.* 45:131-143.
- Fischer, G.H., S.I. Wellen, D. Klimstra, J.M. Lenczowski, J.W. Tichelaar, M.J. Lizak, J.A. Whitsett, A. Korestsky, and H.E. Varmus. 2001. Induction and apoptotic regression of lung adenocarcinomas by regulation of a K-Ras transgene in the presence and absence of tumor suppressor genes. *Genes and Dev.* 15:3249-3262.
- Folcher, M., R.P. Morris, G. Dale, K. Salah-Bey-Hocini, P.H. Viollier, and C.J. Thompson. 2001. A transcriptional regulator of a pristinamycin resistance gene in *Streptomyces coelicolor*. *J. Biol. Chem.* 276:1479-1485.
- Fomenko, D.E., A.Z. Metlitskaya, J. Péduzzi, C. Goulard, G.S. Katrukha, L.V. Gening, S. Rebuffat, and I.A. Khmel. 2003. Microcin C51 plasmid genes: possible source of horizontal gene transfer. *Antimicrob. Agents Chemother.* 47:2868-2874.
- Fong, S.T., J. Camakaris, and B.T. Lee. 1995. Molecular genetics of a chromosomal locus involved in copper tolerance in *Escherichia coli* K-12. *Mol. Microbiol.* 15:1127-1137.

- Fralick, J.A. 1996. Evidence that TolC is required for functioning of the Mar/AcrAB efflux pump of *Escherichia coli*. *J. Bacteriol.* 178:5803-5805.
- Friedberg, D., M. Midkiff M., and J.M. Calvo. 2001. Global versus local regulatory roles for Lrp- related proteins: *Haemophilus influenzae* as a case study. *J. Bacteriol.* 183:4004-4011.
- Frisk, A., Jyot, J., Arora, S.K., and Ramphal, R. 2002. Identification and functional characterization of flgM, a gene encoding the anti-sigma 28 factor in *Pseudomonas aeruginosa*. *J. Bacteriol.* 184: 1514–1521.
- Fromknecht, K., P.D. Vogel, and J.G. Wise. 2003. Combinatorial redesign of the DNA binding specificity of a prokaryotic helix-turn-helix repressor. *J. Bacteriol.* 185: 475-481.
- Fujita, M., H. Aramaki, T. Horiuchi, and A. Amemura. 1993. Transcription of the cam operon and *camR* genes in *Pseudomonas putida* PpG1. *J. Bacteriol.* 175:6953-6958.
- Fuqua, C., S.C. Winans, and E.P. Greenberg. 1996. Census and consensus in bacterial ecosystems: the LuxR-LuxI family of quorum-sensing transcriptional regulators. *Annu. Rev. Microbiol.* 50:727-751.
- Fuqua, W.C., S.C. Winans, and E.P. Greenberg. 1994. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J. Bacteriol.* 176:269-275.
- Gallegos, M.T., C. Michán, and J.L. Ramos. 1993. The XylS/AraC family of regulators. *Nucl. Acids Res.* 21:807–810.
- Gallegos, M.T., R. Schleif, A. Bairoch, K. Hofmann, and J.L. Ramos. 1997. AraC/XylS of transcriptional regulators. *Microbiol. Mol. Biol. Rev.* 61:393-410.
- Gambino, L., S.J. Gracheck, and P.F. Miller. 1993 Overexpression of the MarA positive regulator is sufficient to confer multiple antibiotic resistance in *Escherichia coli*. *J. Bacteriol.* 175:2888–2894.
- Garnier, T., K. Eiglmeier, J.C. Camus, N. Medina, H. Mansoor, M. Pryor, S. Duthoy, S. Grondin, C. Lacroix, C. Monsempe, S. Simon, B. Harris, R. Atkin, J. Doggett, R. Mayes, L. Keating, P.R. Wheeler, J. Parkhill, B.G. Barrell, S.T. Cole, S.V. Gordon, and R.G. Hewinson. 2003. The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci. USA* 100:7877-7882.
- Garrett, E.S., Perlegas, D., and Wozniak, D.J. 1999. Negative control of flagellum synthesis in *Pseudomonas aeruginosa* is modulated by the alternative sigma factor AlgT AlgU. *J. Bacteriol.* 181: 7401–7404.
- Gattiker, A., Michoud, K., Rivoire, C., Auchincloss, A.H., Coudert, E., Lima, T., Kersey, P., Pagni, M., Sigrist, C.J., Lachaize, C. et al. 2003. Automated annotation of microbial proteomes in SWISS-PROT. *Comput. Biol. Chem.*, 27, 49–58.
- George, A.M., and S.B. Levy. 1983. Gene in the major cotransduction gap of the *Escherichia coli* K-12 linkage map required for the expression of chromosomal resistance to tetracycline and other antibiotics. *J. Bacteriol.* 178:2507-2513.
- González-Pasayo, R., and E. Martínez-Romero. 2000. Multiresistance genes of *Rhizobium etli* CFN42. *Mol. Plant Microbe. Interact.* 13:572-577.
- Gossen, M., A.L. Bonin, and H. Bujard. 1993. Control of gene activity in higher eukaryotes cells by prokaryotic regulatory elements. *Trends Biochem. Sci.* 18:471-475.
- Gossen, M., and H. Bujard. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA* 89:5547-5551.
- Gotz, F. 2002. *Staphylococcus* and biofilms. *Mol. Microbiol.* 43:1367-1378.
- Gray, K.M., and J.R. Garey. 2001. The evolution of bacterial LuxI and LuxR quorum sensing regulators. *Microbiology* 147:2379-2387.
- Green, J., C. Scott, and J.R. Guest. 2001. Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. *Adv. Microb. Physiol.* 44:1-34.
- Grkovic, S., K.M. Hardie, M.H. Brown, and R.A. Skurray. 2003. Interactions of the QacR multidrug-binding protein with structurally diverse ligands: implications for the evolution of the binding pocket. *Biochemistry* 42:15226-15236.

- Grkovic, S., M.H. Brown, and R.A. Skurray. 2001. Transcriptional regulation of multidrug efflux pumps in bacteria. *Cell Develop. Biol.* 12:225-237.
- Grkovic, S., M.H. Brown, M.A. Schumacher, R.G. Brennan, and R.A. Skurray. 2001. The Staphylococcal QacR multidrug regulator binds a correctly spaced operator as a pair of dimers. *J. Bacteriol.* 183:7102-7109.
- Grkovic, S., M.H. Brown, N.J. Roberts, I.T. Paulsen, and R.A. Skurray. 1998. QacR is a repressor protein that regulates expression of the *Staphylococcus aureus* multidrug efflux pump QacA. *J. Biol. Chem.* 273:18665-18673.
- Grogan, G., G.A. Roberts, D. Bougioukou, N.J. Turner, and S.L. Flitsch. 2001. The desymmetrization of bicyclic beta-diketones by enzymatic retro-claisen reaction. A new reaction of the crotonase superfamily. *J. Biol. Chem.* 276:12565-12572.
- Grohmann, E., G. Muth, and M. Espinosa. 2003. Conjugative plasmid transfer in gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* 67:277-301.
- Grossman, A.D., Straus, D.B., Walter, W.A., and Gross, C.A. 1987. Sigma 32 synthesis can regulate the synthesis of heat shock proteins in *Escherichia coli*. *Genes Dev* 1: 179-184.
- Guengerich, F.P. 1991. Reactions and significance of cytochrome P-450 enzymes. *J. Biol. Chem.* 266:10019-10022.
- Guilfoile, P.G., and C.R. Hutchinson. 1992. Sequence and transcriptional analysis of the *Streptomyces glaucescens tcmAR* tetracenomycin resistance and repressor gene loci. *J. Bacteriol.* 174:3651-3658.
- Guilfoile, P.G., C.R. Hutchinson. 1992. The *Streptomyces glaucescens TcmR* protein represses transcription of the divergently oriented *tcmR* and *tcmA* genes by binding to an intergenic operator region. *J. Bacteriol.* 174:3659-3666.
- Haft, D.H., B.J. Loftus, D.L. Richardson, F. Yang, J.A. Eisen, I.T. Paulsen, and O. White. 2001. TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucl. Acids. Res.* 29:41-43.
- Hagen, K.D., and J.C. Meeks. 1999. Biochemical and genetic evidence for participation of DevR in a phosphorelay signal transduction pathway essential for heterocyst maturation in *Nostoc punctiforme* ATCC 29133. *J. Bacteriol.* 181:4430-4434.
- Hagman, K.E., and W.M. Shafer. 1995. Transcriptional control of the *mtr* efflux system of *Neisseria gonorrhoeae*. *J. Bacteriol.* 177:4162-4165.
- Hagman, K.E., W. Pan, B.G. Spratt, J.T. Balthazar, R.C. Judd, and W.M. Shafer. 1995. Resistance of *Neisseria gonorrhoeae* to antimicrobial agents is modulated by the *mtrRCDE* efflux system. *Microbiology* 141:611-622.
- Hansson, M., and L. Hederstedt. 1992. Cloning and characterization of the *Bacillus subtilis hemEHY* gene cluster, which encodes protoheme IX biosynthetic enzymes. *J. Bacteriol.* 174:8081-8093.
- Hara, O., and T. Beppu. 1982. Induction of streptomycin-inactivating enzyme by A-factor in *Streptomyces griseus*. *J. Antibiot.* 35:1208-1215.
- Hara, O., and T. Beppu. 1982. Mutants blocked in streptomycin production in *Streptomyces griseus* — the role of A-factor. *J. Antibiot.* 35:349-358.
- Harrison, S.C. 1991. A structural taxonomy of DNA-binding domains. *Nature* 353:715-719.
- Harrison, S.C., and A.K. Aggarwal. 1990. DNA recognition by proteins with the helix-turn-helix motif. *Annu. Rev. Biochem.* 59:933-969.
- Hashimoto, Y., M. Yamashita, and Y. Murooka. 1997. The *Propionibacterium freudenreichii hemYHBXRL* gene cluster, which encodes enzymes and a regulator involved in the biosynthetic pathway from glutamate to protoheme. *Appl. Microbiol. Biotechnol.* 47:385-392.
- Haydon, D.J., and J.R. Guest. 1991. A new family of bacterial regulatory proteins. *FEMS Microbiol. Lett.* 63: 291-295.
- Hayes, A., G. Hobbs, C.P. Smith, S.G. Oliver, and P.R. Butler. 1997. Environmental signals triggering methylenomycin production by *Streptomyces coelicolor* A3(2). *J. Bacteriol.* 179:5511-5515.

