

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 576 114**

51 Int. Cl.:

G01N 33/68 (2006.01)

C07K 14/35 (2006.01)

C12N 15/31 (2006.01)

C12N 15/63 (2006.01)

C12R 1/32 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **22.09.2009 E 09815404 (0)**

97 Fecha y número de publicación de la concesión europea: **16.03.2016 EP 2344891**

54 Título: **Nanoporos MSP y procedimientos relacionados**

30 Prioridad:

22.09.2008 US 98938 P

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

05.07.2016

73 Titular/es:

UNIVERSITY OF WASHINGTON (50.0%)

4311 11th Avenue NE, Suite 500

Seattle, WA 98105-4608, US y

THE UAB RESEARCH FOUNDATION (50.0%)

72 Inventor/es:

GUNDLACH, JENS, H.;

NIEDERWEIS, MICHAEL;

BUTLER, THOMAS, Z.;

PAVLENOK, MIKHAIL;

TROLL, MARK, A. y

SUKUMARAN, SUJA

74 Agente/Representante:

PONTI SALES, Adelaida

Observaciones :

Véase nota informativa (Remarks) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 576 114 T3

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

DESCRIPCIÓN

Nanoporos MSP y procedimientos relacionados

5 REFERENCIA CRUZADA A LA SOLICITUD DE RELACIONADA

[0001] La presente solicitud reivindica el beneficio de la Solicitud Provisional de Estados Unidos N° de Serie 61/098.938, presentada el 22 de septiembre, 2008.

10 DECLARACIÓN DE DERECHOS DE LICENCIA DEL GOBIERNO

[0002] La presente invención se realizó con el apoyo del Gobierno bajo la subvención número 5R21HG004145 otorgada por el Instituto Nacional de Salud. El Gobierno tiene ciertos derechos en la invención.

15 ANTECEDENTES

[0003] Las tecnologías de secuenciación de ADN establecidas requieren grandes cantidades de ADN y varias etapas largas para construir sólo varias decenas de bases de la secuencia completa. Esta información debe entonces montarse en estilo "shotgun", un esfuerzo que depende no linealmente del tamaño del genoma y de la longitud de los fragmentos de los que se construye el genoma completo. Estas etapas son caras y consumen mucho tiempo, especialmente cuando se secuencian los genomas de mamíferos. Butler, T ("Nanopore analysis of nucleic acids" Tesis (Ph.D.) Universidad de Washington, 2007, Dissertation Abstracts International, Vol. 68, núm. 5B 2007, página 2889) describe una porina mutante MspA D90s/D91 S/D93N.

25 DESCRIPCIÓN RESUMIDA

[0004] La presente invención se define de acuerdo con las reivindicaciones. Por lo tanto, la invención proporciona una porina de *Mycobacterium smegmatis* (Msp) que tiene un vestíbulo y una zona de constricción que definen un túnel, en el que la Msp comprende una MspA mutante que comprende una mutación en las posiciones 90, 91 y 93, en la que la mutación en la posición 90 es D90N, la mutación en la posición 91 es D91N, y la mutación en la posición 93 es D93N. En algunas realizaciones, la Msp comprende al menos un aminoácido con carga positiva adicional en comparación con el vestíbulo de una Msp de tipo salvaje, tal como una mutación en la posición 118, en la que la mutación en la posición 118 es D118R, y opcionalmente una mutación en la posición 139, en la que la mutación en la posición 139 es E139K o E139R. En algunas realizaciones, la Msp comprende mutaciones en las posiciones 118, 139 y 134, en la que la mutación en la posición 118 es D118R, la mutación en la posición 134 es D134R, y la mutación en la posición 139 es E139K.

[0005] La presente invención también proporciona el uso de una Msp de la invención, para detectar la presencia de un analito.

[0006] La presente invención también proporciona un procedimiento para detectar la presencia de un analito, que comprende: aplicar un campo eléctrico suficiente para trasladar un analito de un primer medio líquido conductor a un segundo medio líquido conductor en comunicación de líquidos a través de una Msp de la invención; y medir una corriente de iones, en la que la aparición de un bloqueo en el patrón de corriente de iones indica la presencia del analito en el primer medio líquido conductor. En algunas realizaciones, dicho primer medio líquido conductor se mantiene en tierra, y se aplica una tensión positiva al segundo medio líquido conductor. En algunas realizaciones, el bloqueo o bloqueos en el patrón de corriente de iones se caracterizan por: (a) reducción de la corriente de iones hasta entre el 80% y el 50% del nivel de no bloqueado; o (b) reducción de la corriente de iones hasta menos del 50% del nivel no bloqueado. En algunas realizaciones, el procedimiento se lleva a cabo a 180mV o más.

[0007] La presente invención también proporciona un aparato para usar en el procedimiento de la invención, que comprende una Msp de la invención, en el que el túnel de la Msp se encuentra entre un primer medio líquido conductor y un segundo medio líquido conductor; en el que al menos un medio líquido conductor comprende un analito, y en el que el sistema es operativo para detectar el analito.

[0008] En algunas realizaciones de la invención, el analito es un polímero, por ejemplo ácido nucleico. En realizaciones de la invención en el que el analito es un polímero que comprende más de una unidad, el procedimiento de la invención puede comprender además: identificar una o más unidades del polímero en un procedimiento que comprende medir la corriente de iones o proporcionar un patrón de corriente que comprende un bloqueo para cada unidad de polímero; y comparar uno o más bloqueos en cada patrón de corriente con (i) uno o más bloqueos en el patrón de corriente o (ii) uno o más bloqueos en un patrón de corriente obtenido usando un polímero que tiene unidades conocidas. En algunas realizaciones, el uso, el procedimiento o aparato de la invención es para la secuenciación de ácidos nucleicos. El ácido nucleico puede ser ADN, por ejemplo ADN monocatenario, o ARN.

[0009] La presente invención también proporciona un ácido nucleico que codifica una Msp de la invención.

5 [0010] La presente invención también proporciona un vector que comprende un ácido nucleico que codifica una Msp de la invención. El vector comprende opcionalmente un promotor unido operativamente a dicho ácido nucleico que codifica una Msp de la invención.

10 [0011] La presente invención también proporciona una bacteria *Mycobacterium smegmatis* mutante capaz de la expresión de Msp, comprendiendo la bacteria: (a) una delección de una MspA de tipo salvaje; (b) una delección de una MspC de tipo salvaje; (c) una delección de una MspD de tipo salvaje; (e) un vector como se describe anteriormente. En algunas realizaciones, la bacteria *Mycobacterium smegmatis* mutante es la cepa ML16 de *Mycobacterium smegmatis*. La invención proporciona también una porina de *Mycobacterium smegmatis* obtenible a partir de la bacteria *Mycobacterium smegmatis* mutante descrita anteriormente.

15 [0012] También se proporciona un procedimiento que comprende aplicar un campo eléctrico a una porina de porina de *Mycobacterium smegmatis* (Msp) que tiene un vestíbulo y una zona de constricción que definen un túnel, en el que la porina de Msp se encuentra entre un primer medio líquido conductor y un segundo medio líquido conductor.

20 [0013] También se proporciona un procedimiento de modificación de la conductancia a través del túnel de una porina de Msp que comprende eliminar, añadir, o sustituir al menos un aminoácido en el vestíbulo o en la zona de constricción de una porina de Msp de tipo salvaje.

25 [0014] También se proporciona un sistema que comprende una porina de Msp que tiene un vestíbulo y una zona de constricción que definen un túnel, en el que el túnel se encuentra entre un primer medio líquido y un segundo medio líquido, en el que al menos un medio líquido comprende un analito, y en el que el sistema es operativo para detectar una propiedad del analito.

30 [0015] También se proporciona un sistema que comprende una porina de Msp que tiene un vestíbulo y una zona de constricción que definen un túnel, en el que el túnel se encuentra en una bicapa lipídica entre un primer medio líquido y un segundo medio líquido, y en el que el único punto de la comunicación de líquidos entre el primer y el segundo medio líquido se produce en el túnel.

35 [0016] También se proporcionan porinas Msp mutantes. Por ejemplo, se proporciona una porina de porina de *Mycobacterium smegmatis* A (MspA) que comprende un vestíbulo y una zona de constricción que definen un túnel, y al menos un primer monómero de MspA mutante que comprende una mutación en la posición 93 y una mutación en la posición 90, posición 91, o ambas posiciones 90 y 91. También se proporciona una porina de MspA mutante que comprende un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a sobre 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en la que el vestíbulo y la zona de constricción juntos definen un túnel, y que comprende además al menos un primer monómero de MspA mutante parálogo u homólogo. También se proporciona un mutante de MspA parálogo u homólogo que comprende un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro desde aproximadamente 0,3 a aproximadamente 3 nm, en la que el vestíbulo y la zona de constricción juntos definen un túnel.

45 [0017] Se describen procedimientos de preparación de porinas Msp mutantes. Por ejemplo, en el presente documento se proporciona un procedimiento de fabricación de una porina de MspA mutante, que comprende modificar un monómero de MspA de tipo salvaje en la posición 93 y en la posición 90, posición 91, o en ambas posiciones 90 y 91. También se proporciona un procedimiento de fabricación de una porina de MspA mutante que tiene un vestíbulo y una zona de constricción que definen un túnel, que comprende la eliminar, añadir o sustituir cualquier aminoácido en el vestíbulo o en la zona de constricción de monómero de MspA de tipo salvaje parálogo u homólogo, de manera que la porina de MspA mutante resultante es capaz de trasladar un analito a través del túnel después de la aplicación de un campo eléctrico.

50 [0018] También se proporciona un procedimiento que comprende trasladar un analito a través de un túnel de una porina de porina de *Mycobacterium smegmatis* (Msp) sin el empleo de un campo eléctrico.

55 [0019] En el presente documento se proporcionan secuencias de ácido nucleico. Opcionalmente, una secuencia de ácido nucleico puede comprender una primera y segunda secuencia de nucleótidos, en el que la primera secuencia de nucleótidos codifica una primera secuencia de monómero de Msp y la segunda secuencia de nucleótidos codifica una segunda secuencia de monómero de Msp. La secuencia de ácido nucleico puede comprender además una tercera secuencia de nucleótidos que codifica una secuencia de aminoácidos enlazadora. Opcionalmente, la secuencia de ácido nucleico comprende además un tercer o más secuencia de nucleótidos que codifica una tercera o más secuencia de monómero de Msp. Por ejemplo, la secuencia de ácido nucleico puede comprender además una tercera, cuarta, quinta, sexta, séptima y octava secuencia de nucleótidos. La primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de nucleótidos codifican una primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de monómero de Msp, y la secuencia de ácido nucleico comprende además una

noveno secuencia de nucleótidos que codifica una secuencia de aminoácidos de enlace. También se proporcionan porinas Msp que comprenden dos o más Msp de cadena única.

5 [0020] También se proporcionan polipéptidos codificados por los ácidos nucleicos descritos en el presente documento. También se proporcionan vectores que comprenden los polipéptidos descritos en el presente documento. También se proporcionan células cultivadas transfectadas con cualquier vector descrito en el presente documento, o progenie de las mismas, en el que la célula es capaz de expresar una porina de Msp o un monómero de porina de Msp. También se proporciona una cepa de *Mycobacterium smegmatis* que comprende cualquier vector descrito en el presente documento.

10 [0021] También se proporciona una cepa bacteriana mutante capaz de expresión inducible de monómero de Msp, comprendiendo la cepa bacteriana: (a) una delección de una MspA de tipo salvaje; (B) una delección de una MspC de tipo salvaje; (c) una delección de una MspD de tipo salvaje; y (d) un vector que comprende un promotor inducible unido operativamente a una secuencia de ácido nucleico de monómero de Msp.

15 [0022] También se proporciona un procedimiento de producción de una porina de Msp de una cadena, comprendiendo el procedimiento: (a) transformar una cepa bacteriana mutante con un vector que comprende una secuencia de ácido nucleico capaz de codificar una una porina de Msp de una cadena; y opcionalmente (b) purificar la porina de Msp de una cadena de la bacteria. La cepa mutante puede incluir delecciones de una MspA de tipo salvaje, una MspB de tipo salvaje, una MspC de tipo salvaje, y una MspD de tipo salvaje, y un vector que comprende un promotor inducible unido operativamente a una secuencia de ácido nucleico de Msp. La cepa mutante puede ser transformada con un vector que comprende una secuencia de ácido nucleico capaz de codificar una porina de Msp de una cadena.

20 [0023] Se proporcionan además procedimientos de uso de porinas Mps, tales como una porina de Msp de una cadena. Por ejemplo, el procedimiento puede comprender crear una bicapa lipídica que tiene una primera cara y una segunda cara, añadir una porina de Msp, tal como una porina de Msp de una cadena purificada, a la primera cara de la bicapa lipídica, aplicar un voltaje positivo a la segunda cara de la bicapa, trasladar una secuencia de ácido nucleico experimental o secuencia de polipéptidos a través de la porina de Msp, medir la corriente de bloqueo de la secuencia de traslado pasada a través de la porina de Msp, y comparar la corriente de bloqueo experimental con un patrón de bloqueo de corriente y determinar la secuencia experimental.

DESCRIPCIÓN DE LOS DIBUJOS

35 [0024] Los aspectos anteriores y muchas de las ventajas concomitantes resultarán más fácilmente apreciadas, ya que los mismos se entenderán mejor por referencia a la siguiente descripción detallada, cuando se toma conjuntamente con los dibujos adjuntos.

40 La FIGURA 1 muestra la estructura y distribución de carga de porina de MspA de tipo salvaje (WTMspA). A pH 8, se espera que los residuos ácidos estén predominantemente con carga negativa y los residuos básicos estén con carga positiva. Los lugares y las identidades de la mutación se indican con flechas y etiquetas. Véase Faller y otros, Science, 303: 1189 (2004).

45 La FIGURA 2 muestra los resultados de ensayos de actividad formadora de túnel y conductancia de un solo túnel para porinas WTMspA, mutante D90N/D91N/D93N (M1MspA, también llamada M1-NNN), y mutante D90N/D91N/D93N/D118R/E139K/D134R (M2MspA, también llamada M2-NNN). Los paneles de la izquierda muestran la conductancia en la bicapa con el tiempo cuando una porina de MspA está presente en la solución (1 M KCl, 20°C) que baña la bicapa. Aumentos graduales en la conductancia se interpretan como inserciones de porinas MspA en la bicapa. A la derecha están los histogramas de los tamaños de estas etapas de conductancia. Los histogramas de porina WTMspA, M1MspA, y M2MspA resumen 40 inserciones de 3 experimentos repetidos, 144 inserciones de 3 experimentos repetidos, y 169 inserciones de 5 experimentos repetidos, respectivamente.

50 Las FIGURAS 3A y 3B muestran el comportamiento espontáneo de bloqueo de porinas WTMspA. La FIGURA 3A es un diagrama esquemático de experimentos. La FIGURA 3B muestra señales de corriente iónica representativas observadas para porinas WTMspA a 60 mV (izquierda) y 100 mV (derecha) sin ADN presente. Los intervalos de flujo de corriente negativa corresponden a la inversión de la tensión aplicada, que a menudo se requería para restablecer el nivel de corriente iónica no bloqueado.

55 La FIGURA 4 muestra la expresión de monómeros de MspA mutantes en un gel electroforético. Se añadió extracto crudo (13 µl) a cada carril. El gel se tiñó con azul de Coomassie. Carril 1: marcador de masa de proteína; Carril 2: WTMspA; Carril 3: sin MspA; Carril 4: mutante M1MspA; Carril 5: mutante D90N/D91N/D93N/D118R; Carril 6: mutante D90N/D91N/D93N/D118R/E139R; Carril 7: mutante D90N/D91N/D93N/D118R/E139K; Carril 8: mutante M2MspA. Los mutantes en los carriles 5-7 se construyeron, extrajeron y estudiaron para garantizar que la expresión y la actividad de formación de túnel se mantenían para cada sustitución sucesiva de aminoácidos. Los diagramas por encima del gel muestran esquemáticamente la ubicación aproximada y la polaridad de los aminoácidos mutados en este experimento.

Las FIGURAS 5A-5C muestran la detección de construcciones de ssADN en horquilla con porinas M1MspA. La FIGURA 5A es un diagrama esquemático de experimentos. La FIGURA 5B muestra la señal de corriente iónica representativa observada para porinas M1MspA en ausencia de ADN y presencia de 8 μ M de ADN en horquilla hp08 M (SEQ ID NO: 4) a 180 y 140 mV. La FIGURA 5C muestra bloqueos numerados de trazas en la FIGURA 5B a escalas de tiempo más amplias.

La FIGURA 6 muestra las características de bloqueos profundos de las construcciones de horquilla en la porina M1MspA. Las coordenadas de cada punto proporcionan la duración y la corriente media de 1 bloqueo profundo. Los datos en negro y gris se adquirieron a 140 y 180 mV, respectivamente. El modo del \log_{10} de los tiempos de permanencia del bloqueo profundo, t_D , se indica para cada conjunto de datos. Los diagramas de la derecha muestran la secuencia de cada construcción de horquilla: hp08 (5' GCTGTTGC TCTCTC GCAACAGC A₅₀ 3') (SEQ ID NO: 4), hp10 (5' GCTCTGTTGC TCTCTC GCAACAGAGC A₅₀ 3') (SEQ ID NO: 5), y hp12 (5' GCTGTCTGTTGC TCTCTC GCAACAGACAGC A₅₀-3') (SEQ ID NO: 6).

La FIGURA 7 es un gráfico que muestra las distribuciones de los tiempos del bloqueo parcial para hp08 (SEQ ID NO: 4) en la porina M1MspA. Las distribuciones están bien ajustados por exponenciales individuales. Los bloqueos parciales a 180 mV tienen una constante de tiempo que es un factor de ~ 3 más largo que a 140 mV.

La FIGURA 8 proporciona una visión detallada de las distribuciones del tiempo de permanencia de los bloqueos profundos de las construcciones de horquilla en la porina M1MspA. Los paneles de la izquierda muestran los histogramas de tiempo con intervalos espaciados logarítmicamente (representaciones en escalera) y las correspondientes estimaciones de la densidad Kernel suavizado de la distribución de probabilidad del \log_{10} de los tiempos de permanencia (x). El máximo de estas estimaciones de densidad suavizadas, t_D , se utilizó para parametrizar las distribuciones del tiempo de permanencia. Las líneas verticales muestran los valores de t_D . Los paneles de la derecha muestran las curvas de probabilidad de supervivencia derivadas de los datos de tiempo de permanencia (líneas continuas) y exponenciales de descomposición individual, con constantes de tiempo establecidas a los valores de t_D de cada conjunto de datos (líneas de puntos). Los datos se desvían claramente del comportamiento exponencial simple. Sin embargo, es razonable hacer comparaciones cualitativas entre el valor t_D y constantes de tiempo exponenciales utilizadas en otras investigaciones (Kasianowicz et al, Proc Natl Acad Sci USA, 93: 13 770 (1996)), ya que ambos parámetros reflejan aspectos similares de las distribuciones del tiempo de permanencia.

Las FIGURAS 9A-9G muestran los datos adquiridos a partir de experimentos de sonda transbicapa. La FIGURA 9A muestra la animación de configuraciones moleculares: (1) un poro no bloqueado; (2) un ssADN enroscado con neutravidina (nA) que detiene el traslado del complejo nA-ssADN; (3) ADN diana hibridado con nA-ssADN que se disocia a voltaje negativo; y (4) el complejo nA-ssADN que sale del poro a una tensión que depende de la hibridación del ADN diana. La FIGURA 9B es una serie con el tiempo de la tensión aplicada. Un bloqueo de corriente provoca un cambio de la tensión de captura de 180 mV a una tensión de mantenimiento de 40 mV después de un retraso de ~ 200 ms. La tensión de mantenimiento se mantiene durante 5 segundos para permitir la hibridación, y a continuación se realiza una rampa negativa. Las FIGURAS 9C y 9D muestran cada una series con el tiempo de la corriente que demuestran una salida de nA-ssADN a voltajes negativos y positivos, respectivamente. Los grandes picos de corriente se producen debido a cambios de voltaje instantáneos y al cierre espontáneo de los poros a grandes voltajes negativos. Las FIGURAS 9E-9G son histogramas de voltaje de salida (V_{salida}). La FIGURA 9E muestra un experimento en el que la sonda, 5'-C₆A₅₄-CTCTATTCTTATCTC-3' (SEQ ID NO: 7, era complementaria a las moléculas de ssADN diana, 5'-GAGATAAGAATAGAG-3' (SEQ ID NO: 9). La FIGURA 9F muestra el mismo poro que en la FIGURA 9E, pero con una sonda, 5'-C₆A₅₄-CACACACACACACAC-3' (SEQ ID NO: 8), que no es complementaria al ADN diana. La FIGURA 9G muestra los resultados de un control separado utilizando la misma sonda (SEQ ID NO: 7) que en la FIGURA 9E, pero sin ADN diana presente en el compartimiento trans. Se observa un número significativo de sucesos de V_{salida} negativos sólo en la FIGURA 9E, donde la sonda (SEQ ID NO: 7) es complementaria a la diana. La aparición poco frecuente de sucesos de V_{salida} negativo en las FIGURAS 9F y 9G descartan la posibilidad de que una mayoría de V_{salida} negativo en la FIGURA sea causada por una asociación no específica de sonda-diana o por la unión de la sonda al poro.

Las FIGURAS 10A-10C comparan los bloqueos de homopolímeros de dT₅₀ (SEQ ID NO: 32) para porinas M1MspA y M2MspA. La FIGURA 10A es un diagrama esquemático de experimentos. La FIGURA 10B muestra señales representativas de corriente iónica observadas para la porina M1MspA con 8 μ M de dT₅₀ (izquierda) y la porina M2MspA con 2 μ M de dT₅₀ (derecha). La FIGURA 10C muestra los bloqueos numerados de trazas en la FIGURA 10B a escalas de tiempo más amplias.

La FIGURA 11 muestra características estadísticas de bloqueos de dT₅₀ (SEQ ID NO: 32) en la porina M2MspA. Comparación de la estructura media al inicio y al final de los bloqueos. La FIGURA fue creada mediante la superposición de los sucesos en un archivo de datos alineados al inicio del suceso (izquierda) y al final del suceso (derecha). Se muestra la tendencia de los bloqueos para terminar con una breve desviación hacia abajo de la corriente iónica, junto con el aumento de esta tendencia con voltaje.

La FIGURA 12A muestra los histogramas de los niveles de corriente de bloqueo en la porina M1MspA bloqueada por construcciones de ADN. Las construcciones de ADN de arriba a abajo: 3'-A₄₇AAC-hp-5' (SEQ ID NO: 14); 3'-A₄₇ACA hp-5' (SEQ ID NO: 33); 3'-A₄₇CAA-hp-5' (SEQ ID NO: 13); 3'-C₅₀-hp-5' (SEQ ID NO: 16); 3'-A₅₀-hp-5' (SEQ ID NO: 10). La FIGURA 12B muestra un gráfico de los niveles de corriente escalados a la diferencia entre los niveles de poli-C (= 1,0) y poli-A (=0.0) frente a la posición C sola. Un ajuste gaussiano sugiere que la posición de reconocimiento para una C sola es $1,7 \pm 0,8$ nucleótidos (nt) desde el extremo de la horquilla.

La FIGURA 13 muestra una serie de histogramas de corriente de ADN que bloquea la porina M1-NNN MspA (también llamada M1MspA). Las construcciones de ADN de arriba a abajo: 3'-C₅₀-hp-5' (SEQ ID NO: 16); 3'-A₅₀-hp-5' (SEQ ID NO: 10); 3'-T₄₇TTT-hp-5' (SEQ ID NO: 17); 3'-A₄₇AAT-hp-5' (SEQ ID NO: 34); 3'-A₄₇47ATA-hp-5' (SEQ ID NO: 35); 3'-A₄₇TAA-hp-5' (SEQ ID NO: 36); 3'-C₄₇CCA-hp-5' (SEQ ID NO: 37); 3'-C₄₇CAC-hp-5' (SEQ ID NO: 38); 3'-C₄₇ACC-hp-5' (SEQ ID NO: 39). Cada construcción o mezcla se muestra a la izquierda. El conjunto de sucesos en cada histograma se muestra a la derecha. Panel superior: "mezcla de Calibración" (poli-A-hp y poli-C-hp). Paneles 2-5: poli-T-hp y bases T solas en un fondo de poli-A. Tres paneles inferiores: bases A solas en un fondo de poli-A. Poli-A-hp se incluye en la mezcla de referencia (pequeño pico a 19,5%). Todos los datos son con aplicación de 180 mV.

La FIGURA 14 demuestra que la cola de ADN no afecta a las propiedades de reconocimiento. La leyenda es como para la FIGURA 13. Dos colas heterogéneas ('ran1' (SEQ ID NO: 51), 'ran2' (SEQ ID NO: 52), cada una de 47 bases) están unidas a trinucleótidos y la horquilla. El panel central muestra el histograma de corriente que resulta cuando una mezcla de ADN A₅₀-hp (SEQ ID NO: 10) y ADN ran1-C₃-hp se aplica al poro, un punto de referencia para los otros paneles. Los niveles de corriente son idénticos a los de las colas de A₅₀ o C₅₀. Todos los datos son con una aplicación de 180 mV.

Las FIGURAS 15A y 15B muestran los datos de caracterización de la porina M2-qqn, otra porina de MspA mutante. La FIGURA 15A presenta el nivel de expresión de este mutante. Todas las proteínas se expresaron en *M. smegmatis* ML16. Se cargaron 10 μ l de extracto crudo de octilpolioxi-etileno al 0,5% en cada pocillo. Carril 1: WT MspA; Carril 2: Fondo (pMS2, vector vacío); Carril 3: M2-QQN (pML866). La FIGURA 15B muestra trazas de corriente de la porina M2-QQN en una bicapa lipídica de difitanoilfosfatidilcolina que se registraron en KCl 1 M. Se añadieron aproximadamente 70 pg de proteína a la cámara de la bicapa. Aproximadamente 100 poros de cuatro membranas se analizaron en experimentos de bicapa lipídica. La conductancia principal de la porina M2-QQN es de 2,4 nanosegundos (ns).

La FIGURA 16 muestra histogramas de corriente de bloqueo con tres porinas MspA mutantes diferentes expuestas a mezclas de ADN en horquilla de hp-T₅₀ (SEQ ID NO: 17), hp-C₅₀ (SEQ ID NO: 16), y hp-A₅₀ (SEQ ID NO: 10). En cada caso, las corrientes se normalizan a la corriente de estado abierto, mostrada a la derecha para cada mutante. hp-C₅₀ y hp-A₅₀ se utilizaron como una mezcla, y T₅₀ se utilizó por separado.

La FIGURA 17 es un gráfico que muestra la probabilidad de supervivencia de los bloqueos profundos de corriente de dos porinas MspA mutantes. Se muestra la probabilidad de sucesos que duran más de t. Los círculos indican la porina M2-QQN, y las cruces indican la porina M2-NNN. Los voltajes aplicados a través de las bicapas fueron 100, 120, y 140 mV. Se normalizan los datos para el número total de sucesos en cada registro.