- He, J.S., Q. Liang, and A.J. Fulco. 1995. The molecular cloning and characterization of BM1P1 and BM1P2 proteins, putative positive transcription factors involved in barbiturate-mediated induction of the genes encoding cytochrome P450BM-1 of *Bacillus megaterium*. *J. Biol. Chem.* 270:18615-18625.
- Helbl, V., and W. Hillen. 1998. Stepwise selection of TetR variants recognizing *tet* operator 4C with high affinity and specificity. *J. Mol. Biol.* 276:313-318.
- Helbl, V., B. Tiebel, and W. Hillen. 1998. Stepwise selection of TetR variants recognizing *tet* operator 6C with high affinity and specificity. *J. Mol. Biol.* 276:319-324.
- Helbl, W., C. Berens, and W. Hillen. 1995. Proximity probing of Tet repressor to *tet* operator by dimethylsulfate reveals protected and accessible functions for each recognized base-pair in the major groove. *J. Mol. Biol.* 245:538-548.
- Helmann, J.D, B.T. Ballard, and C.T. Walsh. 1990. The MerR metaloregulatory protein binds mercuric ion as a tricoordinate, metal-bridged dimer. *Science* 247:946-948.
- Henikoff, S., G.W. Haughn, J.M. Calvo, and J.C. Wallace. 1988. A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. USA* 85:6602-6606.
- Henikoff, S., J.C. Wallace, and J.P. Brown. 1990. Finding protein similarities with nucleotide sequence databases. *Meth. Enzymol.* 183:111-132.
- Heuer, H., R. Szczepanowski, S. Schneiker, A. Pühler, E.M. Top, and A. Schlüter. 2004. The complete sequences of plasmids pB2 and pB3 provide evidence for a recent ancestor of the IncP-1 β group without any accessory genes. *Microbiology* 150:3591-3599.
- Hillen, W., and C. Berens. 1994. Mechanisms underlying expression of Tn10 encoded tetracycline resistance. *Ann. Rev. Microbiol.* 48:345-369.
- Hillen, W., K. Schollmeier, and C. Gatz. 1984. Control of expression of the Tn10-encoded tetracycline resistance operon. II. Interaction of RNA polymerase and Tet repressor with the *tet* operon regulatory region. *J. Mol. Biol.* 172:185-201.
- Hinrichs, W., C. Kisker, M. Düvel, A. Müller, K. Tovar, W. Hillen, and W. Saenger. 1994. Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. *Science* 264:418-420.
- Hirata, T., A. Saito, K. Nishino, N. Tamura, and A. Yamaguchi. 2004. Effects of efflux transporter genes on susceptibility of *Escherichia coli* to tigecycline GAR-936. *Antimicrob. Agents Chemother.* 48:2179-2184.
- Hitomi, M., H. Nishimura, Y. Tsujimoto, H. Matsui, and K. Watanabe. 2003. Identification of a helix-turn-helix motif of *Bacillus thermoglucosidasius* HrcA essential for binding to the CIRCE element and thermostability of the HrcA-CIRCE complex, indicating a role as a thermosensor. *J. Bacteriol.* 185:381-385.
- Hobbs, G., A.I. Obanye, J. Petty, J.C. Mason, E. Barratt, D.C. Gardner, F. Flett, C.P. Smith, P. Broda, and S.G. Oliver. 1992. An integrated approach to studying regulation of production of the antibiotic methylenomycin by *Streptomyces coelicolor* A3(2). *J. Bacteriol.* 174:1487-1494.
- Hofmann K., P. Bucher, L. Falquet, and A. Bairoch. 1999. The Prosite database, its status in 1999. *Nucl. Acids Res.* 27:215-219.
- Holtel, A., Marqués, S., Mohler, I., Jakubzik, U., and Timmis, K.N. 1994. Carbon source-dependent inhibition of *xyl* operon expression of the *Pseudomonas putida* TOL plasmid. *J. Bacteriol.* 176: 1773-1776.
- Horinouchi, S. 2002. A microbial hormone, A-factor, as a master switch for morphological differentiation and secondary metabolism in *Streptomyces griseus*. *Front. Biosci.* 7:2045-2057.
- Horinouchi, S., and T. Beppu. 1990. Autoregulatory factors of secondary metabolism and morphogenesis in actinomycetes. *Crit. Rev. Biotechnol.* 10:191-204.
- Horinouchi, S., and T. Beppu. 1994. A-factor as a microbial hormone that controls cellular differentiation and secondary metabolism in *Streptomyces griseus*. *Mol. Microbiol.* 12:859-864.
- Horinouchi, S., Y. Ohnishi, and D.K. Kang. 2001. The A-factor regulatory cascade and cAMP in the regulation of physiological and morphological development in *Streptomyces griseus*. *J. Ind. Microbiol. Biotech.* 27:177-182.
- Inouye, S., Yamada, M., Nakazawa, A., and Nakazawa, T. 1989. Cloning and sequence analysis of the *ntrA* (*rpoN*) gene of *Pseudomonas putida*. *Gene* 85: 145-152.

- Isackson, P.J., and K.P. Bertrand. 1985. Dominant negative mutations in the Tn10 Tet repressor: evidence for use of the conserved helix-turn-helix motif in DNA binding. *Proc. Natl. Acad. Sci. USA* 82:6226-6230.
- Ishihama, A. 2000. Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol* 54: 499–518.
- Itoh, K., Y. Tashiro, K. Uobe, Y. Kamagata, K. Suyama, and H. Yamamoto. 2004. Root nodule *Bradyrhizobium* spp. harbor *tfdAa* and *cadA*, homologous with genes encoding 2,4-dichlorophenoxyacetic acid-degrading proteins. *Appl. Environ. Microbiol.* 70:2110-2118.
- Jakoby, M., L. Nolden, J. Meier-Wagner, R. Kramer, and A. Burkovski. 2000. AmtR, a global repressor in the nitrogen regulation system of *Corynebacterium glutamicum*. *Mol. Microbiol.* 37:964- 977.
- Janssen, D.B., F. Pries, J. van der Ploeg, B. Kazemier, P. Terpstra, and B. Witholt. 1989. Cloning of 1,2-dichloroethane degradation genes of *Xanthobacter autotrophicus* GJ10 and expression and sequencing of the *dhIA* gene. *J. Bacteriol.* 171:6791- 6799.
- Jefferson, K.K., S.E. Cramton, F. Gotz, and G.B. Pier. 2003. Identification of a 5-nucleotide sequence that controls expression of the *ica* locus in *Staphylococcus aureus* and characterization of the DNA-binding properties of IcaR. *Mol. Microbiol.* 48:889-899.
- Jellen-Ritter, A.S., and W.V. Kern. 2001. Enhanced expression of the multidrug efflux pumps AcrAB and AcrEF associated with insertion element transposition in *Escherichia coli* mutants selected with a fluoroquinolone. *Antimicrob. Agents Chemother.* 45:1467-1472.
- Jobling, M.G. and R.K. Holmes. 1997. Characterization of *hapR*, a positive regulator of the *Vibrio cholerae* HA/protease gene *hap*, and its identification as a functional homologue of the *Vibrio harveyi* LuxR gene. *Mol. Microbiol.* 26:1023- 1034.
- Jogun, K.H., and J.J. Stezowski. 1976. Coordination and conformational aspects of oxytetracycline metal ion complexation. *J. Am. Chem. Soc.* 98:6018-6026.
- Johnston, T.C., R.B. Thompson and T.O. Baldwin. 1986. Nucleotide sequence of the *luxB* gene of *Vibrio harveyi* and the complete amino acid sequence of the beta subunit of bacterial luciferase. *J. Biol. Chem.* 261:4805-4811.
- Kamionka, A., J. Bogdanska-Urbaniak, O. Scholz, and W. Hillen. 2004. Two mutations in the tetracycline repressor change the inducer anhydrotetracycline to a corepressor. *Nucl. Acids Res.* 32:842-847.
- Kaneko, M., A. Yamaguchi, and T. Sawai. 1985. Energetics of tetracycline efflux system encoded by Tn10 in *Escherichia coli*. *FEBS Lett.* 193:893-899.
- Karniol, B., and R.D. Vierstra. 2004. The HWE histidine kinases, a new family of bacterial two-component sensor kinases with potentially diverse roles in environmental signaling. *J. Bacteriol.* 186:445-453.
- Kasuga, K., H. Nojiri, H. Yamane, T. Kodama, and T. Omori. 1997. Cloning and characterization of genes involved in the degradation of dibenzofuran by *Terrabacter* sp. strain DBF63. *J. Ferment. Bioeng.* 84:387-399.
- Katayama, J., and N. Noguchi. 1999. Nucleotide sequence of the gene cluster containing the *mphB* gene for macrolide 2'-phosphotransferase II. *Biol. Pharm. Bull.* 22:227-228.
- Kato, J., A. Suzuki, H. Yamazaki, Y. Ohnishi, and S. Horinouchi. 2002. Control by A-factor of a metalloendopeptidase gene involved in aerial mycelium formation in *Streptomyces griseus*. *J. Bacteriol.* 184:6010-6025.
- Kato, J.-Y., I. Miyahisa, M. Mashiko, Y. Ohnishi, and S. Horinouchi. 2004. A single target is sufficient to account for biological effects of the A-factor receptor protein of *Streptomyces griseus*. *J. Bacteriol.* 186:2206-221.
- Kawachi, R., T. Akashi, Y. Kamitani, A. Sy, U. Wangchaisoonthorn, T. Nihira, and Y. Yamada. 2000. Identification of an AfsA homologue BarX, from *Streptomyces virginiae* as a pleiotropic regulator controlling autoregulator biosynthesis, virginiamycin biosynthesis and virginiamycin M1 resistance. *Mol. Microbiol.* 36:302-313.
- Kawamura-Sato, K., K. Shibayama, T. Horii, Y. Iimuma, Y. Arakawa, and M. Ohta. 1999. Role of multiple efflux pumps in *Escherichia coli* in indole expulsion. *FEMS Microbiol. Lett.* 179:345-352.

- Keith, L.M.W., and Bender, C.L. 1999. AlgT (σ_{22}) controls alginate production and tolerance to environmental stress in *Pseudomonas syringae*. *J. Bacteriol.* 181: 7176–7184.
- Kieboom, J., and J. de Bont. 2001. Identification and molecular characterization of an efflux system involved in *Pseudomonas putida* S12 multidrug resistance. *Microbiology* 147:43-51.
- Kieboom, J., Bruinenberg, R., Keizer-Gunnink, I., and de Bont, J.A.M. 2001. Transposon mutations in the flagella biosynthetic pathway of the solvent-tolerant *Pseudomonas putida* S12 result in a decreased expression of solvent efflux genes. *FEMS Microbiol. Lett.* 198: 117–122.
- Kieboom, J., J.J. Dennis, G.J. Zylstra, J.A. de Bont. 1998. Active efflux of organic solvents by *Pseudomonas putida* S12 is induced by solvents. *J. Bacteriol.* 180:6769-6772.
- Kieboom, J., J.J. Dennis, J.A. de Bont, and G.J. Zylstra. 1998. Identification and molecular characterization of an efflux pump involved in *Pseudomonas putida* S12 solvent tolerance. *J. Biol. Chem.* 273:85-91.
- Kieboom, J., R. Bruinenberg, I. Keizer-Gunnink, and J.A. de Bont. 2001. Transposon mutations in the flagella biosynthetic pathway of the solvent-tolerant *Pseudomonas putida* S12 result in a decreased expression of solvent efflux genes. *FEMS Microbiol. Lett.* 198:117-122.
- Kim, I., Stiefel, A., Plantor, S., Angerer, A., and Braun, V. 1997. Transcription induction of the ferric citrate transport genes via the N-terminus of the FecA outer membrane protein, the Ton system and the electrochemical potential of the cytoplasmic membrane. *Mol Microbiol* 23: 333–344.
- Kim, Y.B., J. Okuda, C. Matsumoto, N. Takahashi, S. Hashimoto, and M. Nishibuchi. 1999. Identification of *Vibrio parahaemolyticus* strains at the species level by PCR targeted to the *toxR* gene. *J. Clin. Microbiol.* 37:1173-1177.
- Kinoshita, H., H. Ipposhi, S. Okamoto, H. Nakano, T. Nihira, and Y. Yamada. 1997. Butyrolactone auto-regulator receptor protein (BarA) as a transcriptional regulator in *Streptomyces virginiae*. *J. Bacteriol.* 179:6986-6993.
- Kisker C., W. Hinrichs, K. Tovar, W. Hillen, and W. Saenger. 1995. The complex formed between Tet repressor and tetracycline-Mg²⁺ reveals mechanism of antibiotic resistance. *J. Mol. Biol.* 247:260-280.
- Kitani, S., H. Kinoshita, T. Nihira and Y. Yamada Y. 1999. In vitro analysis of the butyrolactone autoregulator receptor protein (FarA) of *Streptomyces lavendulae* FRI-5 reveals that FarA acts as a DNA-binding transcriptional regulator that controls its own synthesis. *J. Bacteriol.* 181:5081-5084.
- Kitani, S., Y. Yamada and T. Nihira. 2001. Gene replacement analysis of the butyrolactone autoregulator receptor (FarA) reveals that FarA acts as a novel regulator in secondary metabolism of *Streptomyces lavendulae* FRI-5. *J. Bacteriol.* 183:4357-4363.
- Klein, J.R., and R. Plapp. 1992. Locations of the *envCD* genes on the physical map of the *Escherichia coli* chromosome. *J. Bacteriol.* 174:3828-3829.
- Klein, J.R., B. Henrich, and R. Plapp. 1991. Molecular analysis and nucleotide sequence of the *envCD* operon of *Escherichia coli*. *Mol. Gen. Genet.* 230:230-240.
- Klinke, S., G. de Roo, B. Witholt, and B. Kessler. 2000. Role of *phaD* in accumulation of medium-chain-length poly3-hydroxyalkanoates. in *Pseudomonas oleovorans*. *Appl. Environ. Microbiol.* 66:3705-3710.
- Klock, G., B. Unger, C. Gatz, W. Hillen, J. Altenbuchner, K. Schmid, and R. Schmitt. 1985. Heterologous repressor-operator recognition among four classes of tetracycline resistance determinants. *J. Bacteriol.* 161:326-332.
- Knobloch, J.K., K. Bartscht, A. Sabottke, H. Rohde, H.H. Feucht and D. Mack. 2001. Biofilm formation by *Staphylococcus epidermidis* depends on functional RsbU, an activator of the *sigB* operon: differential activation mechanisms due to ethanol and salt stress. *J. Bacteriol.* 183:2624-2633.
- Kobayashi, K., N. Tsukagoshi, and R. Aono. 2001. Suppression of hypersensitivity of *Escherichia coli* *acrB* mutant to organic solvents by integrational activation of the *acrEF* operon with the IS1 or IS2 element. *J. Bacteriol.* 183:2646-2653.
- Köhler, T., Harayama, S., Ramos, J.L., and Timmis, K.N. 1989. Involvement of *Pseudomonas putida* RpoN σ factor in regulation of various metabolic functions. *J. Bacteriol.* 171: 4326–4333.
- Kojic, M., and V. Venturi V. 2001. Regulation of *rpoS* gene expression in *Pseudomonas*: involvement of a TetR family regulator. *J. Bacteriol.* 183:3712-3720.

- Koronakis, V., A. Sharff, E. Koronakis, B. Luisi, and C. Hughes. 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* 405:914-919.
- Koster, M., van Klompenburg, W., Bitter, W., Leong, J., and Wesbeek, P. 1994. Role of the outer membrane ferric siderophore receptor PupB in signal transduction across the bacterial cell envelope. *EMBO J* 13: 2805–2813.
- Kovacikova, G. and K. Skorupski. 2002. Regulation of virulence gene expression in *Vibrio cholerae* by quorum sensing: HapR functions at the *aphA* promoter. *Mol. Microbiol.* 46:1135-1147.
- Krug, A., V.F. Wendisch, and M. Bott. 2005. Identification of AcnR, a TetR-type repressor of the acotinase gene *acn* in *Corynebacterium glutamicum*. *J. Biol. Chem.* 280:585-595.
- Krukonis, E.S., R.R. Yu and V.J. Dirita. 2000. The *Vibrio cholerae* *ToxR/TcpP/ToxT* virulence cascade: distinct roles for two membrane-localized transcriptional activators on a single promoter. *Mol. Microbiol.* 38:67-84.
- Kühn, S., Braun, V., and Koster, W. 1996. Ferric rhizoferrin uptake into *Morganella morganii*: characterization of genes involved in the uptake of a polyhydroxycarboxylate siderophore. *J. Bacteriol.* 178: 496–504.
- Kumano, M., A. Tamakoshi, and K. Yamane. 1997. A 32 kb nucleotide sequence from the region of the lincomycin-resistance gene 22 degrees-25 degrees. of the *Bacillus subtilis* chromosome and identification of the site of the *lin-2* mutation. *Microbiology* 143:2775-2782.
- Kumano, M., M. Fujita, K. Nakamura, M. Murata, R. Ohki, and K. Yamane. 2003. Lincomycin resistance mutations in two regions immediately downstream of the -10 region of *Imr* promoter cause overexpression of a putative multidrug efflux pump in *Bacillus subtilis* mutants. *Antimicrob. Agents Chemother.* 47:432-435.
- Kunin, C.M., T.H. Hua, W.L. Van Arsdale, and M. Villarejo. 1992. Growth of *Escherichia coli* in human urine: role of salt tolerance and accumulation of glycine betaine. *J. Infect. Dis.* 166:1311-1315.
- Kustu, S., A.K. North, and D.S. Weiss. 1991. Prokaryotic transcriptional enhancers and enhancer-binding proteins. *Trends Biochem. Sci.* 16:397-402.
- La Teana, A., A. Brandi, M. Falconi, R. Spurio, C.L. Pon, and C.O. Gualerzi. 1991. Identification of a cold shock transcriptional enhancer of the *Escherichia coli* gene encoding nucleoid protein H-NS. *Proc. Natl. Acad. Sci. USA.* 88:10907-10911.
- Lamark, T., I. Kaasen, M.W. Eshoo, P. Falkenberg, J. McDougall, and A.R. Strom. 1995. DNA sequence and analysis of the *bet* genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. *Mol. Microbiol.* 5:1049-1064.
- Lamark, T., O.B. Styrvold, and A.R. Strøm. 1992. Efflux of choline and glycine betaine from osmoregulating cells of *Escherichia coli*. *FEMS Microbiol. Lett.* 96:149–154.
- Lamark, T., T.P. Rokenes, J. McDougall, and A.R. Strom. 1996. The complex *bet* promoters of *Escherichia coli*: regulation by oxygen (*ArcA*), choline (*BetI*), and osmotic stress. *J. Bacteriol.* 178:1655-1662.
- Lamont, I.L., Beare, P.A., Ochsner, U., Vasil, A.I., and Vasil, M.L. 2002. Siderophore-mediated signalling regulates virulence factor production in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci USA* 99: 7072–7077.
- Landfald, B., and A.R. Strom. 1986. Choline- glycine betaine pathway a high level of osmotic tolerance in *Escherichia coli*. *J. Bacteriol.* 165:849-855.
- Lau, P.C.K., Y. Wang, A. Patel, D. Labbe, H. Bergeron, R. Brousseau, Y. Konishi, and M. Rawlings. 1997. A bacterial basic region leucine zipper histidine kinase regulating toluene degradation. *Proc. Natl. Acad. Sci. USA* 94:1453-1458.
- Lauble, H., Y. Georgalis, and U. Heinemann. 1989. Studies on the domain structure of the *Salmonella typhimurium* *AraC* protein. *Eur. J. Biochem.* 185:319– 325.
- Lawson, C.L. Swigon, D. Murakami, K.S. Darst, S.A. Berman, H.M. Ebright, R.H. 2004. Catabolite activator protein: DNA binding and transcription activation. *Curr. Opin. Struct. Biol.* 14:10-20.
- Lee, E.H., C. Rouquette-Loughlin, J.P. Folster, and W.M. Shafer. 2003. *FarR* regulates the *farAB*-encoded efflux pump of *Neisseria gonorrhoeae* via an *MtrR* regulatory mechanism. *J. Bacteriol.* 185: 7145- 7152.

- Leoni, L., Ambrosi, C., Petrucca, A., and Visca, P. 2002. Transcriptional regulation of pseudobactin synthesis in the plant growth-promoting *Pseudomonas* B10. *FEMS Microbiol Lett* 208: 219–225.
- Leoni, L., Orsi, N., de Lorenzo, V., and Visca, P. 2000. Functional analysis of PvdS, an iron starvation sigma factor of *Pseudomonas aeruginosa*. *J. Bacteriol.* 182: 1481–1491.
- Levy S.B., L.M. McMurry, T.M. Barbosa, V. Burdett, P. Courvalin, W. Hillen, M.C. Roberts, J.I. Rood, and D.E. Taylor. 1999. Nomenclature for new tetracycline resistance determinants. *Antimicrob. Agents Chemother.* 43:1523-1524.
- Levy, S.B. 1984. Resistance to tetracyclines. In *Antimicrobial Drug Resistance* Bryan, L.E., ed., pp. 191-204, Academic Press, NY.
- Levy, S.B. 1988. Tetracycline resistance determinants are widespread. *ASM News* 54:418-421.
- Li, X.-Z., L. Zhang, and H. Nikaido. 2004. Efflux pump-mediated intrinsic drug resistance in *Mycobacterium smegmatis*. *Antimicrob. Agents Chemother.* 48:2415-2423.
- Liang, Q., and A.J. Fulco. 1995. Transcriptional regulation of the genes encoding cytochromes P450BM-1 and P450BM-3 in *Bacillus megaterium* by the binding of Bm3R1 repressor to Barbie box elements and operator sites. *J. Biol. Chem.* 270:18606-18614.
- Liang, Q., J.S. He, and A.J. Fulco A.J. 1995. The role of Barbie box sequences as *cis*-acting elements involved in the barbiturate-mediated induction of cytochromes P450BM-1 and P450BM-3 in *Bacillus megaterium*. *J. Biol. Chem.* 270:4438-4450.
- Liang, Q., L. Chen, and A.J. Fulco. 1998. In vivo roles of Bm3R1 repressor in the barbiturate-mediated induction of the cytochrome P450 genes P450BM-3. and P450BM-.1. of *Bacillus megaterium*. *Biochim. Biophys. Acta.* 1380:183-197.
- Lin, Y.H., C. Miyamoto, and E.A. Meighen. 2000. Purification and characterization of a *luxO* promoter binding protein LuxT from *Vibrio harveyi*. *Protein Expr. Purif.* 20:87-94.
- Lin, Z., K. Kumagai, K. Baba, J.J. Mekalanos, and M. Nishibuchi. 1993. *Vibrio parahaemolyticus* has a homolog of the *Vibrio cholerae* *toxRS* operon that mediates environmentally induced regulation of the thermostable direct hemolysin gene. *J. Bacteriol.* 175:3844-3855.
- Liu, H., and K.A. Reynolds. 1999. Role of crotonyl coenzyme A reductase in determining the ratio of polyketides monensin A and monensin B produced by *Streptomyces cinnamonensis*. *J. Bacteriol.* 181:6806-6813.
- Liu, H., and K.A. Reynolds. 2001. Precursor supply for polyketide biosynthesis: the role of crotonyl-CoA reductase. *Metab. Eng.* 3:40-48.
- Lonetto, M., Gribskov, M., and Gross, C.A. 1992. The sigma 70 family: sequence conservation and evolutionary relationship. *J. Bacteriol.* 174: 3843–3849.
- Lonetto, M.A., Brown, K.L., Rudd, K.E., and Buttner, M.J. 1994. Analysis of the *Streptomyces coelicolor* *sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase σ factors involved in the regulation of extracytoplasmatic functions. *Proc. Natl. Acad. Sci. USA* 91: 7573–7577.
- Lucas, C.E., J.T. Balthazar JT, K.E. Hagman KE, and W.M. Shafer. 1997. The *MtrR* repressor binds the DNA sequence between the *mtrR* and *mtrC* genes of *Neisseria gonorrhoeae*. *J. Bacteriol.* 179:4123-4128.
- Lucas, C.E., K.E. Hagman, J.C. Levin, D.C. Stein, and W.M. Shafer. 1995. Importance of lipooligosaccharide structure in determining gonococcal resistance to hydrophobic antimicrobial agents resulting from the *mtr* efflux system. *Mol. Microbiol.* 16:1001-1009.
- Ly, A., J. Henderson, A. Lu, D.E. Culham, and J.M. Wood. 2004. Osmoregulatory systems of *Escherichia coli*: identification of betaine- carnitine-choline transporter family member BetU and distributions of betU and *trkG* among pathogenic and nonpathogenic isolates. *J. Bacteriol.* 186:296-306.
- Ma, D., D.N. Cook, J. Hearst, and H.N. Nikaido. 1994. Efflux pumps and drug resistance in Gram-negative bacteria. *Trends Microbiol.* 2:489-493.
- Ma, D., D.N. Cook, M. Alberti, N.G. Pon, H. Nikaido, and J.E. Hearst. 1993. Molecular cloning and characterization of *acrA* and *acrE* genes of *Escherichia coli*. *J. Bacteriol.* 175:6299-6313.