La FIGURA 18 muestra un alineamiento de monómeros de MspA, MspB, MspC y MspD de *M. smegmatis*. El primer codón ATG o GTG de los marcos de lectura abierta se tomaron como el supuesto codón de inicio. La numeración de la proteína comienza con el primer aminoácido de la parte madura. La secuencia de aminoácidos del monómero de MspA es SEQ ID NO: 28, la secuencia de aminoácidos del monómero de MspB es SEQ ID NO: 29, la secuencia de aminoácidos del monómero MspC es SEQ ID NO: 30, y la secuencia de aminoácidos del monómero MspD es SEQ ID NO: 31.

La FIGURA 19 es una imagen de un gel que muestra la eliminación de cada uno de los genes de porina en el mutante cuádruple de porina ML59 de *M. smegmatis*.

La FIGURA 20 muestra una transferencia Western que demuestra la expresión de la porina de Msp en porinas mutantes de *M. smegmatis* y *M. smegmatis*. El carril 1 es una dilución 1:10 de extracto de proteína para la *M. smegmatis* de tipo salvaje, carril 2 es el mutante MN01 (Δ mspA), carril 3 es el mutante ML10 (Δ mspAC), carril 4 es el mutante ML16 (Δ mspACD), y el carril 5 es el mutante ML180 (Δ mspABCD).

Las FIGURAS 21A y 21B muestran mapas de los plásmidos para la construcción de un mutante cuádruple de porina. Hyg: gen de resistencia a higromicina; ColE1: origen de la replicación de *E. coli*. La FIGURA 21A es el mapa de plásmido de integración para la expresión de MspA. Se requieren AmiC, A, D, S para la expresión inducible por acetamida de MspA. attP: sitio de unión a cromosoma del fago L5; int: L5 integrasa; FRT: sitio de Flp recombinasa. La FIGURA 21B es el mapa del plásmido para el vector de delección MspB. MspBup, MspBdown: regiones aguas arriba y aguas abajo de MspB; loxP: Cre sitio de recombinación; SacB: levansucrasa; XylE: catecol-2,3-dioxigenasa; Gfp2+: proteína fluorescente verde; tsPAL5000: origen sensible a la temperatura de la replicación para micobacterias.

La FIGURA 22 es una imagen de un gel teñido con azul de Coomassie que muestra la expresión inducible de monómeros de MspA en *M. smegmatis*.

5 La FIGURA 23 es una imagen que demuestra el crecimiento del mutante cuádruple de Msp ML705 en placas de agar Middlebrook 7H10.

La FIGURA 24 es un gráfico que muestra la tasa de crecimiento de ML705 en medio líquido rico.

10 La FIGURA 25 es una imagen de una transferencia Western que demuestra la expresión de monómeros de MspA en el mutante cuádruple ML705 tras la inducción con acetamida. El carril 1 es de *M. smegmatis* de tipo salvaje, el carril 2 es la cepa mutante cuádruple ML705 con acetamida, el carril 3 es la cepa mutante de msp cuádruple ML705 sin acetamida, y el carril 4 es la cepa mutante triple ML16. Las proteínas se detectaron utilizando un anticuerpo policlonal para MspA.

15 Las FIGURAS 26A-26D muestran la estructura y actividad de túnel del dímero de nanoporos de MspA de una cadena. La FIGURA 26A es una imagen del modelo molecular del dímero de nanoporos de MspA de una cadena. La FIGURA 26B muestra el esquema de la construcción del gen del de nanoporos de MspA de una cadena (scMspA). La región de aminoácidos enlazadora (GGGGS)₃ (SEQ ID NO: 3) está ampliada. También se muestra la secuencia de ADN del enlazador de aminoácidos: (5 'GGCGGTGGCGGTAGCGGCGGTG
20 GCGGTAGCGGCGGTGGCGGTAGC-3') (SEQ ID 19 NO). La FIGURA 26C es una imagen de una transferencia Western que demuestra la expresión del dímero de nanoporos de scMspA en *M. smegmatis*. El carril 1 es el marcador de masa molecular (M), el carril 2 es *M. smegmatis* de tipo salvaje (WT *Msmeg*), el carril 3 es la cepa ML16 sin la construcción génica de scMspA (ML16), el carril 4 es la cepa ML16 con una construcción génica de MspA de tipo salvaje (WTMspA), y el carril 5 es la cepa ML16 con la construcción génica de dímeros de nanoporos de scMspA (scMspA). La FIGURA 26D muestra una traza de corriente para el dímero de nanoporos de scMspA.

25 La FIGURA 27 muestra un esquema de transporte de ssADN dC₅₈ (SEQ ID NO: 40) a través de la porina de MspA de tipo salvaje. El transporte de ADN se compone de los siguientes pasos: a) inicio de la simulación; b) y c) conformaciones de ADN antes y después del rápido avance; y d) el ADN se adhiere a la superficie de la porina de MspA.
30

La FIGURA 28 es un gráfico que muestra la corriente iónica acumulada del transporte de ssADN dC₅₈ (SEQ ID NO: 40) de la FIGURA 27. El transporte se realizó bajo un margen transmembrana de 1,2V.

35 La FIGURA 29 muestra el diseño de la secuencia de octámeros de nanoporos de MspA de una cadena (scMspA). El octámero dr scMspA consiste en: un monómero de gen de MspA de tipo salvaje, un monómero de MspA1, un monómero de MspA2, un monómero de MspA3, un monómero de MspA4, un monómero de MspA5, un monómero de MspA6, y un monómero de MspA7. Los sitios de restricción Pacl y HindIII flanquean la secuencia de octámeros de nanoporos de scMspA. X1-X14 son sitios de restricción únicos que flanquean las secuencias monoméricas individuales. Las líneas negras que conectan cada monómero representan el enlazador (GGGGS)₃ (SEQ ID NO: 3).
40

La FIGURA 30 muestra la zona de constricción (la caja rectangular) de un monómero de MspA de tipo salvaje y una variedad de monómeros de MspA parálogos y homólogos.

45 La FIGURA 31 muestra histogramas de los niveles de corriente de bloqueo en M1MspA bloqueadas por construcciones de ADN. La construcciones de ADN de arriba a abajo: 3'-A₄₀AAAAAAAAAAAA-hp-5' (SEQ ID NO: 10); 3'-A₄₀CCCCAAAAAA-hp-5' (SEQ ID NO: 11); 3'-A₄₀AAACCCCAAA-hp-5' (SEQ ID NO: 12); 3'-A₄₀AAAAAAACAA-hp-5' (SEQ ID NO: 13); 3'-A₄₀AAAAAAAAC-hp-5' (SEQ ID NO: 14); 3'-A₄₀AAAAAACCCC-hp-5' (SEQ ID NO: 15); 3'-C₄₀40CCCCCCCC-hp-5' (SEQ ID NO: 16); 3'-T₄₀TTTTTTTTTTT-hp-5' (SEQ ID NO: 17); 3'-A₄₀AAAAAAGGG-hp-5' (SEQ ID NO: 18).
50

DESCRIPCIÓN DETALLADA

55 **[0025]** La presente invención se define según las reivindicaciones.

[0026] El presente documento proporciona un procedimiento que comprende aplicar un campo eléctrico a una porina de porinas de *Mycobacterium smegmatis* (Msp) que tiene un vestíbulo y una zona de constricción que definen un túnel, en el que la porina de Msp se encuentra entre un primer medio líquido conductor y un segundo medio líquido conductor. Opcionalmente, el primer y segundo medio líquido conductor son los mismos. Opcionalmente, el primer y segundo medio líquido conductor son diferentes. La porina de Msp puede ser cualquier porina de Msp descrita en el presente documento. Por ejemplo, la porina de Msp se puede seleccionar del grupo que consiste en una porina de MspA de tipo salvaje, una porina de MspA mutante, una porina de MspA de tipo salvaje paróloga u homóloga, y una porina de MspA mutante paróloga u homóloga.
60

65 **[0027]** En cualquier realización en el presente documento, una porina de Msp puede comprender además un motor molecular. El motor molecular puede ser capaz de mover un analito en o a través de un túnel con una velocidad de

5 traslado (traslado) o una velocidad de traslado promedio que es menor que la velocidad de traslado o la velocidad de traslado promedio a la que el analito se traslada por electroforesis en o a través del túnel en ausencia del motor molecular. En consecuencia, en cualquier realización en el presente documento que comprende la aplicación de un campo eléctrico, el campo eléctrico puede ser suficiente para hacer que el analito se traslade por electroforesis a través del túnel.

10 [0028] Cualquier medio líquido descrito en el presente documento, tal como un medio líquido conductor, puede comprender un analito. El analito puede ser cualquier analito aquí descrito. Las realizaciones en la presente memoria pueden comprender además la detección del analito, tal como en un procedimiento que comprende medir una corriente iónica, ya que el analito interacciona con un túnel de porinas de Msp para proporcionar un patrón de corriente, en el que la aparición de un bloqueo en el patrón de corriente indica la presencia del analito.

15 [0029] Opcionalmente, una porina de Msp es una porina de MspA mutante o una porina homóloga o paróloga de MspA mutante, y el analito tiene una velocidad de traslado o una velocidad de traslado promedio a través del túnel de porinas que es menor que, o es mayor que, la velocidad de traslado o la velocidad de traslado promedio del analito a través del túnel de una porina de MspA de tipo salvaje o una porina paróloga u homóloga de MspA de tipo salvaje.

20 [0030] En cualquier realización en el presente documento, un analito puede tener una velocidad de traslado o una velocidad de traslado promedio a través de un túnel de menos de 0,5 nm/ μ s. Opcionalmente, un analito puede tener una velocidad de traslado o una velocidad de traslado promedio a través de un túnel de menos de 0,05 nm/ μ s.

25 [0031] Cualquier porina de Msp descrita en el presente documento puede estar comprendido en una bicapa lipídica. En tales realizaciones o cualquier otra realización en el presente documento, la porina de Msp puede tener una cara cis y una cara trans. Opcionalmente, un analito por electroforesis o de otra manera se traslada desde la cara cis a través de un túnel a la cara trans. Opcionalmente, un analito por electroforesis o de otra manera se traslada desde la cara trans a través de un túnel a la cara cis. Opcionalmente, un analito por electroforesis o de otra manera es inducido desde la cara cis o la cara trans en un túnel y se mantiene en el túnel o a continuación, se retrae hacia la cara cis o la cara trans, respectivamente.

30 [0032] Cualquier realización en el presente documento puede comprender además la identificación de un analito. Tales procedimientos pueden comprender comparar el patrón de corriente obtenido con respecto a un analito desconocido al patrón de corriente conocido obtenido usando un analito conocido bajo las mismas condiciones.

35 [0033] En cualquier realización en el presente documento, un analito puede ser un nucleótido, un ácido nucleico, un aminoácido, un péptido, una proteína, un polímero, un fármaco, un ion, un contaminante, un objeto nanoscópico, o un agente de guerra biológica. Opcionalmente, un analito es un polímero, tal como una proteína, un péptido, o un ácido nucleico. Opcionalmente, el polímero es un ácido nucleico. Opcionalmente, un ácido nucleico tiene una velocidad de traslado o una velocidad de traslado promedio a través de un túnel de menos de 1 nucleótido/ μ s. Opcionalmente, un ácido nucleico tiene una velocidad de traslado o una velocidad de traslado promedio a través del túnel de menos de 0,1 nucleótido/ μ s. Un ácido nucleico puede ser ssADN, dsADN, ARN, o una combinación de los mismos.

45 [0034] Las realizaciones en el presente documento pueden comprender distinguir al menos una primera unidad dentro de un polímero de al menos una segunda unidad dentro del polímero. Distinguir puede comprender la medición de la corriente de iones producida como la primera y segunda unidades trasladadas por separado a través de un túnel para producir un primer y un segundo patrón de corriente, respectivamente, donde el primer y segundo patrones de corriente difieren entre sí.

50 [0035] Las realizaciones en el presente documento pueden comprender, además, secuenciar un polímero. La secuenciación puede comprender la medición de la corriente de iones o de señales ópticas, ya que cada unidad del polímero se traslada por separado a través del túnel para proporcionar un patrón de corriente que está asociado con cada unidad, y la comparación de cada patrón de corriente con el patrón de corriente de una unidad conocida obtenida bajo la mismas condiciones, de manera que se secuencian el polímero.

55 [0036] Cualquier realización en el presente documento puede comprender además la determinación de la concentración, el tamaño, el peso molecular, la forma o la orientación de un analito, o cualquier combinación de los mismos. Cualquier medio líquido descrito en el presente documento, tal como un medio líquido conductor, puede comprender una pluralidad de analitos. Cualquier analito descrito en el presente documento puede comprender una microesfera óptica o una microesfera magnética.

60 [0037] Cualquier porina de Msp descrita en el presente documento puede definirse además como una porina de MspA mutante. Una porina de MspA mutante puede comprender un vestíbulo y una zona de constricción que definen un túnel, y al menos un primer monómero de MspA mutante que comprende una mutación en la posición 93, 91, 90, o cualquier combinación de los mismos. Un porina de MspA mutante puede comprender una mutación en las posiciones 93 y 91; las posiciones 93 y 90; las posiciones 91 y 90; o posiciones 93, 90, y 91. Una MspA mutante de

la invención comprende una mutación en las posiciones 90, 91 y 93, en el que la mutación en la posición 90 es D90N, la mutación en la posición 91 es D91N, y la mutación en la posición 93 es D93N. Opcionalmente, una porina de MspA mutante comprende una o más mutaciones en cualquiera de las siguientes posiciones de aminoácidos: 88, 105, 108, 118, 134, o 139, o cualquier otra mutación descrita en el presente documento.

5 **[0038]** En cualquier realización en el presente documento, el diámetro de una porina de MspA mutante o un parólogo u homólogo de MspA mutante puede ser menor que el diámetro de la zona de constricción de una correspondiente porina de MspA de tipo salvaje o un parólogo u homólogo de MspA de tipo salvaje. Una porina de MspA mutante o parólogo u homólogo de MspA mutante pueden tener una mutación en el vestíbulo o en la zona de constricción que permite que un analito se traslade, por electroforesis o de otra manera, a través del túnel de la porina de MspA mutante o parólogo u homólogo de MspA mutante con una velocidad de traslado o una velocidad de traslado promedio que es menor que la velocidad de traslado o la velocidad de traslado promedio a la que el analito se traslada a través del túnel de una porina de Msp de tipo salvaje o un parólogo u homólogo de MspA de tipo salvaje.

15 **[0039]** Una porina de Msp mutante, tal como una porina de MspA mutante o una porina homóloga o paróloga de MspA mutante, puede comprender una zona de constricción neutral. Una porina de Msp mutante, tal como una porina de MspA mutante o una porina homóloga o paróloga de MspA mutante, puede comprender una conductancia a través del túnel que es superior, por ejemplo, dos veces mayor, que la conductancia a través del túnel de su correspondiente de porina de Msp de tipo salvaje. Una porina de Msp mutante, tal como una porina de MspA mutante o una porina homóloga o paróloga de MspA mutante, puede comprender una conductancia a través del túnel que es menor que la conductancia a través del túnel de su correspondiente porina de Msp de tipo salvaje.

20 **[0040]** Cualquier porina de Msp descrita en el presente documento puede comprender un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en el que el vestíbulo y la zona de constricción juntos definen un túnel. También se proporciona en el presente documento una porina de MspA mutante que comprende un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en el que el vestíbulo y la zona de constricción juntos definen un túnel, y que comprende además al menos un parólogo u homólogo del primer monómero de MspA mutante.

25 **[0041]** El diámetro de la zona de constricción de una porina de Msp mutante, tal como una porina de MspA mutante u homólogo o parólogo de MspA mutante, puede ser menor que el diámetro de la zona de restricción de su correspondiente porina de Msp de tipo salvaje, tal como una porina de MspA de tipo salvaje o un parólogo u homólogo de MspA de tipo salvaje. Una porina de Msp mutante, tal como una porina de MspA mutante o un parólogo u homólogo de MspA mutante, puede comprender una mutación en el vestíbulo o en la zona de constricción que permite que un analito se traslade, por electroforesis o de otra manera, a través del túnel de la porina con una velocidad de traslado o una velocidad de traslado promedio que es menor que la velocidad de traslado o la velocidad de traslado promedio a la que el analito se traslada a través del túnel de su correspondiente porina de Msp de tipo salvaje, (por ejemplo, porina de MspA de tipo salvaje, parólogo u homólogo de MspA de tipo salvaje).

30 **[0042]** Opcionalmente, una porina de Msp está codificada en su totalidad o en parte por una secuencia de ácido nucleico que codifica una porina parcial o completa de Msp de una cadena, en el que la secuencia de ácido nucleico comprende: (a) una primera y segunda secuencia de nucleótidos, en la que el primera secuencia de nucleótidos codifica una primera secuencia de monómero de Msp y la segunda secuencia de nucleótidos codifica una segunda secuencia de monómero de Msp; y (b) una tercera secuencia de nucleótidos que codifica una secuencia de aminoácidos enlazadora. Las secuencias de monómero pueden ser cualquier secuencia de monómero descrita en el presente documento. Opcionalmente, la primera y segunda secuencias de monómero de Msp se seleccionan independientemente del grupo que consiste en un monómero de MspA de tipo salvaje, un monómero de MspB de tipo salvaje, un monómero de MspC de tipo salvaje, un monómero de MspD de tipo salvaje, y los mutantes de los mismos. Opcionalmente, la primera secuencia de monómero de Msp comprende un monómero de MspA de tipo salvaje o un mutante de la misma. Opcionalmente, la primera secuencia de monómero de Msp comprende un monómero de MspA mutante.

35 **[0043]** En cualquier realización en el presente documento, una porina de Msp puede estar codificada en su totalidad o en parte por una secuencia de ácido nucleico que codifica una porina parcial o completa de Msp de una cadena, en la que la secuencia de ácido nucleico comprende: (a) una primera, segunda, tercera, cuarta, quinta, sexta, séptima, y octava secuencia de nucleótidos o cualquier subconjunto de las mismas, en las que la primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de nucleótidos codifican una primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de monómero de Msp, respectivamente; y (b) una novena secuencia de nucleótidos que codifica una secuencia de aminoácidos enlazadora. Por lo tanto, la porina puede comprender una o más porinas parciales de Msp de una cadena que se hibridan, dimerizan, trimerizan, o similares

con otros monómeros de Msp u otros porinas parciales de Msp de una cadena. Como alternativa, la porina completa de Msp de una cadena puede formar una porina sin asociar con otros elementos de Msp. En cualquier realización en el presente documento, por ejemplo, una porina de Msp puede ser codificada por una secuencia de ácido nucleico que codifica una porina de Msp de una cadena, en la que la secuencia de ácido nucleico comprende: (a) una primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de nucleótidos, en la que la primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de nucleótidos codifican una primera, segunda, tercera cuarta, quinta, sexta, séptima y octava secuencia de monómero de Msp, respectivamente; y (b) una novena secuencia de nucleótidos que codifica una secuencia de aminoácidos enlazadora. Cada monómero de Msp puede comprender un monómero de MspA de tipo salvaje o un mutante del mismo. Opcionalmente, al menos un monómero de Msp comprende un monómero de MspA de tipo salvaje o un mutante del mismo. Por lo tanto, la porina puede ser codificada en su totalidad.

[0044] En cualquier realización en el presente documento, un monómero de Msp puede ser un parólogo u homólogo de MspA de tipo salvaje, tal como MspA/Msmeg0965, MspB/Msmeg0520, MspC/Msmeg5483, MspD/Msmeg6057, MppA, PorM1, PorM2, PorM1, Mmcs4296, Mmcs4297, Mmcs3857, Mmcs4382, Mmcs4383, Mjls3843, Mjls3857, Mjls3931 Mjls4674, Mjls4675, Mjls4677, Map3123c, Mav3943, Mvan1836, Mvan4117, Mvan4839, Mvan4840, Mvan5016, Mvan5017, Mvan5768, MUL_2391, Mflv1734, Mflv1735, Mflv2295, Mflv1891, MCH4691c, MCH4689c, MCH4690c, MAB1080, MAB1081, MAB2800, RHA1 ro08561, RHA1 ro04074 y RHA1 ro03127.

[0045] También se proporciona en el presente documento un procedimiento de modificación de la conductancia a través del túnel de una porina de Msp que comprende eliminar, añadir, o sustituir al menos un aminoácido en el vestíbulo o la zona de constricción de una porina de Msp de tipo salvaje. Por ejemplo, el procedimiento puede comprender el aumento de la conductancia. El procedimiento puede comprender la disminución de la conductancia.

[0046] También se proporciona un procedimiento que comprende el traslado de un analito a través de un túnel de una porina de Msp sin el empleo de un campo eléctrico. En este o cualquier otra realización en el presente documento, una porina de Msp puede comprender además un motor molecular. La porina de Msp puede ser cualquier porina de Msp descrita en el presente documento, tal como una porina de MspA de tipo salvaje, una porina de MspA mutante, una porina paróloga u homóloga de MspA de tipo salvaje y una porina paróloga u homóloga de MspA mutante. La porina de Msp puede ser codificada por una secuencia de ácido nucleico que codifica una porina de Msp de una cadena.

[0047] También se proporciona un sistema que comprende una porina de Msp que tiene un vestíbulo y una zona de constricción que definen un túnel, en el que el túnel se encuentra entre un primer medio líquido y un segundo medio líquido, en el que al menos un medio líquido comprende un analito, y en el que el sistema es operativo para detectar una propiedad del analito. Un sistema puede ser operativo para detectar una característica de cualquier analito que comprende someter una porina de Msp a un campo eléctrico de manera que el analito interacciona con la porina de Msp. Un sistema puede ser operativo para detectar una propiedad del analito que comprende someter la porina de Msp a un campo eléctrico de manera que el analito se traslada por electroforesis a través del túnel de la porina de Msp. También se proporciona un sistema que comprende una porina de Msp que tiene un vestíbulo y una zona de constricción que definen un túnel, en el que el túnel se encuentra en una bicapa lipídica entre un primer medio líquido y un segundo medio líquido, y en el que el único punto de comunicación de líquidos entre el primer y segundo medios líquidos se produce en el túnel. Además, cualquier porina de Msp descrita en el presente documento puede estar comprendida en cualquier sistema que se describe en el presente documento.

[0048] El primer y segundo medios líquidos pueden ser el mismo o diferentes, y, o bien uno o ambos pueden comprender uno o más de una sal, un detergente o un tampón. De hecho, cualquier medio líquido que se describe en el presente documento puede comprender uno o más de una sal, un detergente o un tampón. Opcionalmente, al menos un medio líquido es conductor. Opcionalmente, al menos un medio líquido no es conductor. Cualquier medio líquido descrito en el presente documento puede comprender una sustancia que altera la viscosidad o una sustancia que altera la velocidad. Los medios líquidos pueden comprender cualquier analito descrito en el presente documento. Una propiedad de un analito puede ser una propiedad eléctrica, química o física.

[0049] Una porina de Msp puede estar comprendida en una bicapa lipídica en un sistema o cualquier otra realización descrita en el presente documento. Un sistema puede comprender una pluralidad de porinas de Msp.

[0050] Un sistema puede comprender cualquier porina de Msp descrita en el presente documento, tal como una porina de MspA de tipo salvaje, una porina de MspA mutante, una porina paróloga u homóloga de MspA de tipo salvaje, o una porina paróloga u homóloga de MspA mutante. Opcionalmente, la porina de Msp se define además como una porina de MspA mutante. Un sistema puede comprender una porina de Msp mutante que comprende un vestíbulo y una zona de constricción que definen un túnel, y al menos un primer monómero de MspA mutante que comprende una mutación en la posición 93 y una mutación en la posición 90, posición 91, o en ambas posiciones 90 y 91. Una porina de Msp mutante comprendida en un sistema puede comprender un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en la que el vestíbulo y la zona de constricción juntos definen un

túnel. Una porina de MspA mutante puede comprender además al menos un primer monómero parálogo u homólogo de MspA mutante. Una porina de Msp comprendida en un sistema puede ser codificada por una secuencia de ácido nucleico que codifica una porina de Msp de una cadena.

5 **[0051]** Una porina de Msp comprendida en un sistema puede comprender además un motor molecular. El motor molecular en un sistema o cualquier otra realización en el presente documento puede ser capaz de mover un analito en o por un túnel con una velocidad de traslado o una velocidad de traslado promedio que es menor que la velocidad de traslado o la velocidad de traslado promedio a la que el analito se traslada en o a través del túnel en ausencia del motor molecular.

10 **[0052]** Cualquier sistema descrito en el presente documento puede comprender además un amplificador de "patch-clamp" o un dispositivo de adquisición de datos. Un sistema puede comprender además uno o más dispositivos de regulación de la temperatura en comunicación con el primer medio líquido, el segundo medio líquido, o ambos.

15 **[0053]** Cualquier sistema descrito en el presente documento puede ser operativo para trasladar un analito a través de un túnel de porinas de Msp ya sea por electroforesis o de otra manera.

[0054] También se proporciona una porina de Msp que comprende un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en la que el vestíbulo y la zona de constricción juntos definen un túnel. También se proporciona una porina de Msp mutante que comprende un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en el que el vestíbulo y la zona de constricción juntos definen un túnel. También se proporciona una porina de MspA mutante que comprende un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en el que el vestíbulo y la zona de constricción juntos definen un túnel. También se proporciona una porina paróloga u homóloga de MspA mutante que comprende un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en el que el vestíbulo y la zona de constricción juntos definen un túnel. Cualquier parálogo u homólogo de MspA mutante descrito en el presente documento puede comprender además al menos un primer monómero parálogo u homólogo de MspA mutante. También se proporciona una porina de MspA mutante que comprende un vestíbulo que tiene una longitud de aproximadamente 2 a aproximadamente 6 nm y un diámetro de aproximadamente 2 a aproximadamente 6 nm, y una zona de constricción que tiene una longitud de aproximadamente 0,3 a aproximadamente 3 nm y un diámetro de aproximadamente 0,3 a aproximadamente 3 nm, en el que el vestíbulo y la zona de constricción juntos definen un túnel, y que comprende además al menos un primer monómero parálogo u homólogo de MspA mutante. Cualquiera de estas porinas puede emplearse en cualquier realización en el presente documento.

[0055] También se proporciona una porina de MspA mutante que comprende un vestíbulo y una zona de constricción que definen un túnel, y al menos un primer monómero de MspA mutante que comprende una mutación en la posición 93 y una mutación en la posición 90, posición 91, o en ambas posiciones 90 y 91. Esta porina de MspA mutante, y cualquier otra porina de Msp mutante o porina de MspA descrita en el presente documento, pueden emplearse con cualquier realización descrita en el presente documento. La porina de MspA mutante puede comprender una mutación en las posiciones 93 y 90. La porina de MspA mutante puede comprender una mutación en las posiciones 93 y 91. La porina de MspA mutante puede comprender una mutación en las posiciones 93, 91, y 90. La porina de MspA mutante puede comprender cualquier otra mutación descrita en el presente documento. La MspA mutante de la invención comprende una mutación en las posiciones 90, 91 y 93, en la que la mutación en la posición 90 es D90N, la mutación en la posición 91 es D91N, y la mutación en la posición 93 es D93N.