- Ma, D., M. Alberti, C. Lynch, H. Nikaido, and J.E. Hearst. 1996. The local repressor AcrR plays a modulating role in the regulation of *acrAB* genes of *Escherichia coli* by global stress signals. *Mol. Microbiol.* 19:101-112.
- Ma, D., M. Alberti, C. Lynch, N.G. Pon, H. Nikaido, and J.E. Hearst. 1995. Genes *acrAB* encode a stress-induced efflux system of *Escherichia coli*. *Mol. Microbiol.* 16:45-55.
- Maeda, H., Jishage, M., Nomura, T., Fujita, N., and Ishihama, A. 2000. Two extracytoplasmic function sigma subunits, σ^E and σ^{Fecl} , of *Escherichia coli*: promoter selectivity and intracellular levels. *J. Bacteriol.* 182: 1181–1184.
- Mahren, S., and W. Braun. 2003. The Fecl extracytoplasmic-function sigma factor of *Escherichia coli* interacts with the β ' subunit of RNA polymerase. *J. Bacteriol.* 185:1796-1802.
- Mahren, S., Enz, S., and Braun, V. 2002. Functional interaction of region 4 of the extracytoplasmic function sigma factor Fecl with the cytoplasmic portion of the FecR transmembrane protein of the *E. coli* ferric citrate transport system. *J. Bacteriol.* 184: 3704–3711.
- Maness, M.J., and P.F. Sparling. 1973. Multiple antibiotic resistance due to a single mutation in *Neisseria gonorrhoeae*. *J. Infect. Dis.* 128:321-330.
- Manzanera, M., Marqués, S., and Ramos, J.L. 2001. Molecular characterization of *Pseudomonas putida* KT2440 *rpoH* gene regulation. *Microbiology* 147: 1323–1330.
- Manzanera, M., S. Marqués, and J.L. Ramos. 2000. Mutational analysis of the highly conserved C-terminal residues of the XylS protein, a member of the AraC family of transcriptional regulators. *FEBS Lett.* 476:312-317.
- Marger, M.D., and M.H. Saier Jr. 1993. A major superfamily of transmembrane facilitators that catalyze uniport, symport and antiport. *Trends Biochem. Sci.* 18:13-20.
- Marqués, S., M.T. Gallegos, M. Manzanera, A. Holtel, K.N. Timmis, and J.L. Ramos. 1998. Activation and repression of transcription at the double tandem divergent promoters for the *xylR* and *xylS* genes of the TOL plasmid of *Pseudomonas putida*. *J. Bacteriol.* 180:2889-2894.
- Marqués, S., Manzanera, M., González-Pérez, M.M., Gallegos, M.T., and Ramos, J.L. 1999b. The XylS-dependent Pm promoter is transcribed in vivo by RNA polymerase with σ^{32} or σ^{38} depending on the growth phase. *Mol. Microbiol.* 31: 1105–1113.
- Marqués, S., Manzanera, M., González-Pérez, M.M., Ruíz, R., and Ramos, J.L. 1999a. Biodegradation, plasmid encoded catabolic pathways, host factors and cell metabolism. *Environ. Microbiol.* 1: 107–108.
- Martin, M., R. Showalter and M. Silverman. 1989. Identification of a locus controlling expression of luminescence genes in *Vibrio harveyi*. *J. Bacteriol.* 171:2406-2414.
- Martínez-Argudo, I., and A. Contreras. 2002. PII T-Loop mutations affecting signal transduction to NtrB also abolish yeast two-hybrid interactions. *J. Bacteriol* 184:3746-3748.
- Martínez-Argudo, I., P. Salinas, R. Maldonado, and A. Contreras. 2002. Domain interactions on the *ntr* signal transduction pathway: two-hybrid analysis of mutant and truncated derivatives of histidine kinase NtrB. *J. Bacteriol.* 184:200-206.
- Martínez-Bueno M., Molina-Henares A.J., Pareja, E., Ramos, J.L., Tobes, R. 2004. BacTregulators: a database of transcriptional regulators in bacteria and archaea. *Bioinformatics* 20:2787-2791.
- Martínez-Bueno, M., Tobes, R., Rey, M., and Ramos, J.L. 2002. Detection of multiple extracytoplasmic function ECF-sigma factors in the genome of *Pseudomonas putida* KT2440 and their counterparts in *Pseudomonas aeruginosa* PA01. *Environ. Microbiol.* 4:842-855.
- Martínez-Hackert, E., and A.M. Stock. 1997. Structural relationships in the OmpR family of winged-helix transcription factors. *J. Mol. Biol.* 269:301-312.
- Mavroidi, A., L.S. Tzouveleakis, K.P. Kyriakis, H. Avgerinou, M. Danilidou, and E. Tzelepi. 2001. Multidrug-resistant strains of *Neisseria gonorrhoeae* in Greece. *Antimicrob. Agents Chemother.* 45:2651- 2654.
- Mayer, K., G. Murphy, R. Tarchini, R. Wambutt, G. Volckaert, T. Pohl, A. Duesterhoeft, W. Stiekema, K.D. Entian, N. Terry, K. Lemcke, D. Haase, C.R. Hall, A.M. van Dodeweerd, S.V. Tingey, H.W. Mewes, M.W. Bevan, and I. Bancroft. 2001. Conservation of microstructure between a sequenced region of the genome of rice and multiple segments of the genome of *Arabidopsis thaliana*. *Genome Research* 11:1167-1174.

- McCarter, L. 1999. The multiple identities of *Vibrio parahaemolyticus*. J. Mol. Microbiol. Biotechnol. 1:51-57.
- McCarter, L.L. 1998. OpaR, a homolog of *Vibrio harveyi* LuxR, controls opacity of *Vibrio parahaemolyticus*. J. Bacteriol. 180:3166-3173.
- McDougald, D., S.A. Rice and S.Kjelleberg. 2001. SmcR-dependent regulation of adaptive phenotypes in *Vibrio vulnificus*. J. Bacteriol. 183:758-62.
- McDougald, D., S.A. Rice, and S. Kjelleberg. 2000. The marine pathogen *Vibrio vulnificus* encodes a putative homologue of the *Vibrio harveyi* regulatory gene, LuxR: a genetic and phylogenetic comparison. Gene 248:213-221.
- McKay, D.B., and T.A. Steitz. 1981. Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. Nature 290:744-749.
- McMurry, L.M., M. Oethinger, and S.B. Levy. 1998. Overexpression of *marA*, *soxS*, or *acrAB* produces resistance to triclosan in laboratory and clinical strains of *Escherichia coli*. FEMS Microbiol. Lett. 166:305-309.
- Mecscas, J., Rouviere, P.E., Erickson, J.W., Donohue, T.J., and Gross, C.A. 1993. The activity of σ E, an *Escherichia coli* heat-inducible σ -factor, is modulated by expression of outer membrane proteins. Genes Dev. 7: 2618-2628.
- Méndez, B., C. Tachibana, and S.B. Levy. 1980. Heterogeneity of tetracycline resistance determinants. Plasmid 3:99-108.
- Mermod, N., Lehrbach, P.R., Reineke, W., and Timmis, K.N. 1984. Transcription of the TOL plasmid toluate catabolic pathway operon of *Pseudomonas putida* is determined by a pair of co-ordinately and positively regulated overlapping promoters. EMBO J 3: 2461-2466.
- Merrick, M.J. 1993. In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). Mol. Microbiol. 10:903-9.
- Miller, M.B., K. Skorupski, D.H. Lenz, R.K. Taylor, and B.L. Bassler. 2002. Parallel quorum sensing systems converge to regulate virulence in *Vibrio cholerae*. Cell 110:303-314.
- Milton, D.L., A. Hardman, M. Camara, S.R. Chhabra, B.W. Bycroft, G.S. Stewart and P. William. 1997. Quorum sensing in *Vibrio anguillarum*: characterization of the *vanI/vanR* locus and identification of the autoinducer N-(3-oxodecanoyl)-L-homoserine lactone. J. Bacteriol. 179:3004-3012.
- Mitchell, B.A., I.T. Paulsen, M.H. Brown and R.A. Skurray. 1999. Bioenergetics of the staphylococcal multidrug export protein QacA. Identification of distinct binding sites for monovalent and divalent cations. J. Biol. Chem. 274:3541-3548.
- Miyamoto, C.M., E.A. Meighen, and A.F. Graham. 1990. Transcriptional regulation of *lux* genes transferred into *Vibrio harveyi*. J. Bacteriol. 172:2046-2054.
- Miyamoto, C.M., E.E. Smith, E. Swartzman, J.G. Cao, A.F. Graham, and E.A. Meighen. 1994. Proximal and distal sites bind LuxR independently and activate expression of the *Vibrio harveyi lux* operon. Mol. Microbiol. 14:255-262.
- Miyamoto, C.M., J. Chatterjee, E. Swartzman, R. Szittner, and E.A. Meighen. 1996. The role of *lux* autoinducer in regulating luminescence in *Vibrio harveyi*; control of LuxR expression. Mol. Microbiol. 19:767-775.
- Miyamoto, C.M., M. Boylan, A.F. Graham and E.A. Meighen. 1988. Organization of the *lux* structural genes of *Vibrio harveyi*. Expression under the T7 bacteriophage promoter, mRNA analysis, and nucleotide sequence of the *luxD* gene. J. Biol. Chem. 263:13393- 13399.
- Miyamoto, C.M., W. Sun and E.A. Meighen. 1998. The LuxR regulator protein controls synthesis of polyhydroxybutyrate in *Vibrio harveyi*. Biochim. Biophys. Acta 1384:356-364.
- Miyamoto, C.M., Y.H. Lin, and E.A. Meighen. 2000. Control of bioluminescence in *Vibrio fischeri* by the LuxO signal response regulator. Mol. Microbiol. 36:594-607.
- Mohamed, M.S., W. Ismail, J. Heider, and G. Fuchs. 2002. Aerobic metabolism of phenylacetic acids in *Azoarcus Evansii*. Arch. Microbiol. 178:180-192.

- Morett, E., and L. Segovia. 1993. The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol.* 175:6067-74.
- Mulder, N.J., R. Apweiler, T.K. Attwood, A. Bairoch, D. Barrell, A. Bateman, D. Binns, M. Biswas, P. Bradley, P. Bork, P. Bucher, R.R. Copley, E. Courcelle, U. Das, R. Durbin, L. Falquet, W. Fleischmann, S. Griffiths-Jones, D. Haft, N. Harte, N. Hulo, D. Kahn, A. Kanapin, M. Krestyaninova, R. Lopez, I. Letunic, D. Lonsdale, V. Silventoinen, S.E. Orchard, M. Pagni, D. Peyruc, C.P. Ponting, J.D. Selengut, F. Servant, C.J. Sigrist, R. Vaughan, and E.M. Zdobnov. 2003. The InterPro Database, 2003 brings increased coverage and new features. *Nucl. Acids Res.* 31:315-318.
- Muller, G., B. Hecht, V. Helbl, W. Hinrichs, W. Saenger, and W. Hillen W. 1995. Characterization of non-inducible Tet repressor mutants suggests conformational changes necessary for induction. *Nat. Struct. Biol.* 2:693-703.
- Murata, M., S. Ohno, M. Kumano, K. Yamane, and R. Ohki. 2003. Multidrug resistant phenotype of *Bacillus subtilis* spontaneous mutants isolated in the presence of puromycin and lincomycin. *Can. J. Microbiol.* 49:71-77.
- Murray, D.A., M.A. Schumacher, and R.G. Brennan. 2004. Crystal structures of QacR-diamidine complexes reveal additional multidrug-binding modes and a novel mechanism of drug charge neutralization. *J. Biol. Chem.* 279:14365-14371.
- Nakano, H., E. Takehara, T. Nihira, and Y. Yamada. 1998. Gene replacement analysis of the *Streptomyces virginiae* *barA* gene encoding the butyrolactone autoregulator receptor reveals that BarA acts as a repressor in virginiamycin biosynthesis. *J. Bacteriol.* 180:3317-3322.
- Namwat, W., C.K. Lee, H. Kinoshita, Y. Yamada, and T. Nihira. 2001. Identification of the *varR* gene as a transcriptional regulator of virginiamycin S resistance in *Streptomyces virginiae*. *J. Bacteriol.* 183:2025-2031.
- Natsume, R., Y. Ohnishi, T. Senda, and S. Horinouchi. 2004. Crystal structure of a γ -butyrolactone autoregulator receptor protein in *Streptomyces coelicolor* A3(2). *J. Mol. Biol.* 336:409-419.
- Neal, R.J., and K.F. Chater. 1987. Nucleotide sequence analysis reveals similarities between proteins determining methylenomycin A resistance in *Streptomyces* and tetracycline resistance in eubacteria. *Gene* 58:229-241.
- Neal, R.J., and K.F. Chater. 1991. Bidirectional promoter and terminator regions bracket *mmr*, a resistance gene embedded in the *Streptomyces coelicolor* A3(2) gene cluster encoding methylenomycin production. *Gene* 100:75-83.
- Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A.P., et al. 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol* 4: 799–808.
- Nikaido, H.N. 1996. Multidrug efflux pumps of Gram-negative bacteria. *J. Bacteriol.* 178:5853-5859.
- Nikaido, H.N. 1998. Antibiotic resistance caused by Gram-negative multidrug efflux pumps. *Clin. Infect. Dis.* 27:532-541.
- Nilsson, D., and M. Kilstrup. 1998. Cloning and expression of the *Lactococcus lactis* *purDEK* genes, required for growth in milk. *Appl. Env. Microbiol.* 64:4321-4327.
- Nishino, K., and A. Yamaguchi. 2001. Analysis of a complete library of putative drug transporter genes in *Escherichia coli*. *J. Bacteriol.* 183:5803-5812.
- Nixon, B.T., C.W. Ronson, and F.M. Ausubel. 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. *Proc. Natl. Acad. Sci. USA* 83:7850–7854.
- Noguchi, N., J. Katayama, and M. Sasatsu. 2000. A transposon carrying the gene *mphB* for macrolide 2'-phosphotransferase II. *FEMS Microbiol Lett.* 192:175-178.
- Noguchi, N., K. Takada, J. Katayama, A. Emura Aand M. Sasatsu. 2000. Regulation of transcription of the *mph(A)* gene for macrolide 2'-phosphotransferase I in *Escherichia coli*: characterization of the regulatory gene *mphR(A)*. *J. Bacteriol.* 182:5052-5058.
- North, A.K., K.E. Klose, K.M. Stedman, and S. Kustu. 1993. Prokaryotic enhancer-binding proteins reflect eukaryote-like modularity: the puzzle of nitrogen regulatory protein C. *J. Bacteriol.* 175:4267–4273.

- Novakova, R., J. Bistakova, D. Homerova, B. Rezuchova, and J. Kormanec. 2002. Cloning and characterization of a polyketide synthase gene cluster involved in biosynthesis of a proposed angucycline-like polyketide auricin in *Streptomyces aureofaciens* CCM 3239. *Gene* 297:197-208.
- Ochsner, U.A., and Vasil, M.L. 1996. Gene repression by the ferric uptake regulator in *Pseudomonas aeruginosa*: cycle selection of iron-regulated genes. *Proc. Natl. Acad. Sci. USA* 93: 4409-4414.
- Ochsner, U.A., Johnson, Z., and Vasil, M.L. 2000. Genetics and regulation of two distinct haem-uptake systems, *phu* and *has*, in *Pseudomonas aeruginosa*. *Microbiology* 146: 185-198.
- Oethinger, M., I. Podglajen, W.V. Kern, and S.B. Levy. 1998. Overexpression of the *marA* or *soxS* regulatory gene in clinical topoisomerase mutants of *Escherichia coli*. *Antimicrob. Agents Chemother.* 42:2089-2094.
- Ohki, R., K. Tateno, Y. Okada, H. Okajima, K. Asai, Y. Sadaie, M. Murata, and T. Aiso. 2003. A bacitracin-resistant *Bacillus subtilis* gene encodes a homologue of the membrane-spanning subunit of the *Bacillus licheniformis* ABC transporter. *J. Bacteriol.* 185:51-59.
- Ohnishi, Y., S. Kameyama, H. Onaka, and S. Horinouchi. 1999. The A-factor regulatory cascade leading to streptomycin biosynthesis in *Streptomyces griseus*: identification of a target gene of the A-factor receptor. *Mol. Microbiol.* 34:102-111.
- Ohta, Y., M. Maeda, and T. Kudo. 2001. *Pseudomonas putida* CE2010 can degrade biphenyl by a mosaic pathway encoded by the *tod* operon and *cmtE*, which are identical to those of *P. putida* F1 except for a single base difference in the operator-promoter region of the *cmt* operon. *Microbiology* 147:31-41.
- Okamoto, S., K. Nakamura, T. Nihira, and Y. Yamada. 1995. *Virginiae* butanolide binding protein from *Streptomyces virginiae*. Evidence that VbrA is not the virginiae butanolide binding protein and reidentification of the true binding protein. *J. Biol. Chem.* 270:12319-12326.
- Okusu, H., D. Ma, and H.N. Nikaido. 1996. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. *J. Bacteriol.* 178:306-308.
- Onaka, H., and S. Horinouchi. 1997. DNA-binding activity of the A-factor receptor protein and its recognition DNA sequences. *Mol. Microbiol.* 24:991-1000.
- Onaka, H., M. Sugiyama, and S. Horinouchi. 1997. A mutation at proline-115 in the A-factor receptor protein of *Streptomyces griseus* abolishes DNA-binding ability but not ligand-binding ability. *J. Bacteriol.* 179:2748-2752.
- Onaka, H., N. Ando, T. Nihira, Y. Yamada, T. Beppu, and S. Horinouchi. 1995. Cloning and characterization of the A-factor receptor gene from *Streptomyces griseus*. *J. Bacteriol.* 177:6083-6092.
- Onaka, H., T. Nakagawa, and S. Horinouchi. 1998. Involvement of two A-factor receptor homologues in *Streptomyces coelicolor* A3(2) in the regulation of secondary metabolism and morphogenesis. *Mol. Microbiol.* 28: 743-753.
- Ongena, M., Jacques. P., Delfosse. P., and Thonart, P. 2002. Unusual traits of the pyoverdinin-mediated iron acquisition system in *Pseudomonas putida* strain BTP1. *Biometals* 15: 1-13.
- Orth, P., D. Schnappinger, P.-E. Sum, G.A. Ellestad, W. Hillen, W. Saenger, and W. Hinrichs. 1999b. Crystal structure of the Tet repressor in complex with a novel tetracycline, 9-(N,N-dimethylglycylamido)-6-demethyl-6-deoxy-tetracycline. *J. Mol. Biol.* 285:455-461.
- Orth, P., D. Schnappinger, W. Hillen, W. Saenger, and W. Hinrichs. 2000. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nature Struct. Biol.* 7:215-219.
- Orth, P., F. Cordes, D. Schnappinger, W. Hillen, W. Saenger, and W. Hinrichs. 1998. Conformational changes of the Tet repressor induced by tetracycline trapping. *J. Mol. Biol.* 279:439-447.
- Orth, P., W. Saenger, and W. Hinrichs. 1999a. Tetracycline-chelated Mg²⁺ ion initiates helix unwinding in Tet repressor induction. *Biochemistry* 38:191-198.
- Osteras, M., E. Boncompagni, N. Vincent, M.C. Poggi, and D. Le Rudulier. 1998. Presence of a gene encoding choline sulfatase in *Sinorhizobium meliloti* bet operon: choline-O-sulfate is metabolized into glycine betaine. *Proc. Natl. Acad. Sci. U S A.* 95:11394-11399.
- Otten, S.L., J. Ferguson, and C.R. Hutchinson. 1995. Regulation of daunorubicin production in *Streptomyces peucetius* by the *dnrR2* locus. *J. Bacteriol.* 177:1216-1224.

- Pabo, C.O., and M. Lewis. 1992. The operator-binding domain of lambda repressor: structure and DNA recognition. *Nature* 298:443–447.
- Pabo, C.O., and R.T. Sauer. 1984. Protein-DNA recognition. *Annu. Rev. Biochem.* 53:293–321.
- Pabo, C.O., and R.T. Sauer. 1992. Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Microbiol.* 61:1053-1095.
- Paget, M.S. and Helmann, J.D. 2003. The sigma70 family of sigma factors. *Genome Biol.* 4:203.
- Palmer, C.N., E. Axen, V. Hughes, and C.R. Wolf. 1998. The repressor protein, Bm3R1, mediates an adaptive response to toxic fatty acids in *Bacillus megaterium*. *J. Biol. Chem.* 273:18109-18116.
- Palumbo, J.D., C.I. Kado, and D.A. Phillips. 1998. An isoflavonoid-inducible efflux pump in *Agrobacterium tumefaciens* is involved in competitive colonization of roots. *J. Bacteriol.* 180:3107-3113.
- Pan, W., and B.G. Spratt. 1994. Regulation of the permeability of the gonococcal cell envelope by the mtr system. *Mol. Microbiol.* 11:769–775.
- Pao, G.M., R. Tam, L.S. Lipschitz, and M.H. Saier Jr. 1994. Response regulators: structure, function and evolution. *Res Microbiol.* 145:356-362.
- Paulsen, I.T., M.H. Brown, S.J. Dunstan and R.A. Skurray. 1995. Molecular characterization of the staphylococcal multidrug resistance export protein QacC. *J. Bacteriol.* 177:2827-2833.
- Paulsen, I.T., M.H. Brown, T.G. Littlejohn, B.A. Mitchell, and R.A. Skurray. 1996. Multidrug resistance proteins QacA and QacB from *Staphylococcus aureus*: membrane topology and identification of residues involved in substrate specificity. *Proc. Natl. Acad. Sci. USA* 93:3630-3635.
- Peng, W.T., and E.W. Nester. 2001. Characterization of a putative RND-type efflux system in *Agrobacterium tumefaciens*. *Gene* 270:245-252.
- Perez-Martin, J. de Lorenzo, V. 1997. Clues and consequences of DNA bending in transcription. *Annu. Rev. Microbiol.* 51:593-628.
- Pérez-Rueda, E., and J. Collado-Vides. 2001. Common history at the origin of the position–function correlation in transcriptional regulators in Archaea and Bacteria. *J. Mol. Evol.* 53:172–179.
- Perron, K., O. Caille, C. Rossier, C. van Delden, J.L. Dumas, and T. Köhler. 2004. CzcR-CzcS, a two-component system involved in heavy metal and carbapenem resistance in *Pseudomonas aeruginosa*. *J. Biol. Chem.* 279:8761-8768.
- Pocard, J.A., N. Vincent, E. Boncompagni, L.T. Smith, M.C. Poggi, and D. Le Rudulier. 1997. Molecular characterization of the *bet* genes encoding glycine betaine synthesis in *Sinorhizobium meliloti* 102F34. *Microbiology* 143:1369-1379.
- Pocard, J.A., T. Bernard, L.T. Smith, and D. Le Rudulier. 1989. Characterization of three choline transport activities in *Rhizobium meliloti*: modulation by choline and osmotic stress. *J. Bacteriol.* 171:531-537.
- Poelarends, G.J., L.A. Kulakov, M.J. Larkin, J.E. van Hylckama Vlieg, and D.B. Janssen. 2000. Roles of horizontal gene transfer and gene integration in evolution of 1,3-dichloropropene- and 1,2-dibromoethane-degradative pathways. *J. Bacteriol.* 182:2191-2199.
- Poole, K., D.E. Heinrichs, and S. Neshat. 1993. Cloning and sequence analysis of an EnvCD homologue in *Pseudomonas aeruginosa*: regulation by iron and possible involvement in the secretion of the siderophore pyoverdine. *Mol. Microbiol.* 10:529-544.
- Poole, K., K. Tetro, Q. Zhao, S. Neshat, D.E. Heinrichs, and N. Bianco. 1996. Expression of the multidrug resistance operon *mexA-mexB-oprM* in *Pseudomonas aeruginosa*: *mexR* encodes a regulator of operon expression. *Antimicrob. Agents Chemother.* 40:2021-2028.
- Pragman, A.A., J.M. Yarwood, T.J. Tripp, and P.M. Schlievert. 2004. Characterization of virulence factor regulation by SrrAB, a two-component system in *Staphylococcus aureus*. *J. Bact.* 186:2430-2438.
- Pruitt, K.D. and Maglott, D.R. 2001. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res.*, 29, 137–140.

- Puskas, L.G., M. Inui, Z. Kele, and H. Yukawa. 2000. Cloning of genes participating in aerobic biodegradation of *p*-cumate from *Rhodopseudomonas palustris*. *DNA Seq.* 11:9-20.
- Raina, S., Missiakas, D., and Georgopoulos, C. 1995. The *rpoE* gene encoding the σ^E (σ^{24}) heat shock sigma factor of *Escherichia coli*. *EMBO J* 14: 1043–1055.
- Raivio, T.L., and Silhavy, T.T. 2001. Periplasmic stress and ECF sigma factors. *Annu Rev Microbiol* 55: 591–624.
- Ramos, J.L., E. Duque, M.T. Gallegos, P. Godoy, M.I. Ramos-González, A. Rojas, W. Terán, and A. Segura. 2002. Mechanisms of solvent tolerance in gram-negative bacteria. *Ann. Rev. Microbiol.* 56:743- 768.
- Ramos, J.L. M.T. Gallegos, S. Marqués, M.I. Ramos-González, M. Espinosa-Urgel, and A. Segura. 2001. Response of gram-negative bacteria to certain environmental stresses. *Curr. Opin. Microbiol.* 4:166-171.
- Ramos, J.L., Marqués, S., and Timmis, K.N. 1997. Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. *Annu. Rev. Microbiol.* 51: 341–373.
- Ramos, J.L. Martinez-Bueno, M. Molina-Henares, A.J. Teran, W. Watanabe, K. Zhang, X. Gallegos, M.T. Brennan, R. Tobes, R. 2005. The TetR family of transcriptional repressors. *Microbiol. Mol. Biol. Rev.* 69:326-56.
- Ramos, J.L., F. Rojo, L. Zhou, and K.N. Timmis. 1990. A family of positive regulators related to the *Pseudomonas putida* TOL plasmid XylS and the *Escherichia coli* AraC activators. *Nucleic Acids Res.* 18:2149-2152.
- Ramos-González, M.I., and Molin, S. 1998. Cloning, sequencing, and phenotypic characterization of the *rpoS* gene from *Pseudomonas putida* KT2440. *J. Bacteriol.* 180: 3421–3431.
- Rand, J.D., S.G. Danby, D.L. Greenway, and R.R. England. 2002. Increased expression of the multidrug efflux genes *acrAB* occurs during slow growth of *Escherichia coli*. *FEMS Microbiol. Lett.* 207:91-95.
- Rebets, Y., B. Ostash, A. Luzhetskyy, D. Hoffmeister, A. Brana, C. Mendez, J.A. Salas, A. Bechthold, and V. Fedorenko. 2003. Production of landomycins in *Streptomyces globisporus* 1912 and *S. cyanogenus* S136 is regulated by genes encoding putative transcriptional activators. *FEMS Microbiol. Lett.* 222:149-153.
- Redenbach, M., K. Ikeda, M. Yamasaki, and H. Kinashi. 1998. Cloning and physical mapping of the *EcoRI* fragments of the giant linear plasmid SCP1. *J. Bacteriol.* 180:2796-2799.
- Redfield, R.J. 2002. Is quorum sensing a side effect of diffusion sensing? *Trends Microbiol.* 10:365-370.
- Reizer, A., J. Deutscher, M.H. Saier, and J. Reizer. 1991. Analysis of the gluconate (*gnt*) operon of *Bacillus subtilis*. *Mol. Microbiol.* 5:1081–1089.
- Reizer, J., I.T. Paulsen, A. Reizer, F. Titgemeyer, and M.H. Saier Jr. 1996. Novel phosphotransferase system genes revealed by bacterial genome analysis: the complete complement of pts genes in *Mycoplasma genitalium*. *Microb. Comp. Genomics* 1:151-164.
- Retzlaff, L., and J. Distler. 1995. The regulator of streptomycin gene expression, StrR, of *Streptomyces griseus* is a DNA binding activator protein with multiple recognition sites. *Mol. Microbiol.* 18:151-162.
- Reverchon, S., W. Nasser, and J. Robert-Baudouy. 1991. Characterization of *kdgR*, a gene of *Erwinia chrysanthemi* that regulates pectin degradation. *Mol. Microbiol.* 5:2203-2216.
- Rey, D.A., A. Puhler, and J. Kalinowski. 2003. The putative transcriptional repressor McbR, member of the TetR-family, is involved in the regulation of the metabolic network directing the synthesis of sulfur containing amino acids in *Corynebacterium glutamicum*. *J. Biotechnol.* 103:51-65.
- Rice, P., I. Longden, and A. Bleasby. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16:276-277.
- Richard, C.L. Tandon, A. Sloan, N.R. Kranz, R.G. 2003. RNA polymerase subunit requirements for activation by the enhancer-binding protein *Rhodobacter capsulatus* NtrC. *J. Biol. Chem.* 278:31701-8.
- Rigali, S., A. Derouaux, F. Giannotta, and J. Dusart. 2002. Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies. *J. Biol. Chem.* 277:12507-12515.