[0056] El diámetro de la zona de constricción de la porina de MspA mutante puede ser menor que el diámetro de la zona de constricción de una correspondiente porina de MspA de tipo salvaje. La porina de MspA puede tener una mutación en el vestíbulo o en la zona de constricción que permite que un analito se traslade, por electroforesis o de otra manera, a través del túnel de mutante con una velocidad de traslado o una velocidad de traslado promedio que es menor que la velocidad de traslado o la velocidad de traslado promedio a la que el analito se traslada a través del túnel de una porina de Msp de tipo salvaje. La porina de MspA puede tener una mutación en el vestíbulo o en la zona de constricción que permite que un analito se traslade, por ejemplo, por electroforesis, a través del túnel con una velocidad de traslado promedio de menos de 0,5 nm/μs o menos de 0,05 nm/μs. El analito puede seleccionarse del grupo que consiste en un nucleótido, un ácido nucleico, un aminoácido, un péptido, una proteína, un polímero, un fármaco, un ion, un agente de guerra biológica, un contaminante, un objeto nanoscópico, o una combinación o grupo de los mismos. Opcionalmente, el analito se define además como un ácido nucleico. El ácido nucleico puede trasladarse, por electroforesis o de otra manera, a través del túnel con una velocidad de traslado promedio de menos de 1 nucleótido/μs, o menos de 0,1 nucleótido/μs. Un ácido nucleico se puede definir además como ssADN, dsADN,

ARN, o una combinación de los mismos.

[0057] Un analito en cualquier realización en el presente documento puede comprender además una miroesfera magnética. Una miroesfera magnética puede definirse además como una miroesfera magnética recubierta con estreptavidina. Un analito puede comprender además una microesfera óptica. Cualquier analito descrito en el presente documento puede ser un ión o puede ser neutro. Un analito puede comprender biotina.

[0058] Cualquier porina de Msp descrita en el presente documento, tal como una porina de MspA mutante, puede comprender 2-15 monómeros de Msp que son los mismos o diferentes. Opcionalmente, una porina de Msp, tal como una porina de MspA mutante, comprende 7-9 monómeros de Msp que son los mismos o diferentes. Opcionalmente, se selecciona al menos un segundo monómero del grupo que consiste en un monómero de MspA de tipo salvaje, un segundo monómero de MspA mutante, un monómero parálogo u homólogo de MspA de tipo salvaje, y un monómero parálogo u homólogo de MspA mutante, en la que el segundo monómero de MspA mutante puede ser el mismo o diferente que el primer monómero de MspA mutante. Opcionalmente, el segundo monómero es un monómero parálogo u homólogo de MspA de tipo salvaje. Un monómero parálogo u homólogo de MspA de tipo salvaje puede ser un monómero de MspB de tipo salvaje. Un monómero de MspA puede comprender una o más mutaciones en cualquiera de las siguientes posiciones de aminoácidos: 88, 105, 108, 118, 134, o 139. Un monómero de MspA puede comprender una o más de las siguientes mutaciones: L88W, D90K/N/Q/R, D91N/Q, D93N, I105W, N108W, D118R, D134R o E139K. Un monómero de MspA puede comprender las siguientes mutaciones: D90N/D91N/D93N. Un monómero de MspA puede comprender las siguientes mutaciones: D90N/D91N/D93N/D118R/D134R/E139K. Un monómero de MspA puede comprender las siguientes mutaciones: D90Q/D91Q/D93N. Un monómero de MspA puede comprender las siguientes mutaciones: D90Q/D91Q/D93N/D118R/D134R/E139K. Un monómero de MspA puede comprender las siguientes mutaciones: D90 (K,R)/D91N/D93N. Un monómero de MspA puede comprender las siguientes mutaciones: (L88, I105)W/D91Q/D93N. Un monómero de MspA puede comprender las siguientes mutaciones: I105W/N108W. Además, un monómero de MspA puede comprender cualquier otra mutación descrita en el presente documento.

[0059] En cualquier realización en el presente documento, una porina de Msp mutante, tal como una porina de MspA mutante o un parálogo u homólogo de MspA mutante, puede comprender al menos un aminoácido adicional con carga positiva en comparación con el vestíbulo o la zona de constricción de una porina de Msp de tipo salvaje, respectivamente; al menos un aminoácido adicional con carga negativa en comparación con el vestíbulo o la zona de restricción de una porina de MspA de tipo salvaje, respectivamente; al menos un aminoácido menos cargado positivamente en comparación con el vestíbulo o la zona de restricción de una porina de MspA de tipo salvaje, respectivamente; o al menos un aminoácido menos cargado negativamente en comparación con el vestíbulo o la zona de la constricción de una porina de MspA de tipo salvaje, respectivamente.

[0060] Opcionalmente, cada aminoácido cargado positivamente en el vestíbulo y la zona de constricción de una porina de Msp de tipo salvaje está sustituido por un aminoácido cargado negativamente, y cada aminoácido cargado negativamente es el mismo o diferente; o cada aminoácido cargado negativamente en el vestíbulo y la zona de constricción de una porina de Msp de tipo salvaje está sustituida por un aminoácido cargado positivamente, y cada aminoácido cargado positivamente es el mismo o diferente.

[0061] Opcionalmente, el vestíbulo o la zona de la constricción de una porina de Msp mutante comprende un mayor número de residuos cargados positivamente que el del vestíbulo o la zona de constricción de una porina de Msp de tipo salvaje, respectivamente; o el vestíbulo o la zona de constricción comprende un mayor número de residuos cargados negativamente que el del vestíbulo o la zona de constricción de una porina de Msp de tipo salvaje, respectivamente; o al menos un aminoácido cargado positivamente en el vestíbulo o la zona de restricción de una porina de Msp de tipo salvaje, tal como porina de MspA de tipo salvaje o porina paróloga u homóloga de MspA de tipo salvaje, se elimina o se sustituye por un aminoácido con carga negativa; o al menos un aminoácido con carga negativa en el vestíbulo o la zona de constricción de una porina de Msp de tipo salvaje se elimina o se sustituye por un aminoácido cargado positivamente.

[0062] Al menos un aminoácido en el vestíbulo o en la zona de restricción de una porina de Msp de tipo salvaje, tales como porina de MspA de tipo salvaje o porina paróloga u homóloga de MspA de tipo salvaje, puede ser sustituido por un aminoácido que tiene una cadena lateral estéricamente más grande; un aminoácido que tiene una cadena lateral estéricamente más pequeña; un aminoácido que tiene una cadena lateral más polar; un aminoácido que tiene una cadena lateral menos polar; o un aminoácido que tiene una cadena lateral más hidrófobo; un aminoácido que tiene una cadena lateral menos hidrófoba.

[0063] En cualquier realización en el presente documento, al menos un aminoácido en el vestíbulo o en la zona de constricción de una porina de Msp mutante puede comprender un aminoácido no natural o un aminoácido modificado químicamente.

[0064] Cualquier porina de Msp descrita en el presente documento puede comprender una o más deleciones, adiciones o sustituciones en bucles periplásmicos.

[0065] Tal como se describe en el presente documento, cualquier porina de Msp, como una porina de MspA mutante, puede comprender además un motor molecular. Cualquier motor molecular descrito en el presente documento puede ser capaz de mover un analito en o a través del túnel con una velocidad de traslado o una velocidad de traslado promedio que es menor que la velocidad de traslado o la velocidad de traslado promedio a la que el analito se traslada en o a través del túnel en ausencia del motor molecular. En cualquier realización en el presente documento, el motor molecular puede ser una enzima, tal como una polimerasa, una exonucleasa, o un fragmento de Klenow.

[0066] También se proporcionan procedimientos para producir porinas de Msp descritas en el presente documento. En consecuencia, se proporciona un procedimiento de producción de una porina de MspA mutante que comprende al menos un monómero de MspA mutante, comprendiendo el procedimiento modificar un monómero de MspA de tipo salvaje en la posición 93 y en la posición 90, posición 91, o en ambas posiciones 90 y 91. El procedimiento puede comprender modificar un monómero de MspA de tipo salvaje en las posiciones 93 y 90. El procedimiento puede comprender modificar un monómero de MspA de tipo salvaje en las posiciones 93, 91, y 90. El procedimiento puede comprender además o alternativamente modificar un monómero de MspA de tipo salvaje en una o más de las siguientes posiciones de aminoácidos: 88, 105, 108, 118, 134, o 139, o realizar cualquier otra modificación descrita en el presente documento. Un porina de MspA mutante producida por los procedimientos descritos en el presente documento puede comprender cualquier mutación o propiedad de porina descrita en el presente documento. Por ejemplo, una MspA mutante puede comprender una zona de constricción neutra. Una porina de MspA mutante puede comprender además al menos un monómero de Msp, tal como un monómero de MspA de tipo salvaje, un monómero de MspA mutante, un parólogo u homólogo de MspA de tipo salvaje, o un segundo monómero parólogo u homólogo de MspA mutante. La porina de MspA mutante puede tener una conductancia a través del túnel que es más alta, por ejemplo, dos veces mayor, que la conductancia a través del túnel de su correspondiente porina de MspA de tipo salvaje.

[0067] Cualquier porina de Msp mutante descrita en el presente documento, tal como una porina de MspA mutante o porina paráloga u homóloga de MspA mutante, puede comprender uno o más monómeros de MspB mutante, MspC mutante, o MspD mutante, o una combinación de los mismos.

[0068] También se proporciona un procedimiento de fabricación de una porina de MspA mutante que tiene un vestíbulo y una zona de constricción que definen un túnel, que comprende la supresión, adición o sustitución de cualquier aminoácido en el vestíbulo o en la zona de constricción de un monómero parólogo u homólogo de MspA de tipo salvaje, de manera que la porina de MspA mutante resultante es capaz de trasladar un analito a través del túnel después de la aplicación de un campo eléctrico. La porina de MspA mutante puede ser de cualquier tipo descrito en el presente documento.

[0069] También se proporcionan secuencias de ácido nucleico que codifican porinas de Msp descritas en el presente documento. Por ejemplo, se proporciona una secuencia de ácido nucleico que codifica una porina de MspA mutante o un parólogo u homólogo de MspA mutante. También se contemplan vectores que comprenden secuencias de ácido nucleico descritas en el presente documento, tal como un vector que comprende una secuencia de ácido nucleico que codifica una porina de MspA mutante o un homólogo o parólogo de MspA mutante. Cualquier vector descrito en el presente documento puede comprender además una secuencia de promotor. Cualquier vector descrito en el presente documento puede comprender además un promotor constitutivo. Un promotor constitutivo puede comprender un promotor p_{smc} . Un promotor puede comprender un promotor inducible. Un promotor inducible puede comprender un promotor inducible por acetamida.

[0070] También se proporcionan células cultivadas transfectadas con cualquier vector descrito en el presente documento, o progenie de las mismas en las que la célula es capaz de expresar una porina de Msp, tal como una porina de MspA mutante o un parólogo u homólogo de MspA mutante.

[0071] También se proporciona una cepa de *Mycobacterium smegmatis* que comprende cualquier vector descrito en el presente documento. También se contempla una cepa de *Mycobacterium smegmatis* libre de porinas endógenas, y puede comprender, además, cualquier vector descrito en el presente documento. Por "libre" se entiende que una porina endógena no se puede detectar en una inmunotransferencia utilizando un antisuero específico de Msp apropiado, o que comprende menos del 1% de porinas endógenas.

[0072] También se proporciona un vector que comprende una secuencia de ácido nucleico que codifica un monómero de Msp de tipo salvaje, en el que la secuencia de ácido nucleico está operativamente controlada por un promotor inducible. El vector puede ser un vector de integración. También se proporciona una célula cultivada transfectada con este vector, o progenie de la misma, en la que la célula es capaz de expresar una porina de Msp de tipo salvaje. También se contempla una cepa de *Mycobacterium smegmatis* que comprende este vector.

[0073] También se proporcionan secuencias de ácidos nucleicos que codifican una porina parcial o completa de Msp de una cadena descrita en el presente documento. La secuencia de ácido nucleico puede comprender, por ejemplo: (a) una primera y segunda secuencia de nucleótidos, en el que la primera secuencia de nucleótidos codifica una

primera secuencia de monómero de Msp y la segunda secuencia de nucleótidos codifica una segunda secuencia de monómero de Msp; y (b) una tercera secuencia de nucleótidos que codifica una secuencia de aminoácidos enlazadora. La primera y segunda secuencias de monómero de Msp pueden seleccionarse independientemente del grupo que consiste en un monómero de MspA de tipo salvaje, un monómero de MspA mutante, un monómero parólogo u homólogo de MspA de tipo salvaje, y un monómero parólogo u homólogo de MspA mutante. La primera secuencia de monómero de Msp puede comprender un monómero de MspA de tipo salvaje o un mutante del mismo. Opcionalmente, la primera secuencia de monómero de Msp comprende un monómero de MspA mutante. La primera secuencia de monómero de Msp puede comprender una o más de las mutaciones seleccionadas del grupo que consiste en una sustitución de A a P en el aminoácido 138, una sustitución de E a A o K en el aminoácido 139, una sustitución de D a K o R o Q en el aminoácido 90; una sustitución de D a N o Q en el aminoácido 91, una sustitución de D a N en el aminoácido 93, una sustitución de L a W en el aminoácido 88, una sustitución de I a W en el aminoácido 105, una sustitución de N a W en el aminoácido 108, una sustitución de D a R en el aminoácido 118, y una sustitución de D a R en el aminoácido 134. De hecho, cualquier monómero de Msp descrito en el presente documento puede comprender cualquiera de estas sustituciones.

[0074] Opcionalmente, el monómero de MspA mutante comprende una sustitución de A a P en el aminoácido 138, una sustitución de E a A en el aminoácido 139, o una combinación de los mismos; una sustitución de D a K o R en el aminoácido 90, una sustitución de D a N en el aminoácido 91, una sustitución de D a N en el aminoácido 93, o cualquier combinación de los mismos; una sustitución de D a Q en el aminoácido 90, una sustitución de D a Q en el aminoácido 91, una sustitución de D a N en el aminoácido 93, o cualquier combinación de los mismos; una sustitución de L a W en el aminoácido 88, una sustitución de I a W en el aminoácido 105, una sustitución de D a Q en el aminoácido 91, una sustitución de D a N en el aminoácido 93, o cualquier combinación de los mismos; una sustitución de I a W en el aminoácido 105, una sustitución de N a W en el aminoácido 108, o una combinación de los mismos; o una sustitución de D a R en el aminoácido 118, una sustitución de E a K en el aminoácido 139, una sustitución de D a R en el aminoácido 134, o cualquier combinación de los mismos.

[0075] Cualquier porina de Msp puede comprender una primera, segunda, o más secuencias de monómero de Msp que comprende un parólogo de MspA de tipo salvaje o mutante del mismo, en el que el parólogo o mutante de la misma es un monómero de MspB de tipo salvaje o un mutante del mismo. Una o más secuencias de monómero de Msp pueden comprender SEQ ID NO: 1, SEQ ID NO: 2, o una combinación de los mismos. Opcionalmente, la segunda secuencia de monómero de Msp comprende un monómero de MspB mutante. Opcionalmente, la primera secuencia de monómero de Msp comprende un monómero de MspA de tipo salvaje o un mutante del mismo y la segunda secuencia de monómero de Msp comprende un monómero de MspB de tipo salvaje o un mutante del mismo. Opcionalmente, la primera secuencia de monómero de Msp comprende la SEQ ID NO: 1 y la segunda secuencia de monómero de Msp comprende la SEQ ID NO: 2.

[0076] Las secuencias de aminoácidos enlazadoras se describen en el presente documento. En cualquier realización en el presente documento, una secuencia de aminoácidos enlazadora puede comprender, por ejemplo, de 10 a 20 aminoácidos. Por ejemplo, una secuencia de aminoácidos enlazadora comprende 15 aminoácidos. Opcionalmente, la secuencia de aminoácidos enlazadora comprende una secuencia de péptido (GGGGS)₃ (SEQ ID NO: 3).

[0077] Se contemplan los polipéptidos codificados por cualquier secuencia de ácido nucleico descrita en el presente documento.

[0078] También se proporciona una secuencia de ácido nucleico que codifica una porina parcial o completa de Msp de una cadena, en la que la secuencia de ácido nucleico comprende: (a) una primera, segunda, tercera, cuarta, quinta, sexta, séptima, y octava secuencia de nucleótidos o cualquier subconjunto de las mismas, en la que la primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de nucleótidos codifica una primera, segunda, tercera, cuarta, quinta, sexta, séptima, y octava secuencia de monómero de Msp, respectivamente; y (b) una secuencia de nucleótidos que codifica una novena secuencia de aminoácidos enlazadora. La primera y segunda secuencias de monómero de Msp pueden seleccionarse independientemente del grupo que consiste en un monómero de Msp de tipo salvaje, un monómero de Msp mutante, un monómero parólogo u homólogo de MspA de tipo salvaje, y un monómero parólogo u homólogo de MspA mutante. Cada monómero de Msp puede comprender un monómero de MspA de tipo salvaje o un mutante del mismo. Opcionalmente, al menos un monómero de Msp comprende un monómero de MspA de tipo salvaje o un mutante del mismo. Opcionalmente, al menos un monómero de Msp comprende un monómero de MspA mutante. La secuencia de monómero de Msp mutante puede comprender cualquier mutación descrita en el presente documento. Por ejemplo, una o más de las mutaciones se pueden seleccionar del grupo que consiste en una sustitución de A a P en el aminoácido 138, una sustitución de E a A o K en el aminoácido 139, una sustitución de D a K o R o Q en el aminoácido 90; una sustitución de D a N o Q en el aminoácido 91, una sustitución de D a N en el aminoácido 93, una sustitución de L a W en el aminoácido 88, una sustitución de I a W en el aminoácido 105, una sustitución de N a W en el aminoácido 108, una sustitución de D a R en el aminoácido 118, y una sustitución de D a R en el aminoácido 134. Cada secuencia de monómero de Msp puede comprender SEQ ID NO: 1. Opcionalmente, al menos una secuencia de monómero de Msp comprende un parólogo de MspA de tipo salvaje o mutante del mismo, en el que el parólogo de MspA o mutante del mismo es un monómero de MspB de

tipo salvaje o un mutante del mismo. Opcionalmente, al menos una secuencia de monómero de Msp comprende la SEQ ID NO: 2. Opcionalmente, al menos una secuencia de monómero de Msp comprende un monómero de MspB mutante. Opcionalmente, al menos una secuencia de monómero de Msp comprende un monómero de MspA de tipo salvaje o un mutante del mismo y al menos una secuencia de monómero de Msp comprende un monómero de MspB de tipo salvaje MSPB o un mutante del mismo. Opcionalmente, al menos una secuencia de monómero de Msp comprende la SEQ ID NO: 1 y al menos una secuencia de monómero de Msp comprende la SEQ ID NO: 2. También se proporciona un polipéptido codificado por cualquiera de las secuencias de ácidos nucleicos anteriores. También se proporciona un vector que comprende cualquiera de las secuencias de ácidos nucleicos anteriores. El vector puede comprender además una secuencia de promotor. El promotor puede comprender un promotor constitutivo. El promotor constitutivo puede comprender promotor p_{smyc} . El promotor puede comprender un promotor inducible. El promotor inducible puede comprender un promotor inducible por acetamida.

[0079] También se proporciona una cepa bacteriana mutante capaz de la expresión de Msp inducible, comprendiendo la cepa bacteriana: (a) una delección de MspA de tipo salvaje; (b) una delección de MspC de tipo salvaje; (c) una delección de MspD de tipo salvaje; y (d) un vector que comprende un promotor inducible unido operativamente a una secuencia de ácido nucleico de monómero de Msp. La cepa bacteriana puede comprender, además, la cepa ML16 de *M. smegmatis*. El ácido nucleico de Msp puede codificar un monómero de MspA de tipo salvaje o un monómero parálogo u homólogo de MspA de tipo salvaje. El ácido nucleico de Msp puede codificar un monómero de Msp seleccionado del grupo que consiste en un monómero de MspA de tipo salvaje, un monómero de MspC de tipo salvaje, y un monómero de MspD de tipo salvaje. Opcionalmente, el ácido nucleico de Msp codifica el monómero de MspA de tipo salvaje. El promotor inducible puede comprender un promotor inducible por acetamida. La cepa bacteriana puede comprender además una delección de una MspB de tipo salvaje. La cepa bacteriana puede comprender además un vector tal como se describe en el presente documento, tal como un vector que comprende un promotor constitutivo unido operativamente a una secuencia de ácido nucleico que codifica una porina o monómero de Msp. La Msp puede ser una porina o monómero de MspA de tipo salvaje o una porina o monómero parálogo u homólogo de MspA de tipo salvaje. La porina o monómero de Msp puede seleccionarse del grupo que consiste en una porina o monómero de MspA de tipo salvaje, porina o monómero de MspB de tipo salvaje, porina o monómero de MspC de tipo salvaje, y porina o monómero de MspD de tipo salvaje. Opcionalmente, la porina o monómero de Msp es una porina o monómero de MspA de tipo salvaje.

[0080] La cepa bacteriana puede comprender además un vector que comprende un ácido nucleico que codifica una porina parcial o completa de Msp de una cadena, en el que el ácido nucleico comprende: (a) una primera y segunda secuencia de nucleótidos, en el que la primera secuencia de nucleótidos codifica un primer monómero de Msp y la segunda secuencia de nucleótidos codifica una segunda secuencia de monómero Msp; y (b) una tercera secuencia de nucleótidos que codifica una secuencia de aminoácidos enlazadora. La cepa bacteriana puede comprender además un vector que comprende un ácido nucleico que codifica porina parcial o completa de Msp de una cadena, en el que el ácido nucleico comprende: (a) una primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de nucleótidos o cualquier subconjunto de la misma, en la que la primera, segunda, tercera, cuarta, quinta, sexta, séptima, y octava secuencia de nucleótidos codifican una primera, segunda, tercera, cuarta, quinta, sexta, séptima, y octava secuencia de monómero de Msp, respectivamente; y (b) una secuencia de nucleótidos que codifica una novena secuencia de aminoácidos enlazadora.

[0081] También se proporciona un procedimiento de producción de una porina parcial o completa de Msp de una cadena, comprendiendo el procedimiento: (a) transformar una cepa bacteriana tal como se describe en el presente documento con un vector que comprende una secuencia de ácido nucleico capaz de codificar una porina parcial o completa de Msp de una cadena; y (b) purificar la porina parcial o completa de Msp de una cadena de las bacterias. El vector puede comprender una secuencia de ácido nucleico que codifica una porina parcial o completa de Msp de una cadena, en el que la secuencia de ácido nucleico comprende: (a) una primera y segunda secuencia de nucleótidos, en la que la primera secuencia de nucleótidos codifica una primera secuencia de monómero de Msp y la segunda secuencia de nucleótidos codifica una segunda secuencia de monómero de Msp; y (b) una tercera secuencia de nucleótidos que codifica una secuencia de aminoácidos enlazadora. El vector puede comprender una secuencia de ácido nucleico que codifica una porina parcial o completa de Msp de una cadena, en el que la secuencia de ácido nucleico comprende: (a) una primera, segunda, tercera, cuarta, quinta, sexta, séptima, y octava secuencia de nucleótidos o cualquier subconjunto de las mismas, en la que la primera, segunda, tercera, cuarta, quinta, sexta, séptima y octava secuencia de nucleótidos codifican una primera, segunda, tercera, cuarta, quinta, sexta, séptima, y octava secuencia de monómero de Msp, respectivamente; y (b) una novena secuencia de nucleótidos que codifica un enlazador de aminoácidos. Las secuencias de monómero de Msp pueden seleccionarse independientemente del grupo que consiste en un monómero de MspA de tipo salvaje, un monómero de MspA mutante, un monómero parálogo u homólogo de MspA de tipo salvaje, y un monómero parálogo u homólogo de MspA mutante. Por ejemplo, las secuencias de monómero de Msp son monómeros de MspA de tipo salvaje.

[0082] Una "porina de *Mycobacterium smegmatis* (Msp)" o "porina de Msp" se refiere a un complejo multimérico que comprende dos o más monómeros de Msp. Un monómero de Msp está codificado por un gen en *Mycobacterium smegmatis*. *Mycobacterium smegmatis* tiene cuatro genes de Msp identificados, indicados como MspA, MspB, MspC y MspD. Un porina de Msp puede, por ejemplo, estar constituido por monómeros de MspA de tipo salvaje, monómeros de MspA mutantes, monómeros parálogos u homólogos de MspA de tipo salvaje, o monómeros

parálogos u homólogos de MspA mutante. Opcionalmente, una porina de Msp es una de Msp de una cadena o es un multímero de varias porinas de Msp de una cadena. Una porina de Msp de una cadena puede, por ejemplo comprender un multímero formado por dos o más monómeros de Msp (por ejemplo, ocho monómeros) conectados por uno o más péptidos de aminoácidos enlazadores. Una porina parcial de Msp de una cadena se refiere a un complejo multimérico de una cadena que debe dimerizar, trimerizar, o similar para formar una porina. Una porina completa de Msp de una cadena se refiere a un complejo multimérico de una cadena que forma una porina sin la necesidad de dimerizar, trimerizar o similares para formar una porina.

[0083] La porina de Msp de cualquier realización en el presente documento puede ser cualquier porina de Msp descrita en el presente documento, tal como una porina de MspA de tipo salvaje, una porina de MspA mutante, una porina paróloga u homóloga de MspA de tipo salvaje, o una porina paróloga u homóloga de MspA mutante. La porina de Msp puede ser codificada por una secuencia de ácido nucleico que codifica una porina de Msp de una cadena. Cualquier porina de Msp del presente documento puede comprender cualquier monómero de Msp descrito en el presente documento, tal como un monómero de Msp mutante.

[0084] Los nutrientes pasan a través de porinas de tipo salvaje en micobacterias. Las porinas de MspA de tipo salvaje, porinas de MspB de tipo salvaje, porinas de MspC de tipo salvaje, y porinas de MspD de tipo salvaje son ejemplos de porinas formadoras de túnel de tipo salvaje. Una porina de Msp puede definirse adicionalmente como cualquier porina de Msp descrita en el presente documento, incluyendo porinas parálogas, homólogas, mutantes y de una sola cadena.

[0085] Una "porina de MspA mutante" es un complejo multimérico que tiene al menos o como máximo 70, 75, 80, 85, 90, 95, 98, o 99 por ciento o más de identidad, o cualquier intervalo derivable en el mismo, pero menos del 100%, a su correspondiente porina de MspA de tipo salvaje y retiene la capacidad de formación de túnel. Una porina de MspA mutante puede ser una proteína recombinante. Opcionalmente, una porina de MspA mutante es una que tiene una mutación en la zona de constricción o en el vestíbulo de una porina de MspA de tipo salvaje. Opcionalmente, puede aparecer una mutación en el borde o en el exterior de los bucles periplásmicos de una porina de MspA de tipo salvaje. Una porina de MspA mutante se puede emplear en cualquier realización descrita en el presente documento.