- Rodionov, D.A., M.S. Gelfand, A.A. Mironov, and A.B. Rakhmaninova. 2001. Comparative approach to analysis of regulation in complete genomes: multidrug resistance systems in gamma-proteobacteria. *J. Mol. Microbiol. Biotechnol.* 3:319-324.
- Roessner, C.A., K.X. Huang, M.J. Warren, E. Raux, and A.I. Scott. 2002. Isolation and characterization of 14 additional genes specifying the anaerobic biosynthesis of cobalamin vitamin B12. in *Propionibacterium freudenreichii* (*P. shermanii*). *Microbiology* 148: 1845-1853.
- Rojas, A., E. Duque, G. Mosqueda, G. Golden, A. Hurtado, J.L. Ramos, and A. Segura. 2001. Three efflux pumps are required to provide efficient tolerance to toluene in *Pseudomonas putida* DOT-T1E. *J. Bacteriol.* 183:3967-3973.
- Rojas, A., Duque, E., Schmid, A., A. Hurtado, J.L. Ramos and A. Segura. 2004. Biotransformation in double phase systems: physiological responses of *Pseudomonas putida* DOT T1E to a double phase made of aliphatic alcohols and biosynthesis of substituted catechols. *Appl. Env. Microbiol.* 70:3637-43.
- Rojas, A., A. Segura, M.E. Guazzaroni, W. Terán, A. Hurtado, M.T. Gallegos, and J.L. Ramos. 2003. In vivo and in vitro evidence that TtgV is the specific regulator of the TtgGHI multidrug and solvent efflux pump of *Pseudomonas putida*. *J. Bacteriol.* 185:4755- 4763.
- Rojo, F. 1999. Repression of transcription initiation in bacteria. *J. Bacteriol.* 181:2987-2991.
- Røkenes, T.P., T. Lamark, and A.R. Strøm. 1996. DNA-binding properties of the BetI repressor protein of *Escherichia coli*: the inducer choline stimulates BetI-DNA complex formation. *J. Bacteriol.* 178:1663-1670.
- Rouch, D.A., D.S. Cram, D. DiBerardino, T.G. Littlejohn and R.A. Skurray. 1990. Efflux-mediated anti-septic resistance gene *qacA* from *Staphylococcus aureus*: common ancestry with tetracycline- and sugar-transport proteins. *Mol. Microbiol.* 4:2051-2062.
- Rouquette, C., J.B. Harmon and W.M. Shafer. 1999. Induction of the *mtrCDE*-encoded efflux pump system of *Neisseria gonorrhoeae* requires MtrA, an AraC-like protein. *Mol. Microbiol.* 33:651-658.
- Rouquette-Loughlin, C., I. Stojiljkovic, T. Hrobowski, J.T. Balthazar, and W.M. Shafer. 2002. Inducible, but not constitutive, resistance of gonococci to hydrophobic agents due to the MtrC-MtrD- MtrE efflux pump requires TonB-ExbB-ExbD proteins. *Antimicrob. Agents Chemother.* 46:561-565.
- Rousseau, P., E. Guegen, G. Duval-Valentin, and M. Chandler. 2004. The helix-turn-helix motif of bacterial insertion sequence IS911 transposase is required for DNA binding. *Nucl. Acids Res.* 32:1335- 1344.
- Ryu, J.R., L.K. Olson, and D.N. Arnosti. 2001. Cell-type specificity of short-range transcriptional repressors. *Proc. Natl. Acad. Sci. USA* 98:12960- 12965.
- Saenger, W., P. Orth, C. Kisker, W. Hillen, and W. Hinrichs. 2000. The tetracycline repressor-A paradigm for a biological switch. *Angew. Chem. Int. Ed. Engl.* 39:2042-2052.
- Sala, C., F. Forti, E. di Florio, F. Canneva, A. Milano, G. Riccardi, and D. Ghisotti. 2003. *Mycobacterium tuberculosis* FurA autoregulates its own expression. *J. Bacteriol.* 185:5357-5362.
- Salah-Bey, K., and C.J. Thompson. 1995. Unusual regulatory mechanism for a *Streptomyces* multidrug resistance gene, *ptr*, involving three homologous protein-binding sites overlapping the promoter region. *Mol. Microbiol.* 17:1109-1119.
- Sánchez, P., A. Alonso, and J.L. Martínez. 2002. Cloning and characterization of SmeT, a repressor of the *Stenotrophomonas maltophilia* multidrug efflux pump SmeDEF. *Antimicrob. Agents Chemother.* 46:3386- 3393.
- Schauder, S., and B.L. Bassler. 2001. The languages of bacteria. *Genes Dev.* 15:1468-1480.
- Schell, M.A. 1993. Molecular biology of the LysR family of transcriptional regulators. *Ann. Rev. Microbiol.* 47:597-626.
- Schevitz, R.W., Z. Otwinowski, A. Joachimiak, C.L. Lawson, and P.B. Singler. 1985. The three-dimensional structure of *trp* repressor. *Nature* 317:782-786.
- Schindelin, H., M.A. Marahiel, and U. Heinemann. 1993. Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold- shock protein. *Nature* 364:164-168.

- Schleif, R. 1987. The L-arabinose operon. In F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Um-barger ed., *Escherichia coli* and *Salmonella typhimurium* : cellular and molecular biology. American Society for Microbiology, Washington, D.C. pp. 1473–1481.
- Schnider-Keel, U., A. Seematter, M. Maurhofer, C. Blumer, B. Duffy, C. Gigot-Bonnefoy, C. Reimann, R. Notz, G. Defago, D. Haas, and C. Keel. 2000. Autoinduction of 2,4-diacetylphloroglucinol biosynthesis in the biocontrol agent *Pseudomonas fluorescens* CHA0 and repression by the bacterial metabolites salicylate and pyoluteorin. *J. Bacteriol.* 182:1215-1225.
- Schnider-Keel, U., Lejbolle, K.B., Baehler, E., Haas, D., and Keel, C. 2002. The sigma factor AlgU (AlgT) controls exopolysaccharide production and tolerance towards desiccation and osmotic stress in the biocontrol agent *Pseudomonas fluorescens* CHA0. *Appl Environ Microbiol* 67: 5683–5693.
- Schouls, L.M., C.S. Schot, and J.A. Jacobs. 2003. Horizontal transfer of segments of the 16S rRNA genes between species of the *Streptococcus anginosus* group. *J. Bacteriol.* 185:7241-7246.
- Schultz J., R.R. Copley, T. Doerks, C.P. Ponting, and P. Bork. 2000. Smart: A Web-based tool for the study of genetically mobile domains. *Nucl. Acids Res.* 28:231-234.
- Schumacher, M.A., and R.G. Brennan. 2002. Structural mechanisms of multidrug recognition and regulation by bacterial multidrug transcription factors. *Mol. Microbiol.* 45:885-893.
- Schumacher, M.A., M.C. Miller, S. Grkovic, M.H. Brown, R.A. Skurray, and R.G. Brennan. 2001. Structural mechanisms of QacR induction and multidrug recognition. *Science*, 294, 2158–2163.
- Schumacher, M.A., M.C. Miller, S. Grkovic, M.H. Brown, R.A. Skurray, and R.G. Brennan. 2002. Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J.* 21:1210-1218.
- Segura, A., Duque, E., Hurtado, A., and Ramos, J.L. 2001. Mutations in genes involved in the flagellar export apparatus of the solvent-tolerant *Pseudomonas putida* DOT-T1E strain impair motility and lead to hypersensitivity to solvent shocks. *J. Bacteriol.* 183: 4127–4133.
- Seoane, A.S., and S.B. Levy. 1995. Characterization of MarR, the repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*. *J. Bacteriol.* 177:3414-3419.
- Shafer, W.M., J.T. Balthazar, K.E. Hagman, and S.A. Morse. 1995. Missense mutations that alter the DNA-binding domain of the *MtrR* protein occur frequently in rectal isolates of *Neisseria gonorrhoeae* that are resistant to faecal lipids. *Microbiology* 141:907-911.
- Shankar, N., A.S. Baghdayan, and M.S. Gilmore. 2002. Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis*. *Nature* 417:746-750.
- Shao, C.P. and L.I. Hor. 2001. Regulation of metalloprotease gene expression in *Vibrio vulnificus* by a *Vibrio harveyi* LuxR homologue. *J. Bacteriol.* 183:1369-1375.
- Shaw, G.C., and A.J. Fulco. 1992. Barbiturate-mediated regulation of expression of the cytochrome P450BM-3 gene of *Bacillus megaterium* by Bm3R1 protein. *J. Biol. Chem.* 267:5515-5526.
- Shaw, G.C., and A.J. Fulco. 1993. Inhibition by barbiturates of the binding of Bm3R1 repressor to its operator site on the barbiturate-inducible cytochrome P450BM-3 gene of *Bacillus megaterium*. *J. Biol. Chem.* 268:2997-3004.
- Shaw, G.C., C.C. Sung, C.H. Liu, and C.H. Lin. 1998. Evidence against the Bm1P1 protein as a positive transcription factor for barbiturate-mediated induction of cytochrome P450BM-1 in *Bacillus megaterium*. *J. Biol. Chem.* 273:7996-8002.
- Shaw, G.C., C.C. Sung, C.H. Liu, and H.S. Kao. 1997. A 53-base-pair inverted repeat negatively regulates expression of the adjacent and divergently oriented cytochrome P450(BM-1) gene and its regulatory gene, *bm1P1*, in *Bacillus megaterium*. *J. Bacteriol.* 179:280-283.
- Shaw, G.C., H.S. Kao, C.C. Sung, and A.N. Chiang. 1997. Iron and salicylate induction of cytochrome P450BM-1 in *Bacillus megaterium*. *Curr. Microbiol.* 35:28-31.
- Shaw, G.C., Y.H. Hsueh, and H.S. Kao. 2000. The basal-level expression of the cytochrome P450BM-1 gene is negatively affected by the *bm1P1* gene of *Bacillus megaterium*. *Curr. Microbiol.*40:47-50.
- Shaw, G.C., Y.H. Hsueh, C.C. Sung, Y.S. Chen, and C.H. Liu. 2000. Negative regulation of expression of the *Bacillus megaterium* *bm1P1* gene by the *bm1P1* 3' flanking region. *J. Bacteriol.* 182:2654-2659.