[0086] Los ejemplos parálogos y homólogos de MspA de tipo salvaje se proporcionan en la Tabla 1. Se proporcionan parálogos de MspA de tipo salvaje, que incluyen la MspB de tipo salvaje, MspC de tipo salvaje y MspD de tipo salvaje. Un "parólogo", tal como se define en el presente documento, es un gen de la misma especie bacteriana que tiene estructura y función similares. Un "homólogo", tal como se define en el presente documento, es un gen de otra especie bacteriana que tiene una estructura similar y origen evolutivo. A modo de ejemplo, se proporcionan homólogos de MspA de tipo salvaje, que incluyen MppA, PorM1, PorM2, PorM1 y Mmcs4296.

[0087] Una "porina paróloga u homóloga de MspA mutante" es un complejo multimérico que tiene al menos o como máximo 70, 75, 80, 85, 90, 95, 98, o 99 por ciento o más de identidad, o cualquier intervalo derivable en el mismo, pero menos de 100%, a su correspondiente porina paróloga u homóloga de MspA de tipo salvaje y retiene la capacidad de formación de túnel. Una porina paróloga u homóloga de MspA mutante puede ser una proteína recombinante. Opcionalmente, una porina paróloga u homóloga de MspA mutante es una que tiene una mutación en la zona de constricción o en el vestíbulo de la porina paróloga u homóloga de MspA de tipo salvaje. Opcionalmente, puede aparecer una mutación en el borde o en el exterior de los bucles periplásmicos de una porina paróloga u homóloga de MspA de tipo salvaje. Cualquier porina paróloga u homóloga de MspA mutante se puede emplear en cualquier realización descrita en el presente documento y puede comprender mutación descrita en el presente documento.

[0088] Una porina de Msp puede comprender dos o más monómeros de Msp. Un "monómero de Msp" es un monómero de proteína que es un monómero de MspA de tipo salvaje, un monómero de MspA mutante, un monómero parólogo u homólogo de MspA de tipo salvaje, o un monómero parólogo u homólogo de MspA mutante, y retiene la capacidad de formación de túnel cuando está asociado con uno o más de otros monómeros de Msp. Cualquier porina de Msp descrita en el presente documento puede comprender una o más de cualquier monómero de Msp tal como se describe en el presente documento. Cualquier porina de Msp puede comprender, por ejemplo, 2-15 monómeros de Msp, en la que cada monómero puede ser el mismo o diferente.

[0089] Un "monómero de MspA mutante" se refiere a un monómero de Msp que tiene al menos o como máximo 70, 75, 80, 85, 90, 95, 98, o 99 por ciento o más de identidad, o cualquier intervalo derivable en el mismo, pero menos de 100%, a un monómero de MspA de tipo salvaje, y retiene la capacidad de formación de túnel cuando está asociado con uno o más de otros monómeros de Msp. Opcionalmente, un monómero de MspA mutante se define además como que comprende una mutación en la parte de la secuencia que contribuye a la formación del vestíbulo o la zona de la constricción de una porina formadora de túnel totalmente formada. El monómero de Msp mutante puede ser una proteína recombinante, por ejemplo. Un monómero de MspA mutante puede comprender cualquier mutación descrita en el presente documento.

[0090] Un "monómero parólogo u homólogo de MspA mutante" se refiere a un monómero parólogo u homólogo de

MspA que tiene al menos o como máximo 70, 75, 80, 85, 90, 95, 98, o 99 por ciento o más de identidad, o cualquier intervalo derivable en el mismo, pero menos del 100%, a un monómero parólogo u homólogo de MspA de tipo salvaje, y retiene la capacidad de formación de túnel. Opcionalmente, un monómero parólogo u homólogo de MspA mutante se define además como que comprende una mutación en la parte de la secuencia que contribuye a la formación del vestíbulo y/o la zona de la constricción de una porina formadora de túnel totalmente formada. El monómero parólogo u homólogo de MspA mutante puede ser una proteína recombinante, por ejemplo. Cualquier monómero parólogo u homólogo de MspA mutante se puede emplear opcionalmente en cualquier realización en el presente documento.

5 [0091] Una porina de Msp puede expresarse como una combinación de dos o más monómeros de MspA de tipo salvaje, monómeros de MspA mutantes, monómeros parálogos u homólogos de MspA de tipo salvaje, o monómeros parálogos u homólogos de MspA. Por tanto, una porina de Msp puede ser o comprender un dímero, un trímero, un tetramero, un pentámero, un hexámero, un septámero, un octámero, un nonámero, etc. Por ejemplo, una porina de Msp puede comprender una combinación de monómeros de MspA de tipo salvaje y monómeros de MspB de tipo salvaje. Una porina de Msp puede comprender 1-15 monómeros, donde cada monómero es igual o diferente. De hecho, cualquier porina de Msp descrita en el presente documento puede comprender al menos o como máximo 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, o 15 monómeros, o cualquier intervalo derivable en el mismo, donde cada monómero es igual o diferente. Por ejemplo, una porina de Msp puede comprender uno o más monómeros de MspA mutante que son los mismos o diferentes. Como otro ejemplo, una porina de Msp puede comprender al menos un monómero de MspA mutante y al menos un monómero parólogo u homólogo de MspA.

10 [0092] Tal como se define anteriormente, una porina de Msp de una cadena comprende dos o más monómeros de Msp conectados por uno o más péptidos de aminoácido enlazadores. Una porina de Msp de una cadena que comprende dos monómeros de Msp, en el que los monómeros de Msp están unidos por una secuencia de aminoácidos enlazadora, puede denominarse como un dímero de porinas de Msp de una cadena. Una porina de Msp de una cadena que comprende ocho monómeros de Msp, en la que los monómeros de Msp están unidos por una secuencia de aminoácidos enlazadora, puede denominarse como un octámero de porinas de Msp de una cadena. Una porina de Msp de una cadena puede comprender 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, o más monómeros de Msp, o cualquier intervalo derivable en el mismo, unidos por secuencias de aminoácidos enlazadoras. Opcionalmente, una porina de Msp de una cadena puede comprender, por ejemplo, dos o más dímeros de porina de Msp de una cadena, dos o más trímeros de porina de Msp de una cadena, dos o más cuadrímeros de porina de Msp de una cadena, dos o más pentímeros de porina de Msp de una cadena, uno o más hexímeros de porina de Msp de una cadena, uno o más septímeros de porina de Msp de una cadena, uno o más octámeros de porina de Msp de una cadena, o combinaciones de los mismos. Por ejemplo, una porina de Msp de una cadena puede comprender un dímero de porina de Msp de una cadena y dos trímeros de porina de Msp de una cadena. A modo de otro ejemplo, una porina de Msp de una cadena puede comprender un cuadrímero de porina de Msp de una cadena y dos dímeros de porina de Msp de una cadena.

15 [0093] Una porina de Msp de una cadena de tipo salvaje se compone de monómeros de Msp de tipo salvaje. Opcionalmente, están presentes una o más mutaciones en una porina de Msp de una cadena en el vestíbulo o en la zona de la constricción de una porina de Msp de una cadena. La porina de Msp de una cadena, por ejemplo, tiene al menos una mutación en la secuencia de aminoácidos para el bucle periplásmico, vestíbulo, o zona de constricción (por ejemplo, delección, sustitución o adición) en comparación con una Msp de una cadena de tipo salvaje. Un multímero de las cadenas individuales también puede formar una porina, donde cada cadena individual incluye dos, tres, cuatro, cinco, seis, siete, o más monómeros de Msp.

20 [0094] En el presente documento se proporcionan secuencias de ácidos nucleicos que codifican secuencias de monómero de Msp y mutantes de las mismas. Para las secuencias de monómero de MspA mutante indicadas a continuación, la secuencia de MspA de referencia es la secuencia del monómero de MspA de tipo salvaje madura (SEQ ID NO: 1). Cada secuencia de nucleótidos en las secuencias de ácido nucleico proporcionadas en el presente documento puede comprender, por ejemplo, una secuencia de monómero de MspA mutante. Los ejemplos no limitantes de secuencias de MspA mutantes se proporcionan en la Tabla 7. Opcionalmente, la MspA mutante comprende una sustitución de A a P en el aminoácido 138, una sustitución de E a A en el aminoácido 139, o una combinación de los mismos. Opcionalmente, la MspA mutante comprende una sustitución de D a K o R en el aminoácido 90, una sustitución de D a N en el aminoácido 91, una sustitución de D a N en el aminoácido 93, o cualquier combinación de los mismos. Opcionalmente, la MspA mutante comprende una sustitución de D a Q en el aminoácido 90, una sustitución de D a Q en el aminoácido 91, una sustitución de D a N en el aminoácido 93, o cualquier combinación de los mismos. Opcionalmente, la MspA mutante comprende una sustitución de L a W en el aminoácido 88, una sustitución de I a W en el aminoácido 105, una sustitución de D a Q en el aminoácido 91, una sustitución de D a N en el aminoácido 93, o cualquier combinación de los mismos. Opcionalmente, la MspA mutante comprende una sustitución de I a W en el aminoácido 105, una sustitución de N a W en el aminoácido 108, o una combinación de los mismos. Opcionalmente, la MspA mutante comprende una sustitución de D a R en el aminoácido 118, una sustitución de E a K en el aminoácido 139, una sustitución de D a R en el aminoácido 134, o cualquier combinación de los mismos. Para las secuencias de monómero de MspB mutante indicadas a continuación, la secuencia de MspB de referencia es la secuencia del monómero de MspB de tipo salvaje madura (SEQ ID NO: 2). Opcionalmente, la MspB mutante comprende una sustitución de D a K o R en el aminoácido 90, una sustitución de D

a N en el aminoácido 91, una sustitución de D a N en el aminoácido 93, o cualquier combinación de los mismos.

[0095] Las secuencias de monómeros de Msp tipo salvaje descritas en el presente documento se dan a conocer en el GenBank, que se encuentran en la World Wide Web en pubmed.gov.

[0096] Por ejemplo, las secuencias de nucleótidos y de aminoácidos de un monómero de MspA de tipo salvaje se puede encontrar en GenBank Accession Nos. AJ001442 y CAB56052, respectivamente. Las secuencias de nucleótidos y de aminoácidos de un monómero de MspB de tipo salvaje se pueden encontrar, por ejemplo, en GenBank Accession Nos. NC_008596.1 (desde el nucleótido 600.086 a 600.730) y YP_884932.1, respectivamente. Las secuencias de nucleótidos y de aminoácidos de un monómero de MspC de tipo salvaje se pueden encontrar, por ejemplo, en GenBank Accession Nos. AJ299735 y CAC82509, respectivamente. Las secuencias de nucleótidos y de aminoácidos de un monómero de MspD de tipo salvaje se pueden encontrar, por ejemplo, en GenBank Accession Nos. AJ300774 y CAC83628, respectivamente. De este modo se proporcionan las secuencias de nucleótidos de los monómeros MspA, MspB, MspC y MspD que comprenden una secuencia de nucleótidos de al menos aproximadamente 70, 75, 80, 85, 90, 95, 98, 99 por ciento o más, o cualquier intervalo derivable en el mismo, idéntica a la secuencia de nucleótidos de los números de acceso de GenBank de nucleótidos mencionados anteriormente. También se proporcionan secuencias de aminoácidos de los monómeros MspA, MspB, MspC y MspD (FIGURA 18) que comprenden una secuencia de aminoácidos al menos aproximadamente 70, 75, 80, 85, 90, 95, 98, 99 por ciento o más, o cualquier intervalo derivable en el mismo, idéntica a las secuencias de los números de acceso de GenBank de aminoácidos mencionados anteriormente.

[0097] También se proporcionan secuencias de aminoácidos de monómeros parálogos y homólogos de MspA, que comprenden una secuencia de aminoácidos al menos aproximadamente 70, 75, 80, 85, 90, 95, 98, 99 por ciento o más, o cualquier intervalo derivable en el mismo, a un monómero parálogo u homólogo de MspA de tipo salvaje. Los monómeros parálogos y homólogos de MspA de tipo salvaje son conocidos en la técnica. La Tabla 1 proporciona una lista no limitante de dichos parálogos y homólogos:

Tabla 1. MspA de tipo salvaje y monómeros parálogos y homólogos de MspA de tipo salvaje

Proteína#	Organismo	Identidad/Similitud con MspA (%)	Longitud (aa)	Referencia
MspA/Msmeg0965	<i>M. smegmatis</i>	100/100	211	gb ABK74363.1 , (Stahl <i>et al.</i> , 2001)*
MspB/Msmeg0520	<i>M. smegmatis</i>	94/95	215	gb ABK73437.1 , (Stahl <i>et al.</i> , 2001)*
MspC/Msmeg5483	<i>M. smegmatis</i>	93/95	215	gb ABK74976.1 , (Stahl <i>et al.</i> , 2001)*
MspD/Msmeg6057	<i>M. smegmatis</i>	82/89	207	gb ABK72453.1 , (Stahl <i>et al.</i> , 2001)*
MppA	<i>M. phlei</i>	100/100	211	AJ812030, (Dorner <i>et al.</i> , 2004)**
PorM1	<i>M. fortuitum</i>	95/96	211	emb CAI54228.1
PorM2	<i>M. fortuitum</i>	91/93	215	emb CAL29811.1
PorM1	<i>M. peregrinum</i>	94/96	211	emb CAI54230.1
Mmcs4296	<i>Mycobacterium sp. MCS</i>	85/91	216	gb ABG10401.1
Mmcs4297	<i>Mycobacterium sp. MCS</i>	85/91	216	gb ABG10402.1
Mmcs3857	<i>Mycobacterium sp. MCS</i>	30/44	235	gb ABG09962.1
Mmcs4382	<i>Mycobacterium sp. MCS</i>	85/91	216	gb ABL93573.1
Mmcs4383	<i>Mycobacterium sp. MCS</i>	85/91	216	gb ABL93574.1
Mjls3843	<i>Mycobacterium sp. JLS</i>	26/40	235	gb ABN99619.1
Mjls3857	<i>Mycobacterium sp. JLS</i>	26/40	235	gb ABG09962.1

Mjls3931	<i>Mycobacterium sp. JLS</i>	26/40	235	gb ABL93123.1
Mjls4674	<i>Mycobacterium sp. JLS</i>	85/89	216	gb ABO00440.1
Mjls4675	<i>Mycobacterium sp. JLS</i>	83/89	216	gb ABO00441.1
Mjls4677	<i>Mycobacterium sp. JLS</i>	84/89	216	gb ABO00443.1
Map3123c	<i>M. avium paratuberculosis</i>	24/39	220	gb AAS05671.1
Mav3943	<i>M. avium</i>	24/39	227	gb ABK66660.1
Mvan1836	<i>M. vanbaalenii</i> PYR-1	82/88	209	gb ABM12657.1
Mvan4117	<i>M. vanbaalenii</i> PYR-1	32/43	239	gb ABM14894.1
Mvan4839	<i>M. vanbaalenii</i> PYR-1	83/88	209	gb ABM15612.1
Mvan4840	<i>M. vanbaalenii</i> PYR-1	83/89	209	gb ABM15613.1
Mvan5016	<i>M. vanbaalenii</i> PYR-1	30/41	238	gb ABM15788.1
Mvan5017	<i>M. vanbaalenii</i> PYR-1	25/35	227	gb ABM15789.1
Mvan5768	<i>M. vanbaalenii</i> PYR-1	21/32	216	gb ABM16533.1
MUL_2391	<i>M. ulcerans</i> Agy99	21/34	233	gb ABL04749.1
Mflv1734	<i>M. gilvum</i> PYR-GCK	21/32	225	gb ABP44214.1
Mflv1735	<i>M. gilvum</i> PYR-GCK	32/41	226	gb ABP44215.1
Mflv2295	<i>M. gilvum</i> PYR-GCK	26/40	250	gb ABP44773.1
Mflv1891	<i>M. gilvum</i> PYR-GCK	84/90	217	gb ABP44371.1
MCH4691c	<i>M. chelonae</i>	70/80	223	gb ACV04474.1
MCH4689c	<i>M. chelonae</i>	66/78	223	gb ACV04472.1
MCH4690c	<i>M. chelonae</i>	72/81	217	gb ACV04473.1
MAB1080	<i>M. abscessus</i>	69/79	223	emb CAM61170.1
MAB1081	<i>M. abscessus</i>	68/78	222	emb CAM61171.1
MAB2800	<i>M. abscessus</i>	27/44	246	emb CAM62879.1
RHA1 ro08561	<i>Rhodococcus jostii</i> RHA1	34/51	233	gb ABG99605.1
n.d.	<i>Rhodococcus opacus</i> B4	34/51	233	gbj BAH52196.11
RHA1 ro04074	<i>Rhodococcus sp.</i> RHA1	34/50	233	gb ABG95871.1
RHA1 ro03127	<i>Rhodococcus sp.</i> RHA1	34/50	233	gb ABG94930.1
n.d.	<i>Rhodococcus erythropolis</i> PR4	35/50	229	gbj BAH30938.1

Sólo se incluyeron las proteínas con similitudes significativas de aminoácidos con respecto a la longitud completa. Los datos fueron obtenidos mediante el algoritmo PSI-Blast (matriz BLOSUM62) usando la base de datos GenBank NIH en la World Wide Web en ncbi.nlm.nih.gov/blast/Blast.cgi.

n.d.: "no determinado"

* Stahl et al, Mol. Microbiol. 40: 451 (2001)

** Dorner et al, Biochim. Biophys. Acta. 1667: 47-55 (2004)

5 **[0098]** Los péptidos, polipéptidos, monómeros, multímeros, proteínas, etc. descritos en el presente documento pueden modificarse y variarse adicionalmente, siempre que la función deseada se mantenga o mejore. Se entiende que una manera de definir cualquier modificación y derivado conocido o los que pudieran surgir, de los genes y las proteínas descritas en el presente documento es a través de la definición de las modificaciones y derivados en términos de identidad a secuencias específicas conocidas. Se describen específicamente polipéptidos que tienen al menos 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 por ciento de identidad a una MspA de tipo salvaje y parálogos u homólogos de MspA de tipo salvaje (por

ejemplo, MspB de tipo salvaje, MspC de tipo salvaje, MspD de tipo salvaje, MppA, PorM1, Mmcs4296), y los mutantes proporcionados en el presente documento.

[0099] Los expertos en la técnica entienden fácilmente cómo determinar la identidad de dos polipéptidos. Por ejemplo, la identidad puede calcularse después de alinear las dos secuencias de manera que la identidad se encuentra en su nivel más alto. Por ejemplo, para determinar el "porcentaje de identidad" de dos secuencias de aminoácidos o de dos ácidos nucleicos, las secuencias se alinean para propósitos de comparación óptima (por ejemplo, pueden introducirse huecos en la secuencia de una primera secuencia de aminoácidos o de ácido nucleico para una óptima alineación con una segunda secuencia de aminoácidos o ácido nucleico). Los residuos de aminoácidos o nucleótidos en las posiciones de aminoácidos o posiciones de nucleótidos correspondientes se comparan a continuación. Cuando una posición en la primera secuencia está ocupada por el mismo residuo de aminoácido o nucleótido que la posición correspondiente en la segunda secuencia, entonces las moléculas son idénticas en esa posición. El porcentaje de identidad entre las dos secuencias es función del número de posiciones idénticas compartidas por las secuencias (es decir, porcentaje de identidad = número de posiciones idénticas/número total de posiciones (por ejemplo, posiciones superpuestas) x 100). En una realización, las dos secuencias tienen la misma longitud.

[0100] Existen varios procedimientos para determinar el porcentaje de identidad. Se puede determinar el porcentaje de identidad de la manera siguiente. Se compara una secuencia de ácido nucleico o aminoácido diana con la secuencia de ácido nucleico o de aminoácidos identificada utilizando el programa de BLAST 2 Sequences (B12seq) de la versión independiente de BLASTZ que contiene BLASTN versión 2.0.14 y BLASTP versión 2.0.14. Esta versión independiente de BLASTZ se puede obtener del sitio web del Centro Nacional de Información sobre Biotecnología del gobierno de Estados Unidos (World Wide Web en ncbi.nlm.nih.gov). Las instrucciones que explican cómo utilizar el programa B12seq se pueden encontrar en el archivo "readme" que acompaña a BLASTZ.

[0101] B12seq lleva a cabo una comparación entre dos secuencias utilizando el algoritmo BLASTN o BLASTP. BLASTN se utiliza para comparar secuencias de ácido nucleico, mientras que BLASTP se utiliza para comparar secuencias de aminoácidos. Para comparar dos secuencias de ácidos nucleicos, las opciones se pueden establecer de la siguiente manera: -i se ajusta a un archivo que contiene la primera secuencia de ácido nucleico que se compara (por ejemplo, C:\seq1.txt); -j se ajusta a un archivo que contiene la segunda secuencia de ácido nucleico a comparar (por ejemplo, C:\seq2.txt); -p se ajusta a blastn; -o se ajusta a cualquier nombre de archivo deseado (por ejemplo, C:\output.txt); -q se ajusta a -1; -r se ajusta a 2; y todas las demás opciones se dejan en su configuración por defecto. El siguiente comando generará un archivo de salida que contiene una comparación entre dos secuencias: C:\B12seq -i c:\seq1.txt -j c:\seq2.txt -p blastn -o c:\output.txt -q -1-r 2. Si la secuencia diana comparte homología de secuencia con cualquier porción de la secuencia identificada, entonces el archivo de salida designado presentará esas regiones de homología como secuencias alineadas. Si la secuencia diana no comparte homología con ninguna porción de la secuencia identificada, entonces el archivo de salida designado no presentará secuencias alineadas.

[0102] Una vez alineado, se determina la longitud contando el número de nucleótidos consecutivos de la secuencia diana presentada en alineación con la secuencia de la secuencia identificada empezando con cualquier posición coincidente y terminando con cualquier otra posición coincidente. Una posición coincidente es cualquier posición en la que un nucleótido idéntico se presenta tanto en la secuencia diana como la secuencia identificada. Los huecos que se presentan en la secuencia diana no se cuentan, ya que los huecos no son nucleótidos. Del mismo modo, los huecos presentados en la secuencia identificada no se cuentan, ya que los nucleótidos de la secuencia diana se cuentan, ni los nucleótidos de la secuencia identificada.

[0103] El porcentaje de identidad sobre una longitud particular se puede determinar contando el número de posiciones coincidentes sobre esa longitud y dividiendo ese número por la longitud seguido por la multiplicación del valor resultante por 100. Por ejemplo, si (1) una secuencia diana de 50 nucleótidos se compara con la secuencia de codificación de MspA de tipo salvaje (2) el programa B12seq presenta 45 nucleótidos de la secuencia diana alineados con una región de la secuencia de codificación de MspA de tipo salvaje donde el primer y último nucleótidos de esa región de 45 nucleótidos son coincidencias, y (3) el número de coincidencias sobre esos 45 nucleótidos alineados es 40, entonces la secuencia diana de 50 nucleótidos contiene una longitud de 45 y un porcentaje de identidad sobre esa longitud de 89 (es decir, $40/45 \times 100 = 89$).

[0104] Otra forma de calcular la identidad se puede realizar mediante algoritmos publicados. El alineamiento óptimo de secuencias para comparación puede realizarse mediante el algoritmo de identidad local de Smith y Waterman, Adv. Appl. Math 2: 482 (1981), mediante el algoritmo de alineación de identidad de Needleman y Wunsch, J. Mol. Biol. 48: 443 (1970), mediante la búsqueda del procedimiento de similitud de Pearson y Lipman, Proc. Natl. Acad. Sci. USA. 85: 2444 (1988), mediante implementaciones informáticas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el paquete de software Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante inspección.

[0105] Los mismos tipos de identidad pueden obtenerse para ácidos nucleicos mediante, por ejemplo, los algoritmos descritos en Zuker, Science 244: 48-52 (1989); Jaeger et al., Proc. Natl. Acad. Sci. USA 86: 7706-10 (1989); Jaeger

et al., *Methods Enzymol.* 183: 281-306 (1989).

[0106] Se entiende que cualquiera de los procedimientos se puede utilizar normalmente y que en ciertos casos los resultados de estos diversos procedimientos pueden ser diferentes, pero el experto en la materia entiende si la identidad se encuentra con al menos uno de estos procedimientos, se diría que las secuencias tienen la identidad indicada y se describen en el presente documento.

[0107] Los ácidos nucleicos que codifican secuencias de proteínas descritos en el presente documento, así como variantes y fragmentos de los mismos, están también descritas. Estas secuencias incluyen todas las secuencias degeneradas relacionadas con una secuencia de proteína específica, es decir, todos los ácidos nucleicos que tienen una secuencia que codifica una secuencia de proteína particular, así como todos los ácidos nucleicos, incluyendo los ácidos nucleicos degenerados, que codifican las variantes descritas y derivados de las secuencias de proteínas. Por lo tanto, mientras que cada secuencia de ácido nucleico en particular puede no estar descrita en el presente documento, se entiende que todas y cada una de las secuencias está de hecho descrita y dada conocer en el presente documento a través de las secuencias de proteínas descritas.

[0108] Los fragmentos y secuencias parciales de una porina o monómero de Msp pueden ser útiles en los procedimientos descritos en el presente documento. Al igual que con todos los péptidos, polipéptidos y proteínas, incluyendo fragmentos de los mismos, se entiende que las modificaciones adicionales en la secuencia de aminoácidos de los polipéptidos de Msp descritos en el presente documento pueden aparecer sin alterar la naturaleza o función de los péptidos, polipéptidos y proteínas. Se entenderá que la única limitación en éstas es práctica, deben comprender los elementos funcionales necesarios (por ejemplo, capacidad de formación de túnel) para usar en la realización pertinente. Tales modificaciones incluyen sustituciones de aminoácidos conservadoras y se discuten en mayor detalle a continuación.

[0109] Los procedimientos para determinar si una proteína es una proteína formadora de túneles son bien conocidos en la técnica. Se puede determinar si una Msp forma un túnel mediante la determinación de si la proteína se inserta en una bicapa, tal como se describe en el Ejemplo 2 a continuación: si la proteína se inserta en la bicapa, a continuación, la porina es una proteína formadora de túneles. Típicamente, la formación de túneles se detecta mediante la observación de un cambio discreto en la conductividad. Véase, por ejemplo, FIGURA 2, Ejemplo 2, y Niederweis et al., *Mol. Microbiol.* 33: 933 (1999). Las bicapas se describen en el presente documento.

[0110] Como se sugirió anteriormente, una porina de Msp típicamente será capaz de ser insertada en una bicapa lipídica o película delgada de otro tipo, que son conocidas en la técnica. Un ejemplo de inserción de una porina de MspA mutante en una bicapa lipídica se explica en el presente documento; esta técnica se puede aplicar también a otras porinas de Msp. Además, la patente de Estados Unidos N° 6.746.594 describe una variedad de bicapas lipídicas y películas delgadas, incluyendo materiales inorgánicos, que se pueden emplear con respecto a las porinas de Msp descritas en el presente documento. Los procedimientos, aparatos y técnicas que se describen en la patente de Estados Unidos N° 6.267.872, también se pueden emplear con respecto a las porinas de Msp descritas en el presente documento.

[0111] Además, más de una porina de Msp puede estar comprendida en una bicapa lipídica. Por ejemplo, 2, 3, 4, 5, 10, 20, 200, 2000, o más pueden estar comprendidas en una bicapa lipídica. Opcionalmente, cualquiera de 2 a 10^{10} porinas de Msp se puede emplear en los procedimientos descritos en el presente documento. Dicha pluralidad de porinas de Msp pueden estar en forma de grupos de porinas de Msp. Los grupos pueden estar ensamblados al azar o pueden adoptar un patrón. Como se usa en el presente documento, un "grupo" se refiere moléculas que se agrupan juntas y se mueven como una unidad, pero no están unidas covalentemente entre sí.

[0112] Opcionalmente, las porinas de Msp no son reguladas ("gate") espontáneamente. "Regular" o "gating" se refiere al cambio espontáneo de la conductancia eléctrica a través del túnel de la proteína que es generalmente temporal (por ejemplo, una duración de tan sólo 1 a 10 milisegundos hasta un segundo). Los sucesos de regulación de larga duración a menudo pueden invertirse mediante el cambio de la polaridad. Bajo la mayoría de circunstancias, la probabilidad de regulación aumenta con la aplicación de voltajes más altos. La regulación y el grado de conductancia a través del cambio de túnel son muy variables entre porinas de Msp, dependiendo de, por ejemplo, la formación del vestíbulo y la zona de constricción, así como de las propiedades del medio líquido en el que se sumerge la proteína. Típicamente, la proteína se hace menos conductora durante la regulación, y la conductancia puede parar permanentemente (es decir, el túnel puede cerrar de forma permanente) como resultado, de manera que el proceso es irreversible. Opcionalmente, la regulación se refiere a la conductancia a través del túnel de una proteína que cambia espontáneamente a menos del 75% de su de corriente en estado abierto.

[0113] Varias condiciones, tales como la luz y el medio líquido que contacta con una porina de Msp, incluyendo su pH, composición del tampón, composición detergente, y la temperatura, pueden afectar el comportamiento de una porina de Msp, particularmente con respecto a su conductancia a través del túnel como, así como el movimiento de un analito con respecto al túnel, ya sea temporal o permanentemente.

[0114] De especial relevancia es la geometría de los túneles de porina de Msp, en particular la porina de MspA. La

geometría de la porina de Msp puede proporcionar una mejor resolución espacial. Además, la porina de MspA de tipo salvaje es muy robusta y retiene la actividad de formación de túnel después de la exposición a cualquier pH y después de la extracción a temperaturas extremas (por ejemplo, hasta 100°C durante un máximo de 30 minutos y una incubación hasta 80°C durante hasta 15 minutos). Los polipéptidos pueden ser analizados por su actividad deseada usando los ensayos in vitro descritos en el presente documento.

[0115] En cuanto a la porina de MspA en particular, opcionalmente, la porina de MspA es un octámero que consiste en ocho monómeros de MspA de 184 aminoácidos. Una o más mutaciones pueden tener lugar en uno o más de los monómeros de MspA en los aminoácidos de una porina de MspA de tipo salvaje para producir una porina de MspA mutante. Además, una porina de MspA puede tener menos o más de ocho monómeros, uno o más de los cuales puede comprender una mutación.

[0116] Además, la porina de MspA de tipo salvaje comprende un bucle periplásmico que consiste en trece aminoácidos y está directamente adyacente a la zona de constricción. Véase Huff et al., J. Biol. Chem. 284: 10223 (2009). Las porinas de tipo salvaje MspB, C, y D también contienen un bucle periplásmico. Una o más mutaciones pueden aparecer en el bucle periplásmico de una porina de Msp de tipo salvaje para generar una porina de Msp mutante. Por ejemplo, pueden aparecer deleciones de hasta los trece aminoácidos en el bucle periplásmico de la porina de MspA de tipo salvaje. Típicamente, las deleciones en el bucle periplásmico no afectan a la capacidad de formación de túnel de una porina de Msp.

[0117] Una porina de Msp o monómero Msp también se pueden modificar química o biológicamente. Por ejemplo, se puede modificar una porina de Msp o monómero Msp con productos químicos para producir puentes disulfuro, tal como es conocido por los expertos en la técnica.

[0118] Una porina de Msp puede comprender un sitio de unión a nucleótidos. Como se usa en el presente documento, un "sitio de unión a nucleótidos" se refiere a un sitio en una porina de Msp donde un nucleótido permanece en contacto con, o reside en, un aminoácido durante un período de tiempo que es más largo que lo atribuible al movimiento de difusión, tal como mayor de un picosegundo o un nanosegundo. Se pueden emplear cálculos de dinámica molecular para evaluar estos tiempos de reposo temporales.

[0119] Un "vestíbulo" se refiere a la porción en forma de cono del interior de una porina de Msp cuyo diámetro disminuye generalmente desde un extremo al otro a lo largo de un eje central, donde la porción más estrecha del vestíbulo está conectada a la zona de constricción. Un vestíbulo también puede denominarse como una "copa". Ver la FIGURA 1 para un ejemplo del vestíbulo de una porina MspA de tipo salvaje. El vestíbulo y la zona de constricción definen conjuntamente el túnel de una porina de MSP.

[0120] Cuando se hace referencia a un diámetro del vestíbulo, se entiende que debido a que el vestíbulo tiene forma de cono, el diámetro cambia a lo largo de la trayectoria de un eje central, en el que el diámetro es mayor en un extremo que el extremo opuesto. El diámetro puede variar de aproximadamente 2 nm a aproximadamente 6 nm. Opcionalmente, el diámetro es de aproximadamente, al menos aproximadamente, o como máximo de aproximadamente 2, 2,1, 2,2, 2,3, 2,4, 2,5, 2,6, 2,7, 2,8, 2,9, 3,0, 3,1, 3,2, 3,3, 3,4, 3,5, 3,6, 3,7, 3,8, 3,9, 4,0, 4,1, 4,2, 4,3, 4,4, 4,5, 4,6, 4,7, 4,8, 4,9, 5,0, 5,1, 5,2, 5,3, 5,4, 5,5, 5,6, 5,7, 5,8, 5,9, o 6,0 nm, o cualquier intervalo derivable en el mismo. La longitud del eje central puede variar de aproximadamente 2 nm a aproximadamente 6 nm. Opcionalmente, la longitud es de aproximadamente, al menos aproximadamente, o como máximo de aproximadamente 2, 2,1, 2,2, 2,3, 2,4, 2,5, 2,6, 2,7, 2,8, 2,9, 3,0, 3,1, 3,2, 3,3, 3,4, 3,5, 3,6, 3,7, 3,8, 3,9, 4,0, 4,1, 4,2, 4,3, 4,4, 4,5, 4,6, 4,7, 4,8, 4,9, 5,0, 5,1, 5,2, 5,3, 5,4, 5,5, 5,6, 5,7, 5,8, 5,9, o 6,0 nm, o cualquier intervalo derivable en el mismo. Cuando se hace referencia a "diámetro" en el presente documento, se puede determinar un diámetro mediante la medición de distancias de centro a centro o distancias de superficie a superficie atómica.

[0121] Una "zona de constricción" se refiere a la porción más estrecha del túnel de una porina de Msp, en términos de diámetro, que se conecta a los vestíbulos. La zona de constricción de una porina de MspA de tipo salvaje se muestra en la FIGURA 1 (marcada como "constricción interior"). La longitud de la zona de constricción puede variar de aproximadamente 0,3 nm a aproximadamente 2 nm. Opcionalmente, la longitud es de aproximadamente, como máximo, aproximadamente, o al menos aproximadamente 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2 o 3 nm, o cualquier intervalo derivable en el mismo. El diámetro de la zona de constricción puede variar de aproximadamente 0,3 nm a aproximadamente 2 nm. Opcionalmente, el diámetro es de aproximadamente, como máximo, aproximadamente, o al menos aproximadamente 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2 o 3 nm, o cualquier intervalo derivable en el mismo.

[0122] Una "zona de constricción neutra" se refiere a una zona de constricción que comprende cadenas laterales de aminoácidos que acumulativamente exhiben una carga eléctrica no neta cuando se sumerge en una solución acuosa. El pH del medio líquido (por ejemplo, una solución acuosa tamponada) en contacto con la zona de constricción puede afectar si la zona de constricción se caracteriza como neutra o no.

[0123] Un "túnel" se refiere a la parte central, vacía de una Msp que se define por el vestíbulo y la zona de constricción, a través del cual puede pasar un gas, líquido, ion, o analito.

[0124] Tal como se usa en el presente documento, "cis" se refiere a la cara de un túnel de Msp a través de la cual un analito entra en el túnel o a lo largo de la cara en la cual se mueven los analitos.

5 **[0125]** Tal como se usa en el presente documento, "trans" se refiere a la cara de un túnel de Msp a través de la cual un analito (o fragmentos de los mismos) sale del túnel o a lo largo de la cara en la cual no se mueven los analitos.

10 **[0126]** Tal como se usa en el presente documento, "trasladar o trasladar electroforéticamente un analito," y las variantes gramaticales del mismo, se refiere a la aplicación de un campo eléctrico a una porina de Msp que está en contacto con una o más soluciones (por ejemplo, sumergida en una solución), de tal manera que la corriente fluye a través del túnel de la porina de Msp. El campo eléctrico mueve un analito de tal manera que interactúa con el túnel. Por "interactúa", se entiende que el analito se mueve en y, opcionalmente, a través del túnel, donde "a través del túnel de Msp" (o "traslada" o "traslada") significa entrar por una cara del túnel y pasar a y salir de la otra cara del túnel.

15 **[0127]** Se contempla específicamente que cualquier analito descrito en el presente documento puede trasladarse a través de un túnel de porina de Msp, ya sea por electroforesis o no, en cualquier realización descrita en el presente documento. En este sentido, se contempla específicamente que cualquier realización del presente documento que comprende el traslado puede referirse al traslado electroforético o el traslado no electroforético, a menos que se indique específicamente. Opcionalmente, se contemplan procedimientos que no emplean traslado electroforética.

20 **[0128]** Un "medio líquido" incluye medios líquidos acuosos, orgánico-acuosos y solo orgánicos. Los medios orgánicos incluyen, por ejemplo, metanol, etanol, dimetilsulfóxido, y mezclas de los mismos. Los líquidos empleables en procedimientos descritos aquí son bien conocidos en la técnica. Las descripciones y ejemplos de tales medios, incluyendo medios líquidos conductores, se proporcionan en la patente de Estados Unidos N° 7.189.503, por ejemplo.

25 **[0129]** Las sales, detergentes, o tampones pueden añadirse a tales medios. Tales agentes se pueden emplear para alterar el pH o la fuerza iónica del medio líquido. Las sustancias que alteran la viscosidad, tales como glicerol o varios polímeros (por ejemplo, polivinilpirrolidona, polietilenglicol, alcohol polivinílico, polímeros de celulosa), y mezclas de los mismos, se pueden incluir en medios líquidos. Los procedimientos de medición de la viscosidad son bien conocidos en la técnica. Cualquier agente que puede añadirse a un medio líquido también puede alterar la velocidad de un analito que se está estudiando. Por tanto, un agente que altera la velocidad puede ser una sal, un detergente, un tampón, una sustancia que altera la viscosidad, o cualquier otro agente añadido a un medio líquido que aumenta o disminuye la velocidad de un analito.

30 **[0130]** Típicamente, un analito empleado en el presente documento es soluble o parcialmente soluble en al menos un medio líquido que está en contacto con una Msp descrita en el presente documento. Cualquier analito se puede usar en el presente documento, incluyendo, por ejemplo, un nucleótido, un ácido nucleico, un aminoácido, un péptido, una proteína, un polímero, un fármaco, un ion, un agente de guerra biológica, un contaminante, un objeto nanoscópico, o cualquier otra molécula que comprende uno de estos analitos o una combinación de los mismos. Un analito puede ser un grupo de moléculas, en que el grupo en su conjunto se considera un analito. Típicamente, el tamaño de un analito no será tan grande de manera que no pueda entrar en un túnel de una Msp: en otras palabras, un analito típico será más pequeño en tamaño que la apertura de un túnel de una Msp. Sin embargo, un analito que tiene un tamaño más grande que la apertura de un túnel puede emplearse, y puede determinarse usando procedimientos descritos en el presente documento que el tamaño del analito es demasiado grande para entrar en el túnel. Opcionalmente, el peso molecular del analito está a menos de un millón de Da. Opcionalmente, el peso molecular del analito es de aproximadamente, como máximo aproximadamente, o al menos aproximadamente 1.000.000, 950.000, 900.000, 850.000, 800.000, 750.000, 700.000, 650.000, 600.000, 550.000, 500.000, 450.000, 400.000, 350.000, 300.000, 250.000, 200.000, 150.000, 100.000, 75.000, 50.000, 25.000, 20.000, 15.000, 10.000, 7.500, 5.000, 2.500, 2.000, 1.500, 1.000, o 500 Da o menos, o cualquier intervalo derivable en el mismo.

35 **[0131]** Las modificaciones de las proteínas incluyen modificaciones en las secuencias de aminoácidos. Las modificaciones en la secuencia de aminoácidos pueden surgir de forma natural como variaciones alélicas (por ejemplo, debido al polimorfismo genético), pueden surgir debido a la influencia del medio ambiente (por ejemplo, debido a la exposición a la radiación ultravioleta), o pueden producirse por intervención humana (por ejemplo, por mutagénesis de secuencias de ADN clonadas), tales como mutantes de punto inducido, delección, inserción, y sustitución. Estas modificaciones pueden dar como resultado cambios en la secuencia de aminoácidos, proporcionar mutaciones silenciosas, modificar un sitio de restricción o proporcionar otras mutaciones específicas. Las modificaciones en la secuencia de aminoácidos normalmente se encuentran en una o más de tres clases: modificaciones por sustitución, inserción o delección. Las inserciones incluyen fusiones amino y/o terminales así como inserciones intrasecuencia de residuos de aminoácidos individuales o múltiples. Las inserciones normalmente serán inserciones más pequeñas que las fusiones de amino o carboxi terminales, por ejemplo, del orden de uno a cuatro residuos. Las delecciones se caracterizan por la eliminación de uno o más residuos de aminoácidos de la secuencia de la proteína. Normalmente, no más de aproximadamente de 2 a 6 residuos se eliminan en cualquier sitio dentro de la molécula de proteína. Las sustituciones de aminoácidos son típicamente residuos individuales, pero

pueden aparecer en un número de localizaciones diferentes a la vez; las inserciones normalmente serán del orden de aproximadamente de 1 a 10 residuos de aminoácidos; y las deleciones oscilarán aproximadamente de 1 a 30 residuos. Las deleciones o inserciones preferiblemente se hacen en pares adyacentes, es decir, una deleción de 2 residuos o inserción de 2 residuos. Las sustituciones, deleciones, inserciones o cualquier combinación de los mismos pueden combinarse para llegar a una construcción final. Las mutaciones pueden colocar o no la secuencia fuera del marco de lectura y pueden crear o no regiones complementarias que puedan producir una estructura secundaria de ARNm. Las modificaciones de sustitución son aquellas en las que al menos un residuo se ha eliminado y un residuo diferente se ha insertado en su lugar.

[0132] Las modificaciones, incluyendo las sustituciones de aminoácidos específicos, se realizan por procedimientos conocidos. A modo de ejemplo, las modificaciones se realizan mediante mutagénesis específicas de sitio de nucleótidos en el ADN que codifica la proteína, produciendo de este modo ADN que codifica la modificación, y a continuación expresando el ADN en cultivo de células recombinantes. Las técnicas para realizar mutaciones de sustitución en sitios predeterminados en ADN que tiene una secuencia conocida son bien conocidas, por ejemplo mutagénesis del cebador M13 y la mutagénesis por PCR.

[0133] Una o más mutaciones en una porina de Msp pueden aparecer en el vestíbulo o en la zona de constricción de la proteína. Opcionalmente, una porina de Msp mutante tiene al menos una diferencia en su secuencia de aminoácidos en el bucle periplásmico, vestíbulo, o zona de constricción (por ejemplo, deleción, sustitución, adición) en comparación con la porina de Msp de tipo salvaje.

[0134] Tal como se usa en el presente documento, un "aminoácido" se refiere a cualquiera de los 20 aminoácidos de origen natural encontrados en las proteínas, D-estereoisómeros de los aminoácidos de origen natural (por ejemplo, D-treonina), aminoácidos no naturales, y aminoácidos modificados químicamente. Cada uno de estos tipos de aminoácidos no es mutuamente excluyentes. Los α -aminoácidos comprenden un átomo de carbono al que está unido un grupo amino, un grupo carboxilo, un átomo de hidrógeno y un grupo distinto referido como una "cadena lateral". Las cadenas laterales de aminoácidos de origen natural son bien conocidos en la técnica e incluyen, por ejemplo, hidrógeno (por ejemplo, como en glicina), alquilo (por ejemplo, como en alanina, valina, leucina, isoleucina, prolina), alquilo sustituido (por ejemplo, como en treonina, serina, metionina, cisteína, ácido aspártico, asparagina, ácido glutámico, glutamina, arginina y lisina), arilalquilo (por ejemplo, como en fenilalanina y triptófano), arilalquilo sustituido (por ejemplo, como en tirosina), y heteroarilalquilo (por ejemplo, como en histidina).

[0135] Las siguientes abreviaturas se usan para los 20 aminoácidos de origen natural: alanina (Ala; A), asparagina (Asn; N), ácido aspártico (Asp; D), arginina (Arg; R), cisteína (Cys; C), ácido glutámico (Glu; E), glutamina (Gln; Q), glicina (Gly; G), histidina (His; H), isoleucina (Ile; I), leucina (Leu; L), lisina (Lys; K), metionina (Met; M), fenilalanina (Phe; F), prolina (Pro; P), serina (Ser; S), treonina (Thr; T), triptófano (Trp; W), tirosina (Tyr; Y) y valina (Val; V).

[0136] También son conocidos aminoácidos no naturales (es decir, aquellos que no se encuentran de forma natural en las proteínas) en la técnica, como se establece en, por ejemplo, Williams et al., Mol. Cell Biol. 9: 2574 (1989); Evans et al., J. Amer. Chem. Soc. 112: 4.011 a 4030 (1990); Pu et al., J. Amer. Chem. Soc. 56: 1280-1283 (1991); Williams et al., J. Amer. Chem. Soc. 113: 9.276-9.286 (1991); y todas las referencias citadas en los mismos. Los β - y γ -aminoácidos son conocidos en la técnica y también se contemplan en el presente documento como aminoácidos no naturales. La siguiente tabla muestra ejemplos no limitativos de aminoácidos no naturales que se contemplan en el presente documento.

Tabla 2. Aminoácidos no naturales de ejemplo

Abreviatura	Aminoácido	Abreviatura	Aminoácido
Aad	ácido 2-aminoadípico	EtAsn	N-Etilasparagina
Baad	ácido 3-aminoadípico	Hyl	Hidroxilisina
Bala	β -alanina, ácido β -amino-propiónico	AHyl	alo-Hidroxilisina
Abu	ácido 2-Aminobutírrico	3Hyp	3-Hidroxiprolina
4Abu	ácido 4-aminobutírrico, ácido piperidínico	4Hyp	4-Hidroxiprolina
Acp	ácido 6-aminocaproico	Ide	Isodesmosina
Ahe	ácido 2-aminoheptanoico	Alle	alo-Isoleucina
Aib	ácido 2-aminoisobutírrico	MeGly	N-Methylglicina, sarcosina
Baib	ácido 3-aminoisobutírrico	Melle	N-Metilisoleucina
Apm	ácido 2-aminopiméico	MeLys	6-N-Metil lisina

Dbu	ácido 2,4-diaminobutírico	MeVal	N-Metilvalina
Des	Desmosina	Nva	Norvalina
Dpm	ácido 2,2'-diaminopimélico	Nle	Norleucina
Dpr	ácido 2,3-diaminopropiónico	Orn	Ornitina
EtGly	N-Etilglicina		

[0137] Tal como se usa en el presente documento, un "aminoácido modificado químicamente" se refiere a un aminoácido cuya cadena lateral ha sido modificada químicamente. Por ejemplo, una cadena lateral puede ser modificada para comprender un grupo de señalización, tal como un fluoróforo o un radiomarcador. Una cadena lateral puede modificarse para comprender un nuevo grupo funcional, tal como un grupo tiol, ácido carboxílico, o amino. Los aminoácidos modificados después de la traducción también se incluyen en la definición de aminoácidos modificados químicamente.

[0138] Los aminoácidos, y, más específicamente, sus cadenas laterales, se puede caracterizar por su característica o características químicas. Por ejemplo, las cadenas laterales de aminoácidos pueden estar cargadas positivamente, negativamente o ser neutras. El pH de una solución afecta a la naturaleza cargada de ciertas cadenas laterales, tal como se conoce por los expertos en la técnica. Los ejemplos no limitantes de cadenas laterales que pueden estar cargadas positivamente incluyen histidina, arginina y lisina. Los ejemplos no limitantes de cadenas laterales que pueden estar cargadas negativamente incluyen ácido aspártico y ácido glutámico. Los ejemplos no limitantes de cadenas laterales que pueden ser caracterizadas como neutras incluyen glicina, alanina, fenilalanina, valina, leucina, isoleucina, cisteína, asparagina, glutamina, serina, treonina, tirosina, metionina, prolina y triptófano.

[0139] El nivel estérico de cadenas laterales también se puede usar para caracterizar un aminoácido. Las tablas de diámetros de los átomos pueden ayudar en la determinación de si una cadena lateral es más grande que otra. Los modelos informáticos también pueden ayudar con esta determinación.

[0140] Los aminoácidos se pueden caracterizar por la polaridad de sus cadenas laterales. Las cadenas laterales polares, que son típicamente más hidrófilas que las cadenas laterales no polares, incluyen, por ejemplo, las de serina, treonina, tirosina, cisteína, asparagina y glutamina. Las cadenas laterales no polares, que son típicamente más hidrófobas que las cadenas laterales polares, incluyen, por ejemplo, las de glicina, alanina, valina, leucina, isoleucina, prolina, metionina, fenilalanina y triptófano. Se puede determinar la polaridad de una cadena lateral usando técnicas convencionales conocidas en el sector que implica determinaciones de electronegatividad del átomo y evaluaciones estructurales tridimensionales de cadenas laterales. También se puede comparar la hidrofobicidad/hidrofilia de las cadenas laterales usando técnicas convencionales conocidas en el sector, tales como la comparación del coeficiente de distribución en octanol/agua de cada aminoácido. Véase Sangster, en: Octanol-Water partition coefficients: Fundamentals and Physical Chemistry, la serie Wiley en Solution Chemistry, Chichester: John Wiley & Sons Ltd., 2: 178 páginas (1997).

[0141] La siguiente tabla proporciona ejemplos no limitativos de propiedades de los aminoácidos que pueden ayudar a un experto en la materia a determinar cómo seleccionar los aminoácidos para las modificaciones de una porina o monómero de Msp tal como se describe en el presente documento.

Tabla 3. Propiedades de aminoácidos

Aminoácido	Porcentaje de residuos escondidos ^a (%)	Volumen promedio ^b (Å ³)	Volumen de van der Waals ^c (Å ³)	Área superficial accesible (Å ²)	Clasificación de las polaridades e los aminoácidos ^e
alanina	38 (12)	92	67	67	9 (7)
arginina	0	225	148	196	15 (19)
asparagina	10 (2)	135	96	113	16 (16)
ácido aspártico	14,5 (3)	125	91	106	19 (18)
cisteína	47 (3)	106	86	104	7 (8)
glutamina	6,3 (2,2)	161	114	144	17 (14)
ácido glutámico	20 (2)	155	109	138	18 (17)

glicina	37 (10)	66	48		11 (9)
histidina	19 (1,2)	167	118	151	10 (13)
isoleucina	65 (12)	169	124	140	1 (2)
leucina	41 (10)	168	124	137	3 (1)
lisina	4.2 (0,1)	171	135	167	20 (15)
metionina	50 (2)	171	124	160	5 (5)
fenilalanina	48 (5)	203	135	175	2 (4)
prolina	24 (3)	129	90	105	13 (-)
serina	24 (8)	99	73	80	14 (12)
treonina	25 (5,5)	122	93	102	12 (11)
triptófano	23 (1,5)	240	163	217	6 (6)
tirosina	13 (2,2)	203	141	187	8 (10)
valina	56 (15)	142	105	117	4 (3)

^a Esta columna representa la tendencia de un aminoácido a estar escondido (definido como <5% de los residuos de disposición al disolvente) en el interior de una proteína y se basa en las estructuras de nueve proteínas (total de ~ 2000 residuos individuales estudiados, con 587 (29%) de estos escondidos). Los valores indican la frecuencia en que cada aminoácido se encontró escondido, en relación con el número total de residuos de este aminoácido que se encuentra en las proteínas. Los valores entre paréntesis indican el número de residuos escondidos de este aminoácido que se encuentran en relación con todos los residuos escondidos en las proteínas. Los datos de Schien, *Biotechnology* 8: 308 (1990); para otros procedimientos de cálculo con resultados similares, véase Janin, *Nature* 277: 491 (1979); y Rose et al, *Science* 229: 834 (1985)

^b Volumen promedio (V_r) de los residuos escondidos, calculados a partir del área superficial de la cadena lateral. Richards, *Annu. Rev. Biophys. Bioeng.* 6: 151 (1977); Baumann, *Protein Eng.* 2: 329 (1989).

^c Los datos de Darby NJ y Creighton T.E. *Protein structure En el foco* (Ed. D. Rickwood), p. 4. IRL Press, Oxford, Reino Unido (1993).

^d Área superficial accesible total (ASA) de la cadena lateral de aminoácido para el residuo X en un tripéptido Gly-X-Gly con la cadena principal en una conformación extendida. Miller et al., *J. Mol. Biol.* 196: 641 (1987)

^e Valores mostrados representan la clasificación promedio de aminoácidos de acuerdo con la frecuencia de su aparición en cada clasificación de secuencia de 38 escalas de hidrofobicidad publicados. Trinquier y Sanejouand, *Protein Eng.* 11: 153 (1998). Aunque la mayoría de estas escalas de hidrofobicidad se derivan de medidas experimentales de comportamiento químico o propiedades fisicoquímicas (por ejemplo, solubilidad en agua, la distribución entre agua y disolvente orgánico, la migración cromatográfica, o efectos sobre la tensión superficial) de aminoácidos aislados, se incluyen varias escalas de hidrofobicidad "operativas" en base a las características de entorno conocidas de aminoácidos en las proteínas, tales como su accesibilidad al disolvente o su inclinación a ocupar el núcleo de proteínas (en base a la posición de los residuos en las estructuras terciarias como se observa por cristalografía de rayos X o RMN). Las clasificaciones más bajas representan los aminoácidos más hidrófobos, y los valores más altos representan los aminoácidos más hidrófilos. Con fines comparativos, la escala de hidrofobicidad de Radzicka y Wolfenden, *Biochem.* 27: 1664 (1988) se muestra entre paréntesis. Esa escala se deriva de la posibilidad de hidratación medido de aminoácidos que se basa en sus energías libres de transferencia desde la fase de vapor a ciclohexano, 1-octanol y solución acuosa neutra.