- Shen, J., Meldrum, A., and Poole, K. 2002. FpvA receptor involvement in pyoverdine biosynthesis in *Pseudomonas aeruginosa*. *J. Bacteriol.* 184: 3268–3275.
- Showalter, R.E., M.O. Martin and M.R. Silverman. 1990. Cloning and nucleotide sequence of LuxR, a regulatory gene controlling bioluminescence in *Vibrio harveyi*. *J. Bacteriol.* 172:2946-54.
- Sigrist, C.J., L. Cerutti, N. Hulo, A. Gattiker, L. Falquet, M. Pagni, A. Bairoch, and P. Bucher. 2002. PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform.* 3:265-274.
- Silva, A.J., K. Pham, and J.A. Benitez. 2003. Haemagglutinin/ protease expression and mucin gel penetration in El Tor biotype *Vibrio cholerae*. *Microbiology* 149:1883-1891.
- Sizemore, C., A. Wissmann, U. Gülland, and W. Hillen. 1990. Quantitative analysis of Tn10 Tet repressor binding to a complete set of *tet* operator mutants. *Nucl. Acids Res.* 18:2875-2880.
- Skarstad, K., B. Thony, D.S. Hwang, and A. Kornberg. 1993. A novel binding protein of the origin of the *Escherichia coli* chromosome. *J. Biol. Chem.* 268:5365-5370.
- Smith, L.T., J.A. Pocard, T. Bernard, and D. Le Rudulier. 1988. Osmotic control of glycine betaine biosynthesis and degradation in *Rhizobium meliloti*. *J. Bacteriol.* 170:3142-3149.
- Soisson, S.M., B. MacDougall-Shackleton, R. Schleif, and C. Wolberger. 1997. Structural basis for ligand-regulated oligomerization of AraC. *Science* 276:421-425.
- Spiro, S. 1994. The FNR family of transcriptional regulators. *Antonie Leeuwenhoek* 66:23-36.
- Steffan, R.J., Timmis, K.N., and Unterman, R. 1994. Designing microorganisms for the treatment of toxic wastes. *Annu Rev Microbiol* 48: 525–557.
- Stiefel, A., Mahren, S., Ochs, M., Schindler, P.T., Enz, S., and Braun, V. 2001. Control of the ferric citrate transport system of *Escherichia coli*: mutations in region 2.1 of the FecI extracytoplasmic-function sigma factor suppress mutations in the FecR transmembrane regulatory protein. *J. Bacteriol.* 183: 162–170.
- Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., et al. 2000. Complete genome sequence of *Pseudomonas aeruginosa*: a versatile bacterium and an opportunistic pathogen. *Nature* 406: 959–964.
- Stratigopoulos, G., A.R. Gandecha, and E. Cundliffe. 2002. Regulation of tylosin production and morphological differentiation in *Streptomyces fradiae* by TylP, a deduced gamma-butyrolactone receptor. *Mol. Microbiol.* 45:735-744.
- Stratigopoulos, G., and E. Cundliffe. 2002. Expression analysis of the tylosin-biosynthetic gene cluster: pivotal regulatory role of the *tylQ* gene product. *Chem. Biol.* 9:71-78.
- Strom, A.R., P. Falkenberg, and B. Landfald. 1986. Genetics of osmoregulation in *Escherichia coli*: uptake and biosynthesis of organic osmolytes. *FEMS Microbiol. Rev.* 39:79-86.
- Styrvoid, O.B., P. Falkenberg, B. Landfald, M.W. Eshoo, T. Bjørnsen, and A.R. Strøm. 1986. Selection, mapping, and characterization of osmoregulatory mutants of *Escherichia coli* blocked in the choline-glycine betaine pathway. *J. Bacteriol.* 165:856–863.
- Sugiyama, M., H. Onaka, T. Nakagawa, and S. Horinouchi. 1998. Site-directed mutagenesis of the A-factor receptor protein: Val-41 important for DNA- binding and Trp-119 important for ligand-binding. *Gene* 222: 133-144.
- Sulavik, M.C., L.F. Gambino, and P.F. Miller. 1995. The MarR repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*: prototypic member of a family of bacterial regulatory proteins involved in sensing phenolic compounds. *Mol. Med.* 1:436-46.
- Summers, A.O. 1992. Untwist and shout: a heavy metal-responsive transcriptional regulator. *J. Bacteriol.* 174:3097-3101.
- Sunnarborg, A., D. Klumpp, T. Chung, and D.C. LaPorte. 1990. Regulation of the glyoxylate bypass operon: cloning and characterization of *icIR*. *J. Bacteriol.* 172:2642-1649.
- Surette, M.G., and B.L. Bassler. 1998. Quorum sensing in *Escherichia coli* and *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. USA.* 95:7046- 7050.

- Swartzman, E., and E.A. Meighen EA. 1993. Purification and characterization of a poly(dA-dT) *lux*-specific DNA-binding protein from *Vibrio harveyi* and identification as LuxR. *J. Biol. Chem.* 268:16706-16716.
- Swartzman, E., C. Miyamoto, A. Graham, and E. Meighen. 1990. Delineation of the transcriptional boundaries of the *lux* operon of *Vibrio harveyi* demonstrates the presence of two new *lux* genes. *J. Biol. Chem.* 265:3513-3517.
- Swartzman, E., M. Silverman, and E.A. Meighen. 1992. The LuxR gene product of *Vibrio harveyi* is a transcriptional activator of the *lux* promoter. *J. Bacteriol.* 174:7490-7493.
- Szczepanowski, R., I. Krahn, B. Linke, A. Goesman, A. Pühler, and A. Schlüter. 2004. Antibiotic multiresistance plasmid pRSB101 isolated from a wastewater treatment plant is related to plasmids residing in phytopathogenic bacteria and carries eight different resistance determinants including a multidrug transport system. *Microbiology* 150:3613- 3630.
- Takahashi, M., L. Altschmied, and W. Hillen. 1986. Kinetic and equilibrium characterization of the Tet-repressor-tetracycline complex by fluorescence measurements; evidence for divalent ion requirements and energy transfer. *J. Mol. Biol.* 187:341-348.
- Takano, E., R. Chakraborty, T. Nihira, Y. Yamada and M.J. Bibb. 2001. A complex role for the gamma-butyrolactone SCB1 in regulating antibiotic production in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.* 41:1015-1028.
- Takano, E., T. Nihira, Y. Hara, J.J. Jones, C.J. Gershater, Y. Yamada and M. Bibb. 2000. Purification and structural determination of SCB1, a gamma-butyrolactone that elicits antibiotic production in *Streptomyces coelicolor* A3(2). *J. Biol. Chem.* 275:11010-11016.
- Tatusov, R.L., D.A. Natale, I.V. Garkavtsev, T.A. Tatusova, U.T. Shankavaram, B.S. Rao, B. Kiryutin, M.Y. Galperin, N.D. Fedorova, and E.V. Koonin. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucl. Acids Res.* 29:22-28.
- Tauch, A., A. Pühler, J. Kalinowski, and G. Thierbach. 2000. TetZ, a new tetracycline resistance determinant discovered in gram-positive bacteria, shows high homology to gram-negative regulated efflux systems. *Plasmid* 44:285-291.
- Tauch, A., S. Gotker, A. Puhler, J. Kalinowski, and G. Thierbach. 2002. The 27.8-kb R-plasmid pTET3 from *Corynebacterium glutamicum* encodes the aminoglycoside adenylyltransferase gene cassette aadA9 and the regulated tetracycline efflux system Tet 33 flanked by active copies of the widespread insertion sequence IS*6100*. *Plasmid* 48:117-129.
- Tennent, J.M., B.R. Lyon, M.T. Gillespie, J.W. May and R.A. Skurray. 1985. Cloning and expression of *Staphylococcus aureus* plasmid-mediated quaternary ammonium resistance in *Escherichia coli*. *Antimicrob. Agents Chemother.* 27:79-83.
- Terán, W., A. Felipe, A. Segura, A. Rojas, J.L. Ramos, and M.T. Gallegos. 2003. Antibiotic-dependent induction of *Pseudomonas putida* DOT-T1E TtgABC efflux pump is mediated by the drug binding repressor TtgR. *Antimicrob. Agents Chemother.* 47:3067-3072.
- Thieffry, D., A.M. Huerta, E. Perez-Rueda, and J. Collado-Vides. 1998. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays* 20:433-440.
- Tiebel, B., K. Garke, and W. Hillen W. 2000. Observing conformational and activity changes of *tet* repressor in vivo. *Nature Struct. Biol.* 7:479-481.
- Tobes, R., and J.L. Ramos. 2002. AraC-XylS database: a family of positive transcriptional regulators in bacteria. *Nucl. Acids Res.* 30:318-321.
- Totten, P.A., Cano-Lara, J., and Lory, S. 1990. The *rpoN* gene product of *Pseudomonas aeruginosa* is required for expression of diverse genes, including the flagellin gene. *J. Bacteriol.* 172: 389-396.
- Tovar, K., A. Ernst, and W. Hillen. 1988. Identification and nucleotide sequence of the class E *tet* regulatory elements and operator and inducer binding of the encoded purified Tet repressor. *Mol. Gen. Genet.* 215:76-80.
- Trefzer, A., S. Pelzer, J. Schimana, S. Stockert, C. Bihlmaier, H.P. Fiedler, K. Welzel, A. Vente, and A. Bechthold. 2002. Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic. *Antimicrob. Agents Chemother.* 46:1174-1182.

- Tropel, D., and J.R. van der Meer. 2004. Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol. Mol. Biol. Rev.* 68:474-500.
- Ueda, K., H. Takano, M. Nishimoto, H. Inaba, and T. Beppu. 2005. Dual transcriptional control of *amfTSBA*, which regulates the onset of cellular differentiation in *Streptomyces griseus*. *J. Bacteriol.* 187: 135-142.
- Ueda, K., K. Matsuda, H. Takano, and T. Beppu. 1999. A putative regulatory element for carbon-source-dependent differentiation in *Streptomyces griseus*. *Microbiology* 145:2265-2271.
- Ulmasou, B., J. Capone, and W. Folk. 1997. Regulated expression of plant RNA genes by the prokaryotic *tet* and *lac* repressors. *Plant Mol. Biol.* 25:417-424.
- Umeyama, T., P.C. Lee, K. Ueda, and S. Horinouchi. 1999. An AfsK/AfsR system involved in the response of aerial mycelium formation to glucose in *Streptomyces griseus*. *Microbiology* 145:2281-2292.
- Unger, B., G. Klock, and W. Hillen. 1984. Nucleotide sequence of the repressor gene of the RA1 tetracycline resistance determinant: structural and functional comparison with three related Tet repressor genes. *Nucl. Acids Res.* 12:7693-7703.
- Urlinger, S., V. Helbl, J. Guthman, E. Pook, S. Grimm, and W. Hillen. 2000. The P65 domain from NF-KB is an efficient human activator in the tetracycline- regulatable gene expression system. *Gene* 247:103-110.
- van Der Geize, R., G.I. Hessels, R. van Gerwen, J.W. Vrijbloed, P. van Der Meijden, and L. Dijkhuizen. 2000. Targeted disruption of the *kstD* gene encoding a 3-ketosteroid delta(1)-dehydrogenase isoenzyme of *Rhodococcus erythropolis* strain SQ1. *Appl. Environ. Microbiol.* 66:2029-2036.
- van Rooijen, R.J., and W.M. de Vos. 1990. Molecular cloning, transcriptional analysis and nucleotide sequence of *lacR*, a gene encoding the repressor of the lactose phosphotransferase system of *Lactococcus lactis*. *J. Biol. Chem.* 265:18499-18503.
- Veal, W.L., A. Yellen, J.T. Balthazar, W. Pan, B.G. Spratt, and W.M. Shafer. 1998. Loss-of-function mutations in the *mtr* efflux system of *Neisseria gonorrhoeae*. *Microbiology* 144:621-627.
- Veal, W.L., and W.M. Shafer. 2003. Identification of a cell envelope protein MtrF. involved in hydrophobic antimicrobial resistance in *Neisseria gonorrhoeae*. *J. Antimicrob. Chemother.* 51:27-32.
- Veal, W.L., R.A. Nicholas, and W.M. Shafer. 2002. Overexpression of the MtrC-MtrD-MtrE efflux pump due to an *mtrR* mutation is required for chromosomally mediated penicillin resistance in *Neisseria gonorrhoeae*. *J. Bact.* 184:5619-5621.
- Ventre, I., A. Filloux, and A. Lazdunski. 2004. Two-component signal transduction systems : a key to the adaptative potential of *Pseudomonas aeruginosa*. In: *Pseudomonas Vol II*, J.L. Ramos Ed.. Kluwer, London, pp. 257-288.
- Venturi, V., Ottevanger, C., Bracke, M., and Weisbeek, P. 1995. Iron regulation of siderophore biosynthesis and transport in *Pseudomonas putida* WCS358: involvement of a transcriptional activator and of the Fur protein. *Mol Microbiol* 15: 1081-1093.
- Vicente, M. Chater, K.F. de Lorenzo, V. 1999. Bacterial transcription factors involved in global regulation. *Mol. Microbiol.* 331:8-17.
- Vimr, E.R., K.A. Kalivoda, E.L. Deszo, and S.M. Steenbergen. 2004. Diversity of microbial sialic acid metabolism. *Microbiol. Mol. Biol. Rev.* 68:132-153.
- Visca, P., Leoni, L., Wilson, M.J., and Lamont, I.L. 2002. Iron transport and regulation, cell signalling and genomics: lessons from *Escherichia coli* and *Pseudomonas*. *Mol. Microbiol.* 45: 1177-1190.
- Vujaklija, D., K. Ueda, S.K. Hong, T. Beppu, and S. Horinouchi. 1991. Identification of an A-factor-dependent promoter in the streptomycin biosynthetic gene cluster of *Streptomyces griseus*. *Mol. Gen. Genet.* 229:119-128.
- Vujaklija, D., S. Horinouchi, and T. Beppu. 1993. Detection of an A-factor-responsive protein that binds to the upstream activation sequence of *strR*, a regulatory gene for streptomycin biosynthesis in *Streptomyces griseus*. *J. Bacteriol.* 175:2652-2661.
- Waki, M., T. Nihira and Y. Yamada. 1997. Cloning and characterization of the gene (*farA*) encoding the receptor for an extracellular regulatory factor (IM-2) from *Streptomyces* sp. strain FRI-5. *J. Bacteriol.* 179:5131-5137.