5 **[0142]** Alternativamente, se puede considerar el índice hidropático de los aminoácidos. A cada aminoácido se le ha asignado un índice hidropático sobre la base de su hidrofobicidad y/o características de carga, éstos son: isoleucina (+4,5); valina (4,2); leucina (3,8); fenilalanina (2,8); cisteína/cistina (2,5); metionina (1,9); alanina (1,8); glicina (-0,4); treonina (-0,7); serina (-0,8); triptófano (-0,9); tirosina (-1,3); prolina (-1,6); histidina (-3,2); glutamato (-3,5); glutamina (-3,5); aspartato (-3,5); asparagina (-3,5); lisina (-3,9); y/o arginina (-4,5). La importancia del índice hidropático de los aminoácidos para conferir una función biológica interactiva a una proteína se entiende generalmente en la técnica. Se sabe que ciertos aminoácidos pueden ser sustituidos por otros aminoácidos que tienen un índice hidropático y/o puntuación similar y/o todavía conservan una actividad biológica similar. Al hacer cambios basados en el índice hidropático, la sustitución de aminoácidos cuyos índices hidropáticos puede estar dentro de ± 2 ; dentro de ± 1 , o dentro de $\pm 0,5$.

15 **[0143]** También se entiende en la técnica que la sustitución de aminoácidos similares puede realizarse eficazmente sobre la base de la hidrofiliidad. Como se detalla en la patente de Estados Unidos 4.554.101, se han asignado los siguientes valores de hidrofiliidad a los residuos de aminoácidos: arginina (+3,0); lisina (3,0); aspartato ($3,0 \pm 1$); glutamato ($3,0 \pm 1$); serina (0,3); asparagina (+0,2); glutamina (0,2); glicina (0); treonina (-0,4); prolina ($-0,5 \pm 1$); alanina (-0,5); histidina (-0,5); cisteína (-1,0); metionina (-1,3); valina (-1,5); leucina (-1,8); isoleucina (-1,8); tirosina (-

2,3); fenilalanina (-2,5); triptófano (-3,4). Al hacer cambios basados en valores de hidrofiliidad similares, se contempla que en la sustitución de aminoácidos los valores de hidrofiliidad pueden estar dentro de ± 2 , dentro de ± 1 , o dentro de $\pm 0,5$.

5 **[0144]** Cualquier porina o monómero de Msp mutante puede comprender una sustitución de aminoácido conservativa en comparación con un porina o monómero de Msp de tipo salvaje. Cualquier mutación de sustitución es conservativa en que mínimamente altera las propiedades bioquímicas de la proteína. Ejemplos no limitativos de mutaciones que se introducen para sustituir residuos de aminoácidos conservadores incluyen: residuos de carga positiva (por ejemplo, H, K, y R) sustituidos por residuos de carga positiva; residuos de carga negativa (por ejemplo, D y E) sustituidos por residuos de carga negativa; residuos polares neutros (por ejemplo, C, G, N, Q, S, T, e Y) sustituidos por residuos polares neutros; y los residuos no polares neutros (por ejemplo, A, F, I, L, M, P, V, y W) sustituidos por residuos no polares neutros. Las sustituciones conservativas pueden realizarse de acuerdo con la siguiente Tabla 4. Las sustituciones no conservativas pueden realizarse también (por ejemplo, prolina por glicina).

15 **Tabla 4. Ejemplos de sustituciones de aminoácidos**

Aminoácido	Sustituciones
Ala	Ser, Gly, Cys
Arg	Lys, Gln, Met, Ile
Asn	Gln, His, Glu, Asp
Asp	Glu, Asn, Gln
Cys	Ser, Met, Thr
Gln	Asn, Lys, Glu, Asp
Glu	Asp, Asn, Gln
Gly	Pro, Ala
His	Asn, Gln
Ile	Leu, Val, Met
Leu	Ile, Val, Met
Lys	Arg, Gln, Met, Ile
Met	Leu, Ile, Val
Phe	Met, Leu, Tyr, Trp, His
Ser	Thr, Met, Cys
Thr	Ser, Met, Val
Trp	Tyr, Phe
Tyr	Trp, Phe, His
Val	Ile, Leu, Met

20 **[0145]** Tal como se usa en el presente documento, un "péptido" se refiere a dos o más aminoácidos unidos entre sí por un enlace amida (es decir, un "enlace peptídico"). Los péptidos comprenden o incluyen hasta 50 aminoácidos. Los péptidos pueden ser lineales o cíclicos. Los péptidos pueden ser α , β , γ , δ , o superior, o mixto. Los péptidos pueden comprender cualquier mezcla de aminoácidos tal como se define en el presente documento, tales como que comprende cualquier combinación de aminoácidos D, L, α , β , γ , δ , o superior.

25 **[0146]** Tal como se usa en el presente documento, una "proteína" se refiere a una secuencia de aminoácidos que tiene 51 o más aminoácidos.

30 **[0147]** Tal como se usa en el presente documento, un "polímero" se refiere a una molécula que comprende dos o más unidades lineales (también conocidas como un "mers"), donde cada unidad puede ser la misma o diferente. Los ejemplos no limitantes de polímeros incluyen ácidos nucleicos, péptidos y proteínas, así como una variedad de polímeros de hidrocarburos (por ejemplo, polietileno, poliestireno) y polímeros de hidrocarburos funcionalizados, en los que la cadena principal del polímero comprende una cadena de carbono (por ejemplo, cloruro de polivinilo, polimetacrilatos). Los polímeros incluyen copolímeros, copolímeros de bloque y polímeros ramificados, tales como

polímeros en estrella y dendrímeros.

[0148] Los procedimientos de secuenciación utilizando porinas de Msp se describen en el presente documento. Además, los procedimientos de secuenciación se pueden realizar en procedimientos análogos a los descritos en la patente de Estados Unidos N° 7.189.503. Véase también la patente de Estados Unidos N° 6.015.714. Puede llevarse a cabo más de una lectura en dichos procedimientos de secuenciación para mejorar la precisión. Los procedimientos de análisis de las características de los polímeros (por ejemplo, tamaño, longitud, concentración, identidad) y la identificación de unidades discretas (o "mers") de los polímeros se discuten en la patente '503 también, y se pueden emplear con respecto a las presentes porinas de Msp. De hecho, una porina de Msp se puede emplear con respecto a cualquier procedimiento descrito en la patente '503.

[0149] En la actualidad, se están explorando varios tipos de señales observables como mecanismos de lectura en la secuenciación de nanoporos y detección de analitos. El procedimiento de lectura originalmente propuesto, más directo, y más explorado se basa en una "corriente de bloqueo iónica", o "corriente de copaso" determinada únicamente por la identidad de un nucleótido o de otro analito que ocupa la constricción más estrecha en el poro. Este procedimiento se conoce como "secuenciación de nanoporos de la corriente de bloqueo" o BCNS. La detección y caracterización de la corriente de bloqueo de ácidos nucleicos se ha demostrado tanto en la α -hemolisina de poro de proteína (α -HL) como en nanoporos de estado sólido. Se ha observado que la detección y caracterización de la corriente de bloqueo proporciona una gran cantidad de información acerca de la estructura del ADN de paso, o que se retiene en, un nanoporo en varios contextos.

[0150] En general, un "bloqueo" se evidencia por un cambio en la corriente de iones que se distingue claramente de las fluctuaciones de ruido y por lo general se asocia con la presencia de una molécula de analito en la abertura central del poro. La fuerza de bloqueo dependerá del tipo de analito que está presente. Más particularmente, un "bloqueo" se refiere a un intervalo en el que la corriente iónica cae por debajo de un umbral de aproximadamente 5-100% del nivel de corriente no bloqueado, se mantiene ahí durante al menos 1,0 μ s, y vuelve espontáneamente al nivel no bloqueado. Por ejemplo, la corriente iónica puede caer por debajo de un umbral de aproximadamente, al menos aproximadamente, o como máximo aproximadamente de 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50 %, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, o 100%, o cualquier intervalo derivable en el mismo. Los bloques se rechazan si la señal no bloqueada inmediatamente anterior o siguiente tiene una corriente media que se desvía del nivel no bloqueado típico en más de dos veces el ruido rms de la señal no bloqueada. "Bloqueos profundos" son identificados como intervalos en los que la corriente iónica cae <50% del nivel no bloqueado. Los intervalos donde la corriente se mantiene entre 80% y 50% del nivel de no bloqueado se identifican como "bloqueos parciales."

[0151] Tal como se utiliza en el presente documento, el término "sujeto" se refiere a un organismo mamífero vivo, tal como un humano, mono, vaca, oveja, cabra, perros, gato, ratón, rata, cobaya, o especies transgénicas de los mismos. Opcionalmente, el paciente o sujeto es un primate. Los ejemplos no limitantes de los sujetos humanos son adultos, jóvenes, bebés y fetos.

[0152] El término "ácido nucleico" se refiere a un polímero de desoxirribonucleótidos o ribonucleótidos en forma monocatenaria o bicatenaria, y a menos que se limite de otro modo, abarca análogos conocidos de nucleótidos naturales que se hibridan con ácidos nucleicos de manera similar a nucleótidos de origen natural, tales como ácidos nucleicos peptídicos (PNA) y ADN fosforotioato. A menos que se indique lo contrario, una secuencia de ácido nucleico particular incluye la secuencia complementaria de la misma. Los nucleótidos incluyen, pero no se limitan a, ATP, dATP, CTP, dCTP, GTP, dGTP, UTP, TTP, dUTP, 5-metil-CTP, 5-metil-dCTP, ITP, dITP, 2-amino-adenosina-TP, 2-amino-desoxiadenosina-TP, trifosfato de 2-tiotimidina, trifosfato de pirrolo-pirimidina, y 2-tiocitidina, así como los alfatiotriofosfatos de todos los anteriores, y trifosfatos de 2'-O-metil-ribonucleótido para todas las bases anteriores. Las bases modificadas incluyen, pero no se limitan a, 5-Br-UTP, 5-Br-dUTP, 5-F-UTP, 5-F-dUTP, 5-propinil dCTP, y 5-propinil-dUTP.

[0153] Tal como se usa en el presente documento, un "fármaco" se refiere a cualquier sustancia que pueda alterar un proceso biológico de un sujeto. Los medicamentos pueden ser diseñados o utilizados para o en el diagnóstico, tratamiento o prevención de una enfermedad, trastorno, síndrome, u otra afección de salud de un sujeto. Los fármacos pueden ser de recreativos por naturaleza, es decir, utilizados simplemente para alterar un proceso biológico y no utilizarse para o en el diagnóstico, tratamiento o prevención de una enfermedad, trastorno, síndrome, u otra afección de salud de un sujeto. Los productos biológicos, que son sustancias producidas por mecanismos biológicos relacionados con la tecnología del ADN recombinante, también se incluyen en el término "fármaco". Los fármacos incluyen, por ejemplo, antibacterianos, antiinflamatorios, anticoagulantes, antivirales, antihipertensivos, antidepresivos, antimicrobianos, analgésicos, anestésicos, beta bloqueantes, bisfosfonatos, agentes quimioterapéuticos, agentes de contraste, medicamentos para la fertilidad, alucinógenos, hormonas, narcóticos, opiáceos, sedantes, estatinas, esteroides y vasodilatadores. Ejemplos de fármacos no limitativos también se pueden encontrar en el Índice Merck. Los fármacos antibacterianos utilizados en el tratamiento de la tuberculosis, por ejemplo, incluyen isoniazida, rifampicina, pirazinamida y etambutol.

[0154] Los procedimientos que emplean un medicamento como un analito pueden comprender además el cribado de

- fármacos. Por ejemplo, la absorción de un fármaco en una célula o un organismo puede ser investigada utilizando una porina de Msp mediante la observación de bloqueos de corrientes de iones. Las zonas de constricción y/o vestíbulos específicos de una porina de Msp con diferentes tamaños, propiedades electrostáticas, y propiedades químicas pueden ser construidos para imitar estrechamente el mecanismo deseado para fármacos para entrar o salir de una célula u organismo. Estos procedimientos podrían acelerar en gran medida el cribado de fármacos, así como el diseño de fármacos. Tales estudios se han realizado con otras porinas, tal como se ha descrito por Pagel et al, J. Bacteriology 189: 8593 (2007).
- [0155]** Tal como se usa en el presente documento, un "agente de guerra biológica" se refiere a cualquier organismo o cualquier componente de origen natural, biodiseñado o sintetizado de cualquiera de dichos microorganismos capaz de causar la muerte o enfermedad en las plantas o animales (incluyendo seres humanos) o la degradación de los alimentos o suministro de agua, o la degradación del medio ambiente. Los ejemplos no limitantes incluyen virus Ébola, virus de Marburg, *Bacillus anthracis*, *Variola major*, *Variola minor*, ántrax y la ricina.
- [0156]** Tal como se usa en el presente documento, un "contaminante" se refiere a un material que contamina el aire, el agua o el suelo. Ejemplos de contaminantes no limitantes incluyen fertilizantes, pesticidas, insecticidas, detergentes, hidrocarburos de petróleo, humo, y sustancias que contienen metales pesados, tales como las que contienen zinc, cobre, o mercurio (por ejemplo, el metilmercurio).
- [0157]** Un analito puede ser un "objeto nanoscópico", que es un objeto que es menor de 100 nm en dos de sus dimensiones.
- [0158]** Las microesferas que se pueden emplear incluyen microesferas magnéticas y microesferas ópticas. Por ejemplo, se pueden usar microesferas magnéticas recubiertas con estreptavidina para aplicar una fuerza opuesta a las fuerzas electrostáticas que tiran del ADN a través del túnel de una porina de Msp. En esta última técnica una perla magnética se une al ADN biotinilado, y una fuerza comparable a la fuerza de conducción electrostática (~ 10 pN) se aplicaría usando un fuerte gradiente de campo magnético. Ver Gosse y Creoque, Biophys. J. 82: 3314 (2002). De esta manera, la lectura del bloqueo de corriente no se vería afectada, pero las fuerzas sobre el ADN podrían ser controladas independientemente. Decenas o cientos de lecturas completas independientes de cada ADN podrían entonces correlacionarse y ensamblarse para reconstruir una secuencia de ADN exacta.
- [0159]** Las microesferas ópticas manipulados por "pinzas ópticas" también se conocen en la técnica, y tales procedimientos pueden ser aplicados a las porinas de Msp descritas en el presente documento. Las pinzas ópticas son una herramienta común utilizada para ejercer una fuerza sobre un objeto nanoscópico. Un analito se une en un extremo de la microesfera, mientras que el otro extremo puede ser insertado en el túnel de la porina. La posición y la fuerza de la microesfera se controlan y se miden con las pinzas ópticas. Tales procedimientos controlan el paso del analito en el túnel y permiten un mayor control de la lectura del analito, tal como la lectura de las unidades de un polímero. Véase, por ejemplo, Trepagnier et al., Nano Lett. 7: 2824 (2007) para una descripción de tales procedimientos en el contexto de nanoporos artificiales. La patente de Estados Unidos No. 5.795.782, también describe el uso de pinzas ópticas.
- [0160]** La transferencia de energía por resonancia de fluorescencia (FRET), una técnica bien conocida, se puede emplear en los procedimientos de análisis descritos en el presente documento. Por ejemplo, una molécula aceptora FRET o una molécula donadora FRET fluorescentes pueden incorporarse en una porina de Msp. El analito se marca entonces con un emparejamiento donante FRET o aceptor FRET. Cuando el donante FRET de emparejamiento está dentro de la distancia Förster para el aceptor de FRET, es probable que se produzca la transferencia de energía. La señal resultante podría ser utilizada para fines de análisis en lugar de o además de los procedimientos que utilizan la corriente de iones como se describe en el presente documento. En consecuencia, los procedimientos de detección, identificación, o secuenciación pueden comprender la tecnología de FRET.
- [0161]** Otros procedimientos ópticos que se pueden emplear incluyen la introducción de moléculas ópticamente activas en el interior de una porina de Msp (tal como el vestíbulo o la zona de constricción). La luz externa se aplicaría para afectar el interior de la proteína: tales procedimientos podrían usarse para afectar a la velocidad de traslado de un analito o podrían permitir la entrada o salida del analito del túnel, ofreciendo el paso controlado del analito. Alternativamente, los pulsos ópticos enfocados sobre el poro se podrían utilizar para calentar el poro para afectar la forma en que interactúa con el analito. Dicho control podría ser muy rápido ya que el calor de un pequeño volumen de un centro de coordinación se disiparía rápidamente. Los procedimientos de control de la velocidad de traslado de un analito pueden por lo tanto emplear dichas moléculas ópticamente activas o pulsos ópticos.
- [0162]** La manipulación de la velocidad de traslado también puede llevarse a cabo uniendo un objeto a un extremo de un analito, y el otro extremo del analito entonces interactúa con la porina de Msp. El objeto puede ser una microesfera (por ejemplo, una microesfera de poliestireno), una célula, una molécula grande, tal como estreptavidina, neutravidina, ADN, etc., o un objeto nanoscópico. El objeto podría entonces someterse a un flujo de fluido p podría ser objeto de arrastre viscoso pasiva.
- [0163]** Los "motores moleculares" son bien conocidos en la técnica y se refieren a una molécula (por ejemplo, una

enzima) que interactúa físicamente con un analito, tal como un polímero (por ejemplo, un polinucleótido), y es capaz de mover físicamente el analito con respecto a una ubicación fija, tal como el vestíbulo, zona de constricción, o túnel de una porina de Msp. Aunque no se pretende estar limitado por la teoría, los motores moleculares utilizan energía química para generar fuerza mecánica. Un motor molecular puede interactuar con cada unidad (o "mer") de un polímero de una manera secuencial. Los ejemplos no limitantes de los motores moleculares incluyen ADN polimerasas, ARN polimerasas, helicasas, ribosomas, y exonucleasas. También se conocen motores no enzimáticos, tales como motores de virus que empaquetan el ADN. Ver Smith y otros, Nature 413: 748 (2001). Una variedad de motores moleculares y las propiedades deseables de tales motores se describen en la patente de Estados Unidos N° 7.238.485.

[0164] Un motor molecular puede estar dispuesto en la cara cis o la cara trans de una porina de Msp y, opcionalmente, se puede inmovilizar, tal como se ha descrito en la patente '485. Los procedimientos de incorporación de un motor molecular en una porina de Msp pueden realizarse usando procedimientos descritos en la patente '485. Los sistemas y aparatos descritos en la patente '485 se pueden emplear con respecto a una porina de Msp descrita en el presente documento también. De hecho, cualquier realización descrita en la patente '485 se puede emplear utilizando una porina de Msp, tal como se describe en el presente documento. Los motores moleculares también se describen en, por ejemplo, Cockroft et al., J. Amer. Chem. Soc. 130: 818 (2008); Benner et al., Nature Nanotech. 2: 718 (2007); Gyrfas y col, ACS Nano 3: 1457 (2009).

[0165] Un motor molecular se emplea típicamente para regular la tasa o velocidad de traslado a la que un analito interactúa con una porina de Msp. Cualquier proteína Msp descrita en el presente documento puede comprender un motor molecular. Opcionalmente, se emplea un motor molecular para disminuir la velocidad a la que un analito entra en un túnel de porina de Msp o para disminuir la velocidad de traslado a la que un analito se traslada a través de un túnel de porina de Msp. Opcionalmente, la velocidad de traslado o la velocidad de traslado promedio es menor de 0,5 nm/μs. Opcionalmente, la velocidad de traslado o la velocidad de traslado promedio es menor de 0,05 nm/μs. Opcionalmente, la velocidad de traslado o velocidad de traslado promedio es menor de 1 nucleótido/μs. Opcionalmente, la velocidad de traslado o la velocidad de traslado promedio es menor de 0,1 de nucleótidos/μs. Opcionalmente, la velocidad de movimiento de un analito oscila desde más de 0 Hz a 2000 Hz. Aquí, la velocidad se refiere al número de subunidades (o "mers") de un polímero regular que avanza en un segundo (Hz). Opcionalmente, el intervalo es entre aproximadamente 50-1500 Hz, 100-1500 Hz, o 350 a 1500 Hz. Opcionalmente, la velocidad de movimiento es de aproximadamente, como máximo, aproximadamente, o al menos aproximadamente 25, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, o 2000 Hz, o cualquier intervalo derivable en el mismo. La velocidad puede controlarse mediante el uso de un motor molecular que se mueve un analito a una velocidad sustancialmente constante, al menos para un periodo de tiempo durante una caracterización. Además, el intervalo de velocidad de movimiento puede depender del motor molecular. Por ejemplo, para una ARN polimerasa, un intervalo puede ser de 350 a 1500 Hz; para una ADN polimerasa, un intervalo puede ser de 75 a 1500 Hz; y para ribosomas, helicasas, y exonucleasas, un intervalo puede ser 50-1500 Hz.

[0166] Las técnicas de registro y de detección que se pueden emplear en los procedimientos descritos en el presente documento. Además, las patentes de Estados Unidos Nos. 5.795.782 y 7.189.503, también describen procedimientos de registro e instrumentación que se pueden emplear con respecto a las porinas de Msp, así como procedimientos para optimizar las lecturas de conductancia. La patente de Estados Unidos No. 6.746.594, describe un soporte para películas delgadas que contienen nanoporos y procedimientos para el uso de tales soportes que pueden ser empleados con respecto a las porinas de Msp descritas en el presente documento.

[0167] Además se proporcionan vectores que comprenden cualquiera de los ácidos nucleicos descritos en el presente documento. Tal como se usa en el presente documento, un vector puede comprender moléculas de ácido nucleico que codifican un nanoporo de Msp de una cadena (por ejemplo, un dímero de Msp de una cadena o un octámero de Msp de una cadena), en el que la molécula de ácido nucleico está unida operativamente a una secuencia de control de la expresión. Los esqueletos de vectores adecuados incluyen, por ejemplo, los utilizados habitualmente en la técnica, tales como plásmidos, cromosomas artificiales, BAC, o PAC. Numerosos vectores y sistemas de expresión están disponibles comercialmente de compañías, tales como Novagen (Madison, WI), Clontech (Pal Alto, CA), Stratagene (La Jolla, CA), y Invitrogen/Life Technologies (Carlsbad, CA). Los vectores contienen típicamente una o más regiones reguladoras. Las regiones reguladoras incluyen, sin limitación, secuencias promotoras, secuencias potenciadoras, elementos de respuesta, sitios de reconocimiento de proteínas, elementos inducibles, secuencias de unión a proteínas, regiones 5' y 3' no traducidas (UTRs), sitios de inicio transcripcional, secuencias de terminación, secuencias de poliadenilación, e intrones.

[0168] En otro aspecto, se proporciona una célula cultivada que se transfecta con un vector que comprende los ácidos nucleicos descritos en el presente documento. A este respecto, una célula se transfecta con éxito con un vector cuando la maquinaria de transcripción de la célula intacta tiene acceso a la plantilla de ácido nucleico para la producción de ARNm. Los protocolos para facilitar la transfección de vectores en células son bien conocidos en la técnica.

- 5 [0169] En el presente documento se proporciona la progenie de una célula cultivada que fue transfectada de forma estable con el vector tal como se ha descrito anteriormente. Dicha progenie contendrá copias del vector sin haber sufrido el protocolo de transfección y son capaces de transcribir los ácidos nucleicos contenidos en el vector bajo el control de una secuencia de control de expresión. Las técnicas que utilizan células cultivadas transfectadas con vectores de expresión para producir cantidades de polipéptidos son bien conocidas en el sector. Véase, por ejemplo, Wang, H., et al, J. Virology 81: 12.785 (2007).
- 10 [0170] También se proporciona en el presente documento una cepa bacteriana mutante capaz de expresión de Msp inducible. La cepa bacteriana mutante comprende una delección de una MspA de tipo salvaje, una delección de una MspC de tipo salvaje, una delección de una MspD de tipo salvaje, y un vector que comprende un promotor inducible unido operativamente a una secuencia de ácido nucleico de monómero de Msp. Opcionalmente, la cepa bacteriana mutante comprende una cepa de *M. smegmatis* ML16. Opcionalmente, la secuencia de ácido nucleico de monómero de Msp codifica un monómero de Msp seleccionado del grupo que consiste en un monómero de MspA de tipo salvaje, un monómero de MspC de tipo salvaje, un monómero de MspD de tipo salvaje, y monómeros mutantes de los mismos. Opcionalmente, el promotor inducible comprende un promotor inducible por acetamida.
- 20 [0171] De manera opcional, la cepa bacteriana mutante comprende además una delección de una MspB de tipo salvaje. La cepa bacteriana mutante que comprende una delección de una MspB de tipo salvaje puede comprender además un vector con un promotor constitutivo unido operativamente a una secuencia de ácido nucleico que codifica una porina o monómero de Msp. Opcionalmente, la porina o monómero de Msp se selecciona del grupo que consiste en una MspA de tipo salvaje, una MspC de tipo salvaje, una MspD de tipo salvaje, y mutantes de las mismas. Opcionalmente, el vector comprende cualquiera de los ácidos nucleicos descritos en el presente documento.
- 25 [0172] También se proporciona un procedimiento de producción de una porina completa o parcial de Msp de una cadena. El procedimiento comprende la transformación de una cepa bacteriana mutante. La cepa mutante comprende una delección de una MspA de tipo salvaje, una MspB de tipo salvaje, una MspC de tipo salvaje, una MspD de tipo salvaje, y un vector que comprende un promotor inducible unido operativamente a una secuencia de ácido nucleico de monómero de Msp. La cepa mutante se transforma con un vector que comprende una secuencia de ácido nucleico capaz de codificar una porina de Msp de una cadena. La porina de Msp de una cadena se purifica a continuación a partir de la bacteria. Opcionalmente, la porina de Msp de una cadena comprende una porina de MspA de una cadena. Opcionalmente, el vector comprende cualquiera de los ácidos nucleicos descritos en el presente documento.
- 30 [0173] También se proporciona un procedimiento de secuenciación de ácidos nucleicos o polipéptidos utilizando una porina de Msp de una cadena. El procedimiento comprende la creación de una bicapa lipídica que comprende una primera y segunda cara, la adición de una porina de Msp purificada a la primera cara de la bicapa lipídica, la aplicación de un voltaje positivo a la segunda cara de la bicapa lipídica, el traslado de una secuencia de ácido nucleico o polipéptido experimental a través de la porina de Msp de una cadena, la comparación de la corriente de bloqueo experimental con un patrón de corriente de bloqueo, y la determinación de la secuencia experimental. Opcionalmente, la porina de Msp de una cadena comprende un monómero de MspA de tipo salvaje o un monómero mutante del mismo. Opcionalmente, el monómero de Msp comprende un monómero parálogo u homólogo de MspA seleccionado de la Tabla 1.
- 35 [0174] El uso del término "o" en las reivindicaciones se usa para significar "y/o" a menos que se indique expresamente para referirse a alternativas solamente o las alternativas son mutuamente excluyentes, aunque la divulgación apoya una definición que se refiere solo a alternativas y "y/o".
- 40 [0175] A lo largo de esta solicitud, el término "aproximadamente" se utiliza para indicar que un valor incluye la desviación estándar del error para el dispositivo o procedimiento que se emplea para determinar el valor. En cualquier realización descrita en el contexto de un valor numérico que se utiliza en conjunción con el término "aproximadamente", se contempla específicamente que el término aproximadamente puede omitirse.
- 45 [0176] Siguiendo la antigua ley de patentes, las palabras "un" y "una", cuando se utiliza junto con la palabra "comprende" en las reivindicaciones o la memoria, se refiere a uno o más, a menos que se indique específicamente.
- 50 [0177] Se dan a conocer materiales, composiciones, y componentes que se pueden utilizar para, se pueden utilizar en conjunción con, se pueden utilizar en la preparación de, o son productos de los procedimientos y composiciones descritos. Estos y otros materiales se describen en el presente documento, y se entiende que cuando se dan a conocer combinaciones, subconjuntos, interacciones, grupos, etc., de estos materiales, que aunque la referencia específica de cada una de varias combinaciones individuales y colectivas y permutaciones de estos compuestos puede no estar descrita de forma explícita, cada uno se contempla específicamente y descrita en el presente documento. Por ejemplo, si se da a conocer y se describe un procedimiento y se dan a conocer una serie de modificaciones que se pueden realizar a una serie de moléculas que incluye el procedimiento, todas y cada combinación y permutación del procedimiento, y las modificaciones que son posibles se contemplan específicamente a menos que se indique específicamente lo contrario. Del mismo modo, cualquier subconjunto o
- 55 60 65

combinación de éstas también se contemplan y se describen específicamente. Este concepto se aplica a todos los aspectos de esta descripción incluyendo, pero no limitado a, los pasos en los procedimientos que utilizan las composiciones descritas. Por lo tanto, si hay una gran variedad de pasos adicionales que se pueden realizar, se entiende que cada uno de estas etapas adicionales se puede realizar con cualquier etapa de procedimiento específicas o combinación de etapas de procedimiento de los procedimientos descritos, y que cada una de tales combinaciones o subconjuntos de combinaciones se contemplan específicamente y deben considerarse como descritas. Por tanto, se contempla que cualquier realización descrita en esta memoria se puede implementar en relación con cualquier procedimiento, compuesto, proteína, porina, péptido, polipéptido, multímero, monómero, ácido nucleico, vector, cepa, células cultivadas, sistema, o composición, etc., que se describen en el presente documento, y viceversa. Por ejemplo, cualquier proteína descrita en el presente documento se puede emplear en cualquier procedimiento descrito en el presente documento.

[0178] Los siguientes ejemplos se proporcionan con el propósito de ilustrar, no limitar, el material descrito en el presente documento.

EJEMPLOS

Ejemplo 1: Materiales y procedimientos para los Ejemplos 1-7

[0179] Se sintetizaron oligonucleótidos ssADN homogéneos dA₅₀, dC₅₀, y dT₅₀ (SEQ ID NO: 10, SEQ ID NO: 16 y SEQ ID NO: 17, respectivamente) y las construcciones de horquilla hp08 (5' GCTGTTGC TCTCTC GCAACAGC A₅₀ 3') (SEQ ID NO: 4), hp10 (5' GCTCTGTTGC TCTCTC GCAACAGAGC A₅₀ 3') (SEQ ID NO: 5), y HP12 (5' GCTGTCTGTTGC TCTCTC GCAACAGACAGC A₅₀ 3') (SEQ ID NO: 6) por Integrated DNA Technologies, (IDT; Coralville, IA).

[0180] Cepas bacterianas y condiciones de crecimiento. Todas las cepas bacterianas utilizadas en este estudio se enumeran en la Tabla 5. Las micobacterias se cultivaron a 37°C en un medio líquido Middlebrook 7H9 (Difco) suplementado con 0,2% de glicerol, 0,05% de Tween 80® o sobre agar Middlebrook 7H10 (Difco) suplementado con 0,2% de glicerol. Se utilizó *Escherichia coli* DH5α para todos los experimentos de clonación y se cultivó rutinariamente en medio Luria-Bertani (LB) a 37°C. Se utilizó higromicina en concentraciones de 200 µg/ml para *E. coli* y 50 µg/ml para *M. smegmatis*.

Tabla 5. Cepas y plásmidos

Cepa/plásmido	Cepa parental y genotipo pertinente
Cepa	
<i>E. coli</i> DH5α	<i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi</i> , <i>relA1</i> , <i>hsdR17</i> (rK, mK ⁺), <i>supE44</i> , ϕ 80Δ <i>lacZ</i> Δ <i>M15</i> , Δ <i>lacZ</i> (YA-argF)UE169
<i>M. smegmatis</i> ML16	ML15, Δ <i>mspA</i> ::FRT, Δ <i>mspC</i> ::FRT, Δ <i>mspD</i> ::FRT, attB::loxP, FRT
Plásmido	
pMS2	origen CIE1, origen PAL5000 origin, Hyg ^R
pMN016	P _{smyc} - <i>mspA</i> , origen ColE1, origen PAL5000, Hyg ^R
pMN035	P _{smyc} -rv1698, origen ColE1, origen PAL5000, Hyg ^R
pML904	derivado pMN016, <i>mspA</i> D90N/D91N/D93N (<i>m1mspA</i>)
pML840	derivado pML904, <i>mspA</i> D90N/D91N/D93N/D118R
pML841	derivado pML840, <i>mspA</i> D90N/D91N/D93N/D118R/E139R
pML843	derivado pML840, <i>mspA</i> D90N/D91N/D93N/D118R/E139K
pML844	derivado pML843 <i>mspA</i> , D90N/D91N/D93N/D118R/E139K/D134R (<i>m2mspA</i>)

La anotación HygR indica resistencia a higromicina. *MspA*, *MspC* y *MspD* son genes de porinas de *M. smegmatis*.

[0181] Mutagénesis dirigida de sitio de *MspA*. Los monómeros mutantes M1MspA y M2MspA se construyeron de una manera gradual por mutagénesis dirigida de sitio utilizando la reacción en cadena combinada (CCR) como se describe por Bi y Stambrook, Nucl. Acids Res. 25: 2949 (1997). El plásmido pMN016 porta una fusión transcripcional P_{smyc}-*mspA* (Stephan et al, Mol Microbiol 58: 714 (2005)) y se utilizó como una plantilla. Se utilizaron los oligonucleótidos psmyc1 y pMS-seq1 como cebadores directo e inverso, respectivamente, y un cebador de mutagénesis apropiado (Tabla 6) en el CCR. Se introdujeron tres mutaciones posteriores en *mspA* para construir el gen *m1mspA*. Otras tres mutaciones se introdujeron en *m1mspA* para producir *m2mspA*. Todos los plásmidos se verificaron mediante secuenciación de todo el gen de *mspA* antes de que se transformaron en el triple mutante de porina de *M. smegmatis* ML16 (Stephan et al, Mol Microbiol 58: 714 (2005)) para la producción de proteínas.

Tabla 6. Oligonucleótidos.

Oligonucleótido	Secuencia (dirección 5' a 3')	Objetivo
Psmyc1	CGACCAGCACGGCATAACATC (SEQ ID NO:41)	Amplificación y secuenciación
pMS-SEQ1	CGTTCTCGGCTCGATGATCC (SEQ ID NO:42)	Amplificación y secuenciación
MspA909193NFP	CCTGATCAACAACGGTAACATCACCGC (SEQ ID NO:43)	Clonación de pML904
MspA_118R	CTGGGCACGCCTGGGCAACGG (SEQ ID NO:44)	Clonación de pML840
MspA_139R	TCCGGCGCCCGCGGTGGCGTG (SEQ ID NO:45)	Clonación de pML841
MspA_139K	GGCGCCAAGGGTGGCGTGG (SEQ ID NO:46)	Clonación de pML843
MspA_134R	CGTTCTCGGTCCGCGTCTCC (SEQ ID NO:47)	Clonación de pML844

Los codones que se han alterado para introducir las mutaciones MspA están subrayados.

5 **[0182]** Experimentos de túnel individuales. Se produjeron bicapas lipídicas con difitanoi-PA y difitanoi-PC preparados en proporción igual o desigual y se formaron a través de una apertura horizontal, de diámetro de ~ 20 μm , en Teflon tal como se ha descrito (Akeson et al, Biophys J. 77: 3227 (1999)). Se añadieron porinas de MspA a una cara de la bicapa (cara cis) a una concentración de ~ 2,5 ng/mL. La cara cis se puso en tierra, y se aplicó una tensión positiva a la cara trans de la bicapa. Se empleó un amplificador patch-clamp Axopatch-1B (Axon Instruments) para aplicar tensión a través de la bicapa y medir la corriente iónica que fluye a través del poro. La señal analógica se filtró a paso bajo a 50 kHz con un filtro Bessel de 4 polos. La señal filtrada amplificada se digitalizó a 250 kHz. La adquisición de datos se controló con software personalizado desarrollado en LabWindows/CVI (National Instruments). Todos los experimentos se realizaron a $21 \pm 2^\circ\text{C}$ en 1 M KCl, 10 mM Hepes/KOH tamponado a pH 8.

10
15
20
25 **[0183]** Análisis de datos. El análisis de datos se llevó a cabo con el software personalizado desarrollado en Matlab (The MathWorks, Natick, MA). Los bloqueos se identificaron como los intervalos donde la corriente iónica caía por debajo de un umbral de 80% del nivel de corriente no bloqueada, se mantuvo allí durante al menos 12 μs , y volvió de forma espontánea al nivel no bloqueado. Los bloqueos se rechazaron si la señal no bloqueada inmediatamente anterior o siguiente tenía una corriente media que se desviaba del nivel no bloqueado típico en más de dos veces el ruido rms de la señal no bloqueada. Los bloqueos también fueron rechazados si se producían en menos de 26 μs de otro bloqueo. Los bloqueos profundos fueron identificados como intervalos en los que la corriente iónica caía <50% del nivel no bloqueado. Los intervalos donde la corriente se mantuvo entre 80% y 50% del nivel no bloqueado se identificaron como bloqueos parciales. Cada suceso se parametrizó por los tiempos de permanencia y las corrientes promedio de sus subintervalos parciales y profundos constituyentes.

30 **[0184]** Los valores de t_D utilizados para parametrizar las distribuciones del tiempo de permanencia del bloqueo profundo de horquilla se estimaron como el pico de la distribución de densidad de probabilidad del \log_{10} de los tiempos de permanencia (FIGURA 8). Esta distribución se estimó con el estimador de densidad de Kernel suavizado Matlab utilizando una función kernel normal y una anchura de 0,15. Los datos de bicapa trans se analizaron mediante la detección de cambios abruptos en la conductancia de menos de 1 ns a más de 1 ns. La tensión a la que se produjeron estos cambios se registró y se resumió en los histogramas que se muestran en las FIGURAS 9E-9G.

35 **[0185]** En todos los experimentos, los poros estaban orientados de tal manera que la "entrada" (FIGURA 1) fue expuesta al compartimento cis del aparato.

40 **[0186]** Todos los datos de horquilla mostrados en las FIGURAS 5-8 se derivaron de los datos tomados en la misma porina de larga vida M1MspA. Los datos de homopolímero presentados en el Ejemplo 5 se obtuvieron con diferentes porinas de larga vida M1MspA que los datos de la horquilla, aunque se considera una concordancia cuantitativa entre amplios conjuntos de datos de horquilla tomados en los dos poros.

Ejemplo 2: Características de bloqueo de porinas de MspA de tipo salvaje (WTMspA) con y sin analito

45 **[0187]** Purificación de porinas MspA. Las porinas de MspA se extrajeron selectivamente de *M. smegmatis* y se purificaron por el posterior intercambio aniónico y cromatografía de filtración en gel tal como se ha descrito (Heinz y Niederweis, Anal Biochem 285: 113 (2000); Heinz et al, Methods Mol Biol 228: 139 (2003)).

50 **[0188]** De acuerdo con los resultados anteriores (Niederweis et al, Mol Microbiol 33: 933 (1999)), la proteína purificada demostró una alta actividad de formación de túnel con una conductancia más frecuente de 4,9 nS en 1,0 M KCl a ~ 20°C (FIGURA 2). El compartimento cis se mantuvo en tierra y de tensión positiva se aplicó al compartimento trans (FIGURA 3A). Por encima de ~ 60 mV, la porina WTMspA demostró bloqueos frecuentes, espontáneos, de la corriente iónica en ausencia de ssADN (FIGURA 3). Algunos bloqueos espontáneos fueron

transitorios, y otros requirieron la inversión de la tensión para restablecer el nivel de corriente no bloqueado. A pesar de este comportamiento, quedaron intervalos de señal constante y sin obstrucciones durante decenas de segundos de duración para tensiones de hasta ~ 100 mV (FIGURA 3). La adición de $\sim 2\text{-}8 \mu\text{M}$ dC₅₀ (SEQ ID NO: 48) ssADN al compartimento cis no condujo a una mejora notable o alteración de estas características del bloqueo. Por encima de ~ 100 mV los bloqueos espontáneos fueron tan frecuentes que los experimentos de detección de ssADN eran poco prácticos.

[0189] Una explicación para la aparente ausencia de interacciones de ssADN con la porina WTMspA es la alta densidad de carga negativa en el poro (FIGURA 1). La interacción electrostática con el interior del túnel cargado negativamente probablemente inhibe la entrada de ADN en el poro. Para abordar esta cuestión se sustituyeron residuos de aspartato en la zona de constricción por asparaginas (FIGURA 1). La porina de MspA mutante D90N/D91N/D93N (M1MspA) resultante se discute en el Ejemplo 3.

Ejemplo 3: Características del bloqueo de la porina de MspA mutante M1MspA con y sin analito

Experimental.

[0190] Como se ha indicado en el Ejemplo 2, las interacciones electrostáticas entre ssADNA y el túnel de la porina WTMspA pueden afectar el traslado de ssADN a través del poro. La MspA mutante D90N/D91N/D93N (M1MspA, también conocida como M1-NNN) fue diseñada para probar esta teoría. La porina M1MspA se expresó y se purificó a partir del *M. smegmatis* cepa ML16 que carece de porinas más endógenas (Stephan et al, Mol Microbiol 58: 714 (2005)). Los niveles de expresión de la porina M1MspA (FIGURA 4) y su actividad de formación de túnel fueron similares a la porina WTMspA, mientras que la conductancia se redujo en un factor de 2 a 3 (FIGURA 2). Además, la frecuencia de bloqueos espontáneos se redujo drásticamente en la porina M1MspA, por lo que es posible llevar a cabo experimentos de detección de ADN en tensiones de hasta y por encima de 180 mV (FIGURA 5).

[0191] Se utilizaron construcciones horquilla de ssADN para investigar la interacción del ADN con la porina M1MspA. Cada construcción tenía un poli-dA de 50-nt que colgaba del extremo 3', una región dúplex dsADN de longitud variable (8, 10, y 12 pb para las construcciones hp08 (SEQ ID NO: 4), HP10 (SEQ ID NO: 5), y hop12 (SEQ ID NO: 6), respectivamente), y un bucle de 6-nt (FIGURA 6). A 180 mV, la adición de $\sim 8 \mu\text{M}$ de ssADN hp08 al compartimento cis causó el aumento de la velocidad de bloqueos corriente iónica transitoria de 0,1-0,6 bloqueos por segundo a 20-50 bloqueos por segundo (FIGURA 5). La velocidades de bloqueo fueron proporcionales a la concentración de ADN y fueron fuertemente dependientes de la tensión, disminuyendo ~ 3 veces para una disminución de 20 mV en la tensión aplicada. Los bloqueos largos para estar bien separados eran bloqueos parciales, donde la corriente iónica se redujo a entre 80% y 50% del nivel no bloqueado o bloqueos profundos donde la corriente iónica se redujo a menos del 50% del nivel no bloqueado (FIGURA 5C). Los bloqueos que exhiben subsegmentos tanto parciales como profundos eran muy raros. Los bloqueos parciales duraron decenas a cientos de microsegundos y sus tiempos de permanencia aumentaron con el aumento de tensión (FIGURAS 5C y 7). Los bloqueos profundos duraron de cientos de microsegundos a cientos de milisegundos y sus tiempos de permanencia disminuyeron con el aumento de tensión (FIGURAS 6 y 7). No se observaron estas tendencias en los experimentos con las tres horquillas.

Análisis.

[0192] En analogía con señales similares observadas con αHL (Butler et al, Biophys J. 93: 3229-40 (2007)), los bloqueos parciales se interpretan como la entrada de ADN en el vestíbulo de porinas M1MspA sin enroscado del segmento de una cadena a través de la constricción del túnel. Para este mecanismo, se espera una reducción moderada de la corriente iónica. Sin querer limitarse por la teoría, el aumento en el tiempo de permanencia con el voltaje (FIGURA 7) muy probablemente es el resultado de un aumento de barrera electrostática frente al escape de una molécula de ADN desde el vestíbulo de nuevo al compartimento cis. Esta explicación para el aumento del tiempo de permanencia se puede entender dentro de un marco cinético donde la descomposición del polímero del vestíbulo se produce a través de los dos procesos de primer orden de escape contra el gradiente de voltaje aplicado y el enroscado de un extremo a través de la constricción. La vida útil es entonces el inverso de la suma de las constantes de velocidad para estos procesos. Este tiempo de vida se incrementará con el voltaje si (i) la constante de velocidad de escape disminuye con la tensión y (ii) su disminución domina cualquier cambio en la constante de velocidad de enroscado.

[0193] Para los bloqueos profundos, la clara disminución de tiempos de permanencia con el aumento de la tensión es inconsistente con cualquier proceso que implique un escape del ADN de nuevo al compartimento cis. Tanto el grado de reducción de la corriente iónica como la dependencia del voltaje de los tiempos de permanencia son consistentes con un proceso en el que el segmento de polidA monocatenario es impulsado a través de la constricción de ~ 1 nm de diámetro hasta que el dúplex de ADN de $\sim 2,2$ nm de diámetro alcanza la constricción y detiene el traslado (FIGURA 5A). La construcción de horquilla se mantiene en esta configuración de rosca hasta que cualquier de descomposición del dúplex de ADN (Vercoutere et al, Nat Biotech 19: 248-52 (2001); Sauer-Budge et al, Phys Rev. Lett 90: 238 101 (2003); Mathe et al, Biophys J. 87: 3205-12 (2004)) o un reordenamiento conformacional de la zona de constricción de la porina M1MspA permite que se complete el traslado. Sin estar ligado

por la teoría, el mecanismo de descomposición de finalización de el traslado parece más plausible porque el paso de una hélice de ADN de doble cadena requeriría una constricción con aproximadamente doble de diámetro, lo que altera los enlaces de hidrógeno del barril betal que flanquea la constricción (Faller et al., Science 303: 1189 (2004)) y, potencialmente se exponen las regiones hidrófobas de la proteína y la bicapa interior al agua.

[0194] Los bloqueos profundos de horquilla en la porina de M1MspA tenían distribuciones muy amplias del tiempo de permanencia que no fueron bien descritos por exponenciales simples o sumas de exponenciales (FIGURA 8). Para parametrizar las distribuciones, se utilizó el modo del logaritmo de los tiempos de permanencia de bloqueo profundo, t_D , que corresponde en la FIGURA 6 al tiempo de permanencia con la mayor densidad de bloqueos (FIGURA 8). Para todas las tensiones, hp08 tenía el t_D más corto. Por debajo de 160 mV, hp10 y hp12 tenían un t_D similar. Sin embargo, por encima de 160 mV hp10 tenía un t_D consistentemente más largo que hp12. Estas observaciones son algo diferentes que las de α HL, donde las distribuciones del tiempo de permanencia de bloqueo de horquilla se modelaron con exponenciales individuales y horquillas con mayores energías libres estándar de formación producían consistentemente bloqueos profundos más largos (Vercoutere y otros, Nat Biotechnol 19: 248 (2001); Mathe et al, Biophys J. 87: 3205 (2004)). Suponiendo que los bloqueos profundos son producidos por el traslado con disociación del dúplex como el paso limitante de la velocidad, entonces este proceso es 10-100 veces más lento en la porina M1MspA que en α HL (Mathe et al, Biophys J. 87:.. 3205 (2004)). Curiosamente, los bloqueos de hp10 persistieron más que los bloqueos de hp12. En seis experimentos repetidos con hp10 a 180 mV, se observaron un nivel de corriente no bloqueada promedio de 340 ± 7 pA y un t_D promedio de 9 ± 1 ms (media \pm SEM).

Ejemplo 4: Detección transbicapa con la porina de M1MspA

Teoría.

[0195] Para obtener una prueba directa de que el ADN se traslada a través de MspA, se empleó la técnica de detección de transbicapa ilustrada en la FIGURA 9 y por primera vez por Nakane et al. (Nakane et al, Biophys J. 87: 615 (2004)). Una molécula de sonda de ssADN con un complejo de anclaje voluminoso en un extremo es impulsado por electroforesis en el nanoporo. Los extremos de ssADN libres se enroscan a través de los poros en el compartimiento trans hasta que el ancla detiene el traslado. Si el compartimiento trans contiene moléculas cortas diana de ssADN que son complementarias al extremo de la sonda de ssADN, entonces la sonda y la diana se pueden hibridar. Si se produce la hibridación, la sonda es bloqueada en una configuración de enroscado hasta que la aplicación de un voltaje suficientemente negativa hace que la sonda se disocie de la diana y sale del compartimiento cis. Si la hibridación no se produce por razones estocásticas o porque el extremo de la sonda no es complementario a la diana, o si no hay moléculas diana en el compartimiento trans, entonces no es necesario un voltaje negativo para la sonda para salir de nuevo al compartimiento cis. La aparición de bloqueos que sólo se eliminan por voltaje suficientemente negativo es una evidencia de que la sonda de ssADN se ha enroscado a través del nanoporo al compartimiento trans y se ha hibridado con el ADN diana.

Experimental.

[0196] Las moléculas sonda se construyeron comprendiendo moléculas ssADN de 75-nt de largo que fueron unidos a un anclaje de neutravidina (nA) en su extremo 5' biotinilado y tenían una secuencia complementaria de 15 nt de longitud heterogénea en su extremo 3'. nA se obtuvo de Invitrogen (Carlsbad, CA). Se sintetizaron dos construcciones de ssADN biotinilados en 5' diferentes, 5'-bt-dC₆dA₅₄ d (CTCTATTCTTATCTC)-3' (SEQ ID NO: 7) y 5'-bt-dC₆dA₅₄ d(CACACACACACACAC)-3' (SEQ ID NO: 8), por IDT. nA y las construcciones de ssADN se mezclaron a una concentración de 50 μ M en una proporción 1:1 en el tampón experimental 1M KCl y se almacenaron a -20°C hasta inmediatamente antes de su uso. El ADN diana de 15 nt de longitud, 3'-GAGATAAGAATAGAG-5' (SEQ ID NO: 9) se sintetizó por IDT, se suspendió en el tampón experimental, y se almacenó a -20°C hasta inmediatamente antes de su uso. El compartimiento trans se precargó con ~ 100 μ M de ADN diana y el compartimiento cis se llenó con tampón exento de ADN. Después de formar una bicapa, el compartimiento cis se perfundió para eliminar cualquier ADN diana que se difunde a través de la abertura. Una vez se estableció una porina de M1MspA estable, se añadieron los complejos nA-ssADN al compartimiento trans hasta una concentración final de ~ 1 μ M. Se utilizó un software de control experimental personalizado desarrollado en LabWindows para controlar continuamente la corriente y aplicar los voltajes apropiados.

[0197] Se observaron bloqueos de corriente profundos indefinidos cuando las moléculas de la sonda fueron conducidas al poro del compartimiento cis con 180 mV. Para los experimentos de transbicapa, moléculas de la sonda fueron capturadas con 180 mV. Después de un breve retardo para asegurar que el ssADN se había enroscado lo máximo posible a través de la porina de M1MspA, el voltaje se redujo a 40 mV y se mantuvo a ese nivel durante 5 s para permitir que uno de los ssADN diana de 15 nt de longitud se hibride al extremo complementario de la sonda. A continuación, la tensión se redujo en rampa hasta una velocidad de 130 mV/s. Para cada suceso, el voltaje de salida de la sonda, V_{salida} , fue identificado como la tensión a la que se observó un aumento grande y abrupto en la conductancia con la rampa (FIGURAS 9C y 9D).

[0198] Los datos transbicapa se analizaron mediante la detección de cambios abruptos en la conductancia de <1 a $>$

1 ns. La tensión a la que se produjeron estos cambios se registró y se resumió en los histogramas que se muestran en las FIGURAS 9E-9G. Ver Materiales y Procedimientos en el Ejemplo 1 para obtener más información sobre el análisis de datos.

5 Análisis.

[0199] Los histogramas de V_{salida} de experimentos con tres diferentes combinaciones de sonda/diana se muestran en la FIGURA 9. Cuando el ADN de sonda es complementario al ADN diana (FIGURA 9E) un número significativo de V_{salida} son negativos, lo que indica hibridación de la sonda/diana. En seis experimentos repetidos con moléculas de sonda/diana complementarias, se observaron poblaciones similares de V_{salida} negativo. En cinco experimentos repetidos donde el extremo 3' de ssADN no era complementario a las moléculas diana (FIGURA 9F) y en un experimento sin ADN diana (FIGURA 9G), los valores de V_{salida} negativos rara vez se observaron. En dos diferentes nanoporos se utilizaron las combinaciones de sonda/diana complementarias y no complementarias. Los datos de uno de esos poros se muestran en las FIGURAS 10E y 10F. Estos datos proporcionan una evidencia clara y directa de que ssADN puede enroscarse a través de la porina M1MspA, lo que confirma la hipótesis de que los bloqueos profundos observados en la FIGURA 5 están en efecto causados por el traslado de ssADN a través de la porina de M1MspA.

Ejemplo 5: La porina de MspA mutante M1MspA y ssADN lineal, homogéneo

[0200] También se investigó la interacción entre la porina de M1MspA y 50 unidades de ssADN homogéneas lineales. A 180 mV, la adición de $\sim 8 \mu\text{M}$ de dT50 en el compartimento cis causó ~ 5 bloqueos por segundo (FIGURA 10), un aumento en un factor de ~ 20 sobre la velocidad de bloqueo en ausencia de dT50 (SEQ ID NO: 32). La mayoría de estos bloqueos eran más cortos que $30 \mu\text{s}$, que es demasiado breve para resolver la estructura interna o estimar la profundidad del bloqueo. Los experimentos con dA_0 (SEQ ID NO: 49) y dC_{50} (SEQ ID NO: 48) dieron resultados similares. La corta duración de los bloqueos observados sugieren que el traslado de estas 50 unidades de ssADN homogéneas lineales es típicamente inferior a $30 \mu\text{s}$. Los bloqueos son también coherentes con las pequeñas desviaciones de los polímeros en el vestíbulo que terminan con el escape de nuevo en el compartimento cis. Aunque tanto el traslado como el escape se producen probablemente en los experimentos con 50 unidades de ssADN lineales, las estimaciones de la frecuencia relativa de los dos procesos no eran posibles.

Ejemplo 6: Características de bloqueo de la porina de MspA mutante M2MspA con y sin analito

[0201] A fin de examinar el efecto de las cargas en la porina de MspA en sus capacidades de análisis de ADN, se realizaron tres mutaciones adicionales a la porina M1MspA y se sustituyeron residuos cargados negativamente en el vestíbulo y alrededor de la entrada por residuos de carga positiva (FIGURA 1). La porina mutante resultante D90N/D91N/D93N/D118R/D134R/E139K (M2MspA) demostró niveles de expresión (FIGURA 4) y la actividad de formación de túnel similares a la porina WTMspA (FIGURA 2).