- Walczak, R.J., A.J. Woo, W.R. Strohl, and N.D. Priestley. 2000. Nonactin biosynthesis: the potential nonactin biosynthesis gene cluster contains type II polyketide synthase-like genes. *FEMS Microbiol. Lett.* 183:171-175.
- Wang, L., and B. Weiss. 1992. *dcd* dCTP deaminase gene of *Escherichia coli*: mapping, cloning, sequencing, and identification as a locus of suppressors of lethal *dut* (dUTPase) mutations. *J. Bacteriol.* 174:5647-5653.
- Wang, L., R.L. White, and L.C. Vining. 2002. Biosynthesis of the dideoxysugar component of jadomycin B: genes in the jad cluster of *Streptomyces venezuelae* ISP5230 for L-digitoxose assembly and transfer to the angucycline aglycone. *Microbiology* 148:1091-1103.
- Wasylewski, Z., P. Kaszycki, and M. Drwiega. 1996. A fluorescence study of Tn10-encoded tet repressor. *J. Protein Chem.* 15:45-58.
- Waters, S.H., P. Rogowsky, J. Grinstead, J. Altenbuchner, and R. Schmitt. 1983. The tetracycline resistance determinants of RP1 and Tn1721: nucleotide sequence analysis. *Nucleic Acids Res.* 11:6089-6105.
- Weber, W. Schoenmakers, R. Spielmann, M. El-Baba, M.D. Folcher, M. Keller, B. Weber, C.C. Link, N. van de Wetering, P. Heinzen, C. Jolivet, B. Sequin, U. Aubel, D. Thompson, C.J. Fussenegger, M. 2003. *Streptomyces*-derived quorum-sensing systems engineered for adjustable transgene expression in mammalian cells and mice. *Nucleic Acids Res.* 31:e71.
- Weickert, M.J., and S. Adhya. 1992. A family of bacterial regulators homologous to Gal and Lac repressors. *J. Biol. Chem.* 267:15869-15874.
- Wen, L.P., and A.J. Fulco. 1987. Cloning of the gene encoding a catalytically self-sufficient cytochrome P-450 fatty acid monooxygenase induced by barbiturates in *Bacillus megaterium* and its functional expression and regulation in heterologous (*Escherichia coli*) and homologous (*Bacillus megaterium*) hosts. *J. Biol. Chem.* 262:6676-6682.
- Wery, J., B. Hidayat, J. Kieboom, and J.A. de Bont. 2001. An insertion sequence prepares *Pseudomonas putida* S12 for severe solvent stress. *J. Biol. Chem.* 276:5700-5706.
- Westbrock-Wadman, S., D.R. Sherman, M.J. Hickey, S.N. Coulter, Y.Q. Zhu, P. Warrener, L.Y. Nguyen, R.M. Shawar, K.R. Folger, and C.K. Stover. 1999. Characterization of a *Pseudomonas aeruginosa* efflux pump contributing to aminoglycoside impermeability. *Antimicrob. Agents Chemother.* 43:2975-2983.
- Westrich, L., S. Domann, B. Faust, D. Bedford, D.A. Hopwood and A. Bechthold. 1999. Cloning and characterization of a gene cluster from *Streptomyces cyanogenus* S136 probably involved in landomycin biosynthesis. *FEMS Microbiol. Lett.* 170:381-387.
- Wheeler, D.L., Chappey, C., Lash, A.E., Leipe, D.D., Madden, T.L. and Schuler, G.D., Tatusova, T.A. and Rapp, B.A. 2000. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, 28, 10-14.
- White, D.G., J.D. Goldman, B. Demple, and S.B. Levy. 1997. Role of the *acrAB* locus in organic solvent tolerance mediated by expression of *marA*, *soxS*, or *rob* in *Escherichia coli*. *J. Bacteriol.* 179:6122-6126.
- Wigneshweraraj, S.R. Burrows, P.C. Bordes, P. Schumacher, J. Rappas, M. Finn, R.D. Cannon, W.V. Zhang, X. Buck, M. 2005. The Second Paradigm for Activation of Transcription. *Prog. Nucleic Acid Res. Mol. Biol.* 79:339-69.
- Wilson, M.S., J.B. Herrick, C.O. Jeon, D.E. Hinman, and E.L. Madsen. 2003. Horizontal transfer of *phnAc* dioxygenase genes within one of two phenotypically and genotypically distinctive naphthalene-degrading guilds from adjacent soil environments. *Appl. Environ. Microbiol.* 69:2172-2181.
- Winans, S.C., and B.L. Bassler. 2002. Mob psychology. *J. Bacteriol.* 184:873-883.
- Wintjens, R., and M. Rooman. 1996. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. *J. Mol. Biol.* 262:294-313.
- Winzer, K., K. Lorenz, B. Zickner, and P. Durre. 2000. Differential regulation of two thiolase genes from *Clostridium acetobutylicum* DSM 792. *J. Mol. Microbiol. Biotechnol.* 2:531-541.
- Wissmann, A., R. Baumeister, G. Muller, B. Hecht, V. Helbl, K. Pfeleiderer, and W. Hillen. 1991. Amino acids determining operator binding specificity in the helix-turn-helix motif of Tn10 Tet repressor. *EMBO J.* 10:4145-4152.

- Wösten, M.M.S.M., J.A. Wagenaar, and J.P.M. van Putten. 2004. The FlgS/FlgR two-component signal transduction system regulates the *fla* regulon in *Campylobacter jejuni*. *J. Biol. Chem.* 279:16214-16222.
- Wu, J., and B.P. Rosen. 1991. The ArsR protein is a trans-acting regulatory protein. *Mol. Microbiol.* 5:1331-1336.
- Xi, C., E. Schoeters, J. Vanderleyden, and J. Michiels. 2000. Symbiosis-specific expression of *Rhizobium etli casA* encoding a secreted calmodulin-related protein. *Proc. Natl. Acad. Sci. USA* 97:11114- 11119.
- Yamaguchi, A., N. Ono, T. Akasaka, T. Noumi, and T. Sawai. 1999. Metal-tetracycline/H⁺ antiporter of *Escherichia coli* encoded by a transposon, Tn10. *J. Biol. Chem.* 265:15525-15530.
- Yamaguchi, A., T. Udagawa, and T. Sawai. 1999. Transport of divalent cations with tetracycline as mediated by the transposon Tn10-encoded tetracycline resistance protein. *J. Biol. Chem.* 9:4809-4813.
- Yamamoto, K. and Ishihama, A. 2003. Two different modes of transcription repression of *Escherichia coli* acetate operon by IclR. *Mol. Microbiol.* 47:183-194.
- Yamamoto, K., K. Yata, N. Fujita, and A. Ishihama. 2001. Novel mode of transcription regulation by SdiA, an *Escherichia coli* homologue of the quorum-sensing regulator. *Mol. Microbiol.* 41:1187- 1198.
- Yamazaki, H., Y. Ohnishi, and S. Horinouchi. 2000. An A-factor-dependent extracytoplasmic function AdsA sigma factor that is essential for morphological development in *Streptomyces griseus*. *J. Bacteriol.* 182:4596-4605.
- Yamazaki, H., A. Tomono, Y. Ohnishi, and S. Horinouchi. 2004. DNA-binding specificity of AdpA, a transcriptional activator in the A-factor regulatory cascade in *Streptomyces griseus*. *Mol. Microbiol.* 53:555-572.
- Yamazaki, H., Y. Ohnishi, and S. Horinouchi. 2003. Transcriptional switch on *ssgA* by A-factor, which is essential for spore septum formation in *Streptomyces griseus*. *J. Bacteriol.* 185:1273-1283.
- Yamazaki, H., Y. Takano, Y. Ohnishi, and S. Horinouchi. 2003. *amfR*, an essential gene for aerial mycelium formation, is a member of the AdpA regulon in the A-factor regulatory cascade in *Streptomyces griseus*. *Mol. Microbiol.* 50:1173-1187.
- Yamazaki, M., Y. Ikuto, A. Ohira, K. Chater, and H. Kinashi. 2003. Limited regions of homology between linear and circular plasmids encoding methylenomycin biosynthesis in two independently isolated streptomycetes. *Microbiology* 149:1351-1356.
- Yang, K., L. Han, and L.C. Vining. 1995. Regulation of jadomycin B production in *Streptomyces venezuelae* ISP5230: involvement of a repressor gene, jadR2. *J. Bacteriol.* 177:6111-61117.
- Yang, K., L. Han, J. He, L. Wang, and L.C. Vining. 2001. A repressor-response regulator gene pair controlling jadomycin B production in *Streptomyces venezuelae* ISP5230. *Gene* 279:165-173.
- Yoshimura, M., K. Asai, Y. Sadaie, and H. Yoshikawa. 2004. Interaction of *Bacillus subtilis* extracytoplasmic function ECF. sigma factors with the N-terminal regions of their potential anti-sigma factors. *Microbiology* 150:591-599.
- Yu, E.W., J.R. Aires, and H. Nikaido. 2003. AcrB multidrug efflux pump of *Escherichia coli*: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity. *J. Bacteriol.* 185:5657-5664.
- Yu, E.W., G. McDermott, H.I. Zgurskaya, H. Nikaido, D.E. Koshland Jr. 2003. Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump. *Science* 300:976-80.
- Zarantonelli, L., G. Borthagaray, E.H. Lee E, W. Veal, and W.M. Shafer. 2001. Decreased susceptibility to azithromycin and erythromycin mediated by a novel *mtrR* promoter mutation in *Neisseria gonorrhoeae*. *J. Antimicrob. Chemother.* 47:651-654.
- Zarantonelli, L., G. Borthagaray, E.H. Lee, and W.M. Shafer. 1999. Decreased azithromycin susceptibility of *Neisseria gonorrhoeae* due to *mtrR* mutations. *Antimicrob. Agents Chemother.* 43:2468- 2472.
- Zdobnov, E.M., and R. Apweiler. 2001. InterProScan an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17:847-848.
- Zeng, G., S. Ye, and T.J. Larson. 1996. Repressor for the sn-glycerol 3-phosphate regulon of *Escherichia coli* K-12: primary structure and identification of the DNA-binding domain. *J. Bacteriol.* 178:7080-7089.

- Zhang, G., X. Hang, P. Green, K.P. Ho, and G.Q. Chen. 2001. PCR cloning of type polyhydroxyalkanoate biosynthesis genes from two *Pseudomonas* strains. *FEMS Microbiol. Lett.* 198:165-170.
- Zhang, L., X.Z. Li, and K. Poole. 2001. SmeDEF multidrug efflux pump contributes to intrinsic multidrug resistance in *Stenotrophomonas maltophilia*. *Antimicrob. Agents Chemother.* 45:3497-3503.
- Zhao, J., and T. Aoki. 1992. Nucleotide sequence analysis of the class G tetracycline resistance determinant from *Vibrio anguillarum*. *Microbiol. Immunol.* 36:1051-1060.
- Zhou, L., L. Xiang-He, B.R. Bochner, and B.L. Wahner. 2003. Phenotype microarray analysis of *Escherichia coli* K-12 mutants with deletions of all two-component systems. *J. Bacteriol.* 185:4956-4972.
- Zhu, J., M.B. Miller, R.E. Vance, M. Dziejman, B.L. Bassler and J.J. Mekalanos. 2002. Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. USA* 99:3129-3134.
- Ziebuhr, W., V. Krimmer, S. Rachid, I. Lossner, F. Gotz and J. Hacker. 1999. A novel mechanism of phase variation of virulence in *Staphylococcus epidermidis*: evidence for control of the polysaccharide intercellular adhesin synthesis by alternating insertion and excision of the insertion sequence element IS256. *Mol. Microbiol.* 32:345-356.