[0202] Al igual que la porina de M1MspA, la porina de M2MspA presentaba una conductancia más pequeña que la porina WTMspA (FIGURA 2) y exhibió bloqueos espontáneos mínimos para tensiones de hasta y por encima de 180 mV. A 180 mV, la adición de $2 \mu\text{M}$ dT50 (SEQ ID NO: 32) al compartimento cis condujo a velocidades de bloqueo de ~ 25 bloqueos por segundo (FIGURA 10B.). Un bloqueo parcial de $\sim 100 \mu\text{s}$ terminando con un pico claro descendente era un patrón común de bloqueo (FIGURA 10C). Las duraciones de bloqueo parcial y su tendencia a terminar con un pico descendente aumentaron la tensión (FIGURA 11). Estas tendencias son consistentes con un proceso en el que un polímero entra en el vestíbulo y se mantiene allí, produciendo un bloqueo parcial hasta que un extremo entra en la constricción de campo alto y se inicia el traslado. Este mecanismo ha explicado con precisión un patrón de bloqueo parcial a profundo similar al observado con αHL (Butler et al, Biophys J. 93: 3229 (2007)). La corta duración de los picos descendentes sugiere que el traslado de 50 unidades de ssADN lineales a través de la porina de M2MspA es más corta que $\sim 30 \mu\text{s}$. Los bloqueos parciales que no terminan con picos descendentes se interpretan como que escapan de nuevo al compartimento cis o como traslado que es más corta que $\sim 10 \mu\text{s}$, lo cual es demasiado breve para ser observado en estos experimentos.

Ejemplo 7: Comparación de las porinas mutantes de MspA M1 y M2 y todas las propiedades

[0203] Una similitud importante entre las porinas M1MspA y M2MspA es que el traslado de 50 unidades de ssADN lineales parece ser demasiado rápida para producir bloqueos profundos con estructura resoluble. Sin estar ligado por la teoría, esta observación sugiere que la constricción, que es la misma para ambos mutantes, es la región que determina principalmente la velocidad de una molécula de ADN monocatenario lineal que se traslada a través de la porina de MspA. La comparación de velocidades de traslado de MspA de ~ 2 -10 bases/ μs de las porinas M1MspA y M2MspA con las velocidades de traslado de $\sim 0,5$ -1 bases/ μs observadas con αHL (Meller et al., Proc. Natl Acad. Sci. USA. 97: 1079 (2000); Butler et al, Biophys J. 93: 3229 (2007)) apoya la noción de que los detalles de la geometría del túnel y la composición juegan un papel principal en la determinación de la velocidad de traslado.

[0204] En el caso de la porina de MspA y αHL , la gran diferencia en la velocidad de traslado podría resultar de la

anchura de las regiones de túnel que flanquean las constricciones. Si la interacción entre el ADN y las paredes del túnel ralentiza el paso de ADN (Slonkina y Kolomeisky, J. Phys Chem. 118: 7112-8 (2003)), entonces se esperaría un traslado más lento en α HL donde las 10-20 bases que están altamente confinadas en la región de constricción y transmembrana se ven obligadas a interactuar con las paredes del túnel. En la porina de MspA, sólo 2-4 bases en la constricción se ven obligadas a estar en contacto con la proteína. La distribución de carga dentro de la constricción es otra diferencia significativa entre α HL alfa y las porinas mutantes de MspA M1 y M2. La constricción α HL está formada por las cadenas laterales de E111, K147, y M113 (Song et al, Science 274: 1859 (1996)), que fuerzan la estructura de ssADN de carga negativa en muy estrecha proximidad con siete residuos cargados positivamente y siete cargados negativamente. La falta de residuos cargados en la constricción de las porinas de MspA mutantes M1 y M2 también podría ser responsable de las velocidades de traslado más rápidas en comparación con α HL.

[0205] Una comparación adicional de las características de bloqueo de homopolímero entre las dos porinas de MspA mutantes da una idea de cómo la disposición de los residuos cargados en el túnel influye en sus interacciones con el ADN. Las velocidades de bloqueo de la porina M2MspA eran ~ 20 veces más altas que las velocidades de la porina M1MspA para una concentración dada de ssADN (FIGURA 10B). La porina M2MspA también demostró bloqueos fácilmente observables hasta ~ 80 mV, mientras que casi no hubo bloqueos visibles para la porina M1MspA por debajo de ~ 140 mV. Por último, los bloqueos parciales de la porina M2MspA eran al menos ~ 100 veces más largos que para la porina M1MspA (FIGURA 9C). Estas tendencias son consistentes con un modelo electrostático sencillo en el que los residuos cargados positivamente en las porinas M2MspA facilitan la entrada de ssADN en el vestíbulo e inhiben el escape de las moléculas de ssADN del vestíbulo de nuevo al compartimento cis. Estas observaciones demuestran que la colocación adecuada de los residuos cargados ofrece un medio sencillo para adaptar sustancialmente la interacción entre la porina de MspA y el ADN.

Ejemplo 8: La porina M1MspA reconoce un nucleótido único en un ADN mantenido en el poro por una sección de horquilla (hp)

[0206] Los experimentos con la porina M1MspA y (i) una cadena de ADN poli-A con una sola C incrustado dentro y (ii) una sola T incrustada en un fondo de poli-A se desarrollaron como se describe en el Ejemplo 3. Como se señaló anteriormente, la horquilla mantiene la construcción de ADN en la zona de constricción de la porina de MspA durante el tiempo suficiente para obtener las firmas de corriente muy bien definidas.

[0207] Una sola C incrustada en una construcción de horquilla de ADN poli-A. La FIGURA 12A muestra el histograma de corriente debido a una única C en la posición 1, 2, y 3 después de la horquilla, así como una mezcla de poli-A y poli-C. Los histogramas de corriente para cada sitio son muy diferentes y muestran que el "sitio de reconocimiento" está cerca de la posición 2. Para una descripción más cuantitativa, el pico de las distribuciones de corriente se escaló mediante la diferencia de corriente encontrada para poli-C y poli-A (FIGURA 12B). Un ajuste de Gauss revela que la posición de reconocimiento de la porina de MspA para una sola C es de 1,7 nucleótidos (nt) de distancia de donde se encuentra la horquilla. La longitud del sitio de reconocimiento (longitud de la zona de constricción) es comparable a la anchura gaussiana (1,6 nt) ~ 5-6 Å de largo.

[0208] Una sola T incrustada en una construcción de horquilla de ADN poli-A. Los experimentos utilizando una única T en ADN poli-A se llevaron a cabo en una manera similar, centrándose únicamente en las tres primeras posiciones adyacentes a la horquilla (FIGURA 13, paneles 2-4). La especificidad es igualmente impresionante, pero en este caso muestra la sensibilidad más grande cerca de la posición 1. La ubicación de la única T se puede resolver mucho mejor que una posición. Sin estar ligado por la teoría, los inventores especulan que la diferencia en el reconocimiento de la posición en comparación con una C en poli-A es de hecho causada por el propio ADN que contribuye con el medio electrostático que forma la constricción. Los datos con una única A en un fondo de C se muestra en los tres paneles inferiores de la FIGURA 13. Si bien la A sola produce firmas de bloqueo de corriente que sólo están separados débilmente del fondo de poli-C, las distribuciones de corriente son lo suficientemente estrechas como para separar la A sola. La posición óptima de A en la cadena de poli-C parece estar cerca de la posición 2, es decir, similar a una C sola en una cadena A.

[0209] La composición de la cola de ADN más allá de la posición 3 no afecta a las propiedades de reconocimiento de bases. El ADN poli-A forma estructura secundaria, y las diferencias entre los datos de fondo de C en poli A y fondo de A en poli C podrían ser debidas a la interrupción de la estructura secundaria (rigidez) de la cola de poli-A. Las mediciones se llevaron a cabo con una secuencia heterogénea 47 bases después de las tres primeras posiciones ocupadas por trinucleótidos A o C. Se encontraron que los niveles de corriente eran indistinguibles de los niveles de corriente de colas A₅₀ y C₅₀, lo que indica que la estructura secundaria o composición de la cola no afecta el bloqueo de corriente (FIGURA 14).

[0210] Otra serie de experimentos se llevaron a cabo (1) para evaluar la capacidad de la porina M1MspA de distinguir diferentes nucleótidos y (2) evaluar la ubicación y longitud de la región a la que es sensible la porina (resolución espacial). En estos experimentos, se utilizaron diversas construcciones de ADN con una cadena de 50 nucleótidos de ssADN unida a una sección de horquilla de 14 pares de bases para prevenir el traslado inmediato. Los datos se resumen en la FIGURA 31. dA₅₀ (SEQ ID NO: 49) y dC₅₀ (SEQ ID NO: 48) produjeron corrientes de bloqueo significativamente diferentes. A continuación, se ensayaron una serie de construcciones, y el sitio de

reconocimiento se aisló de las primeros cuatro bases siguientes a la horquilla. Estas construcciones tenían secuencias de ssADN de dC₄dA₄₆ (SEQ ID NO: 15), dA₃dC₄dA₄₃ (SEQ ID NO: 12), y dA₆dC₄dA₄₀ (SEQ ID NO: 11) después de la horquilla. dC₄dA₄₆ muestra una distribución de corriente bloqueo casi idéntica a dC₅₀, mientras que dA₃dC₄dA₄₃ y dA₆dC₄dA₄₀ bloquean como dA₅₀. Esto estrechó el sitio de reconocimiento para estar con los primeros 3 nucleótidos después de la horquilla. A continuación, las construcciones se ensayaron con un solo dC en varias posiciones en un fondo de poli-dA. Hp-dC₁dA₄₉ (dC en la posición 1) (SEQ ID NO: 14) bloqueó la corriente a un nivel intermedio entre los valores de poli-dA y poli-dC. La construcción dA₂dC₃dA₄₇ (dC en la posición 3) (SEQ ID NO: 50) bloqueó la corriente intermedia entre poli-dA y poli-dC, pero cercana a poli-dA. Poli-dT₅₀ (SEQ ID NO: 32) bloqueó con la corriente más pequeña, y hp-dG₃dA₄₇ (SEQ ID NO: 18) produce una corriente intermedia entre poli-dC y poli-dA. En un mutante diferente (D90/91Q + D93N + D118/134R + E139K), se midieron las corrientes de bloqueo para poli-dC, poli-dA, y poli-dT y eran distinguibles entre sí. Estos datos demuestran que la porina M1MspA tiene capacidades de reconocimiento y que el sitio de reconocimiento es corto. Además, el sitio de reconocimiento parece estar ubicado en la zona de constricción, suponiendo que la horquilla se detiene a la derecha de la cara cis de la zona de constricción.

Ejemplo 9 Construcción y caracterización de porinas de MspA mutantes MspA M1-QQN y M2-QQN

[0211] En otra serie de experimentos diseñados para reducir el traslado del ADN a través del túnel de porina de MspA, se realizaron dos mutantes adicionales. Una, llamada M1-QQN, se hizo de una manera similar a M1-NNN (o M1MspA) anterior mediante la sustitución de los aminoácidos en las posiciones 90 y 91 del monómero de MspA de tipo salvaje por glutamina y el aminoácido en la posición 90 por asparagina. Con M2-QQN, el tamaño de poro de constricción se redujo mediante la introducción de glutamina más voluminosa en las posiciones 90 y 91 en el fondo del mutante M2MspA (véase el Ejemplo 6; D90Q + D91Q + D93N + D118R + E139K + D134R). Se expresó en el mutante de *M. smegmatis* ML16 descrito en los Ejemplos 1 y 3 anteriores. La cantidad de la porina M2-QQN en extractos de detergente fue tan alta como la de la porina WTMspA (FIGURA 15A), indicando que las nuevas mutaciones no afectaban a la expresión de los poros. Los experimentos de bicapa lipídica mostraron que la porina M2-QQN forma poros abiertos estables como la porina WTMspA (FIGURA 15B). La actividad de formación de poros es similar a la de la porina WTMspA. La conductancia de un solo túnel de porina M2-QQN (2,4 nS) fue mayor que la de su origen M2 (1,4 ns).

[0212] Los mutantes QQN también distinguen entre las bases A, C, y T. Cualitativamente similares a las porinas mutantes M1MspA (también llamadas mutantes M1-NNN), los mutantes QQN muestran niveles de corriente bien separados utilizando hebras de homopolímero-hp pero las separaciones relativas entre los niveles son diferentes en la porina M1-QQN. Para cada poro, los datos se recogieron con el ADN en horquilla con colas A₅₀, T₅₀ y C₅₀ (SEQ ID NO: 49, SEQ ID NO: 32, SEQ ID NO: 48, respectivamente). Las corrientes de bloqueo se representaron como una fracción de la corriente de poros abierta no bloqueada (FIGURA 16). En cada caso de poli-T bloquea más que poli-C y poli-C bloquea más que poli-A. Cada pico está bien separado de los otros. En la porina QQN, los niveles de corriente promedio de poli-A y poli-C están menos separados que en la porina M1-NNN, pero poli-T está más separada de poli-C que en la porina M1-NNN. Sorprendentemente, el nivel relativo de bloqueo de poli-T en las dos porinas mutantes QQN es claramente diferente. Estos dos mutantes difieren sólo en sustituciones de dominio del borde lejos de la constricción. Sin estar ligado por la teoría, esto puede ser debido a las interacciones entre el dominio del borde y la horquilla de anclaje.

[0213] Las porinas mutantes QQN parecen ralentizar el traslado de ADN a través de MspA. La principal motivación para construir los mutantes QQN era ralentizar el paso del ADN. El traslado de un segmento heterogéneo de ssADN de 100 nt (sin horquilla de anclaje) se registró junto con la duración de los estados de bloqueo profundos. La representación de supervivencia (FIGURA 17) muestra la fracción de sucesos de bloqueo que duran más que el tiempo t. Durante los primeros ~ 100 μs el mutante NNN se descompone significativamente más rápido que los mutantes con la zona de constricción de QQN. Estos datos son consistentes con una mayor barrera para el traslado a través de QQN.

Ejemplo 10: Construcción de un mutante cuádruple de Msp por delección de M. smegmatis

[0214] Para la preparación de la porina de MspA, se extrajeron selectivamente proteínas de la cepa mutante de *M. smegmatis* ML16, que contiene sólo uno (*MspB*) de los cuatro genes Msp (los otros son *MspA*, *MspC*, y *MspD*). El procedimiento explota la estabilidad térmica extrema de MspA por ebullición de células de *M. smegmatis* en 0,5% de n-octilpolioxi-etileno (OPOE), un detergente no iónico, y se obtiene la porina de MspA con muy poca contaminación por otras proteínas (Heinz y Niederweis, Anal. Biochem 285: 113-20 (2000)). Sin embargo, la expresión de fondo de MspB es todavía detectable en inmunotransferencias utilizando un antisuero específico de Msp (Stephan et al, Mol Microbiol 58: 714-30 (2005)), lo que indica que oligómeros mixtos de MspA/MspB podrían formar y contribuir a la heterogeneidad del poro observada en los experimentos de reconstitución de los poros. Por lo tanto, uno de los objetivos era construir una cepa de *M. smegmatis* libre de porinas endógenas. Dado que *M. smegmatis* requiere la actividad de porina para la supervivencia, se integró un casete de expresión de MspA flanqueado por *loxP* en el sitio cromosómico *attB* para el micobacteriófago L5 de la porina mutante triple ML16.

[0215] Esto restauró la expresión de monómero de MspA en la cepa ML56 a la mitad del nivel de tipo salvaje. A

continuación, el gen *MspB* fue reemplazado por un gen *hyg* flanqueado por FRT utilizando el vector suicida pMN247 en una estrategia de dos etapas como se ha descrito (Stephan et al, Gene 343: 181-190 (2004)). Después de la escisión del gen *hyg* por la Flp recombinasa, se obtuvo la cepa mutante cuádruple de porina ML59 ($\Delta MspA \Delta MspB \Delta MspC \Delta MspD \text{ attB}::\text{loxP-MspA-loxP}$). La delección del gen *MspB* se confirmó por hibridación de transferencia de Southern. La PCR demostró la ausencia de cada uno de los cuatro genes *Msp* originales (FIGURA 19). La escisión del casete *loxP-MspA-loxP* dio lugar a clones pequeños, viables, uno de los cuales (ML180) fue examinado con más detalle. Las proteínas se extrajeron de células ML180 utilizando el mismo procedimiento de alta temperatura y el análisis Western demostró que las células ML180 no expresaron proteínas de porina *Msp* ni hubo sucesos de reconstitución en experimentos de bicapa lipídica después de la adición de 20 μg de proteína (FIGURA 20). En conjunto, estos resultados demuestran que se ha creado una porina mutante de *M. smegmatis* que carece de las cuatro porinas *Msp*. Sin embargo, no fue posible detectar la expresión de monómero de *MspA* usando vectores de expresión de *MspA*, muy probablemente debido a mutaciones secundarias desconocidas. Por lo tanto, esta cepa de *M. smegmatis* no se puede utilizar para la expresión de poros de *MspA* diseñados para el traslado del ADN.

15 Ejemplo 11: Construcción del mutante cuádruple de *Msp* ML705 de *M. smegmatis* por delección utilizando la expresión inducible de *MspA*

[0216] Para el aislamiento de porinas *MspA* de tipo salvaje y mutantes, se utiliza actualmente la cepa M16 de *M. smegmatis* ML16 ($\Delta MspA, \Delta MspC, \Delta MspD$). Sin embargo, la expresión de base de *MspB* complica la interpretación de los experimentos de traslado. Por lo tanto, la construcción de una cepa de *M. smegmatis* que carece de los cuatro genes *Msp* es necesaria para mejorar los experimentos de un solo poro. Para ello, el gen *MspA*, bajo el control del promotor inducible por acetamida, se integró en el sitio *attB* de L5 de *M. smegmatis* ML16 dando lugar a la eliminación del gen de *MspB* por intercambio alélico. Por lo tanto, en presencia de acetamida, *MspA* se expresó de rescatar el crecimiento del mutante cuádruple de *M. smegmatis*.

[0217] Para lograr esto, se construyó el plásmido de integración pML967, que contiene el gen *MspA* bajo el control del promotor inducible por acetamida (FIGURA 21A). El vector de delección *MspB*, pML1611 (FIGURA 21B), también se construyó y contiene los dos genes indicadores *gfp* y *xyIE* como marcadores para la integración y la sustitución alélica.

[0218] Se obtuvo la cepa ML341 (ML16, *attP*::pML967) después de la integración del plásmido de expresión de monómero de *MspA* pML967 en *M. smegmatis* ML16. El gen de resistencia a higromicina fue extraído de esta cepa por una expresión temporal de la Flp recombinasa del plásmido pML2005 como se ha descrito previamente (Song et al., Mycobacteria protocols (2008)) dando lugar a la cepa ML343 (ML341, *attP*::*p_{acet}-MspA*). Para examinar la funcionalidad del monómero de gen *MspA* integrado, se extrajo *MspA* con un detergente de células no inducidas e inducidas. La FIGURA 22 muestra que *MspA* se expresa en el 20% de los niveles de tipo salvaje de la construcción integrada después de la adición de acetamida al 2%. Este nivel de monómero de *MspA* es suficiente para permitir la supervivencia de *M. smegmatis*. Hubo poca expresión fondo de porinas de *Msp* en células no inducidas (FIGURA 22) lo que demuestra que el sistema de expresión está regulado.

[0219] A continuación, el vector de delección de *MspB* pML1611 se transformó en ML343. Los transformantes se sembraron en placas de agar Middlebrook 7H10 que contienen 10% de sacarosa para la selección directa de dobles candidatos cruzados. Se obtuvieron varias colonias, que mostraron la presencia de GFP mediante fluorescencia verde tras irradiación con luz azul y la ausencia de *XylE*. La PCR de la colonia de uno de los clones confirmó la ausencia del gen *MspB* y la construcción de un mutante cuádruple *Msp* viable. Esta cepa se denominó ML378. La cepa ML378 se transformó con el plásmido pCreSacB1 para eliminar el casete de expresión de *gfp-hyg*. Tras la selección del contador subsiguiente, se obtuvieron varios clones y se examinaron mediante PCR de colonias. Una de los ocho porina mutantes cuádruples sin marcar de *M. smegmatis* se denominó ML705 y se caracterizó adicionalmente.

[0220] Para examinar si monómeros de *MspA* complementan el fenotipo del mutante cuádruple, el plásmido de expresión de *MspA* pMN016 se transformó en ML705. La FIGURA 24 muestra que el crecimiento de ML705 en placas de agar 7H10 se redujo drásticamente; sin embargo, la expresión de *MspA* de pMN016 restauró completamente el crecimiento de ML705 a niveles de tipo salvaje (FIGURA 23). Estos resultados demostraron que ninguna mutación secundaria causó el defecto de crecimiento y que los monómeros de *MspA* se pueden expresar para producir porinas de *MspA* en el mutante cuádruple de *Msp* ML705.

[0221] El crecimiento del mutante cuádruple de porina ML705 en medio Middlebrook 7H9 fue mucho más lento que el *M. smegmatis* de tipo salvaje y significativamente más lento que el de la porina mutante triple ML16 (FIGURA 24). La adición de acetamida al 2% para inducir la expresión del monómero del gen de *MspA* en el sitio de L5 y la expresión de *MspA* en el plásmido pMN016 restauró la velocidad de crecimiento a los niveles de tipo salvaje (FIGURA 24). El crecimiento de ML705 en las placas y en cultivos líquidos fue más lenta que la de la triple mutante indicando que ML705 tenía menos porinas en la membrana externa que el mutante triple de *Msp* ML16. Esta suposición se confirmó en una transferencia Western (FIGURA 25). La cantidad del monómero de *MspA* es menor que el 5% de la comparada con *M. smegmatis* de tipo salvaje (wt), y 50% menos que la del mutante triple. La FIGURA 25 también demuestra que se puede inducir *MspA* hasta 25% del tipo salvaje cuando se añade acetamida

al 2%.

[0222] Los experimentos descritos anteriormente demuestran que un mutante cuádruple *Msp* (M1705) se ha construido, que se puede cultivar en presencia de acetamida para producir temporalmente monómeros de *MspA* de tipo salvaje. La cepa ML705 puede entonces ser transformada con un plásmido que contiene un casete de expresión para monómeros de *MspA* de tipo salvaje o mutantes, o porinas de *Msp* de una cadena de tipo salvaje o mutantes. La producción de monómeros *MspA* de tipo salvaje cortarse mediante el lavado de células de transferencia a un medio sin acetamida. Esto da lugar a la producción de monómeros de *MspA* de tipo salvaje o mutantes o porinas de *Msp* de una cadena de tipo salvaje o mutantes con menos contaminación por *MspA* de tipo salvaje. Por lo tanto, ML705 es adecuado para la producción de porinas *MspA* de tipo salvaje y mutadas para todos los objetivos.

Ejemplo 12: Construcción de un dímero de porinas de *MspA* de una cadena

[0223] El ADN de una sola hebra no es rotacionalmente simétrico. Por lo tanto, sería beneficioso tener un poro asimétrico para los propósitos de secuenciación. Para combinar las capacidades de secuenciación superiores de porinas de *MspA* con una mayor capacidad de adaptación de las propiedades del vestíbulo y la constricción a la secuenciación del ADN, se construye un nanoporo de *MspA* de una cadena. Los extremos de la cadena de *MspA* están muy juntas en el dímero de porinas de *MspA* (FIGURA 26A) y podrían estar conectados por un enlazador peptídico corto. Para probar esta idea, el monómero de gen *MspA* se fusionó junto con el monómero de gen *MspB*, que codifica una proteína con sólo dos alteraciones (A138P, E139A) en comparación con el monómero de *MspA* de tipo salvaje (Stahl et al., Mol. Microbiol. 40: 451 (2001)). El péptido (GGGGS)₃ (SEQ ID NO: 3), a menudo utilizado para enlazar proteínas (Huston et al, Proc Natl Acad Sci USA 85: 5879 (1988)), se utilizó para conectar el extremo C-terminal del monómero de *MspA* al N-terminal del monómero de *MspB* que carece de péptido señal (FIGURA 26B). El dímero de porinas *MspA-MspB* resultante se colocó bajo el control del promotor p_{smyc} constitutivo en el plásmido pML870 y después se expresó en *M. smegmatis* ML16. La proteína se purificó usando el procedimiento de extracción de calor estándar. Aunque el nivel de expresión del dímero de porinas de *MspA* de una cadena era menor que la de la porina de *MspA* de tipo salvaje (FIGURA 26C), la actividad túnel de ambas porinas era similar (FIGURA 26D). El análisis de los registros de corriente mostró que la conductancia de un único túnel del poro formado por el dímero de *MspA* fue de 2,6 ns. Este resultado muestra que el segmento enlazador no afecta al plegado o función de poros de *Msp*.

Ejemplo 13: Construcción de una porina de *MspA* de una cadena

[0224] Para combinar las capacidades de secuenciación superiores de *MspA* con una mayor capacidad de adaptación de las propiedades del vestíbulo y la constricción a la secuenciación del ADN, se construye un octámero de porinas de *MspA* de una cadena que permite las propiedades óptimas del vestíbulo y la zona de restricción para la secuenciación de ADN. Los extremos de la cadena de *MspA* están muy juntos en la porina de *MspA* y están conectados por un enlazador peptídico corto. El péptido (GGGGS)₃ (SEQ ID NO: 3) se utiliza para conectar el extremo carboxi-terminal del monómero de *MspA* anterior al extremo amino-terminal del siguiente monómero de *MspA*, que carece de péptido señal.

[0225] Para crear un vector que comprende la secuencia de porina de *MspA*, cada secuencia de monómero de *MspA* está flanqueada por un sitio de restricción único, que permite la capacidad de mutar cualquier monómero individual. La secuencia de porina de *MspA* completa está flanqueada por sitios de restricción *PacI* y *HindIII*. Los sitios de restricción entre las secuencias de monómero de *MspA* comprenden: *BamHI*, *ClaI*, *EcoRV*, *HpaI*, *KpnI*, *MluI*, *NdeI*, *NheI*, *PstI*, *SacI*, *SpeI*, *XbaI*, *NotI* y *SphI* (FIGURA 31.). Para crear la secuencia de porina de *MspA*, cada secuencia de *MspA* se ensambla por etapas para formar una *MspA* de una cadena dimérica, tetramérica y octamérica utilizando los sitios de restricción únicos. Para evitar problemas de recombinación en la creación del multímero de *MspA* de una cadena, siete genes de *MspA* se sintetizan con diferentes usos de codones (SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27), es decir, los genes codifican la misma secuencia de aminoácido, sin embargo, la secuencia de ADN ha sido alterada de la secuencia de nucleótidos del gen nativo de *MspA* (SEQ ID NO: 20). Para crear la secuencia de porina de *MspA*, el primer monómero de *Msp* debe contener la secuencia líder como se muestra en la FIGURA 18 (por ejemplo, los aminoácidos 1 a 27 de SEQ ID NO: 28)). Cada una de las siete secuencias de monómero de *Msp* después de la primera secuencia de monómero de *Msp* puede comprender SEQ ID NO: 1 o una mutación de la SEQ ID NO: 1 seleccionada de entre cualquiera de las mutaciones enumeradas en la Tabla 7. El vector de expresión pML2604 es el vector origen que comprende la secuencia de porin de *MspA* clonada en los sitios de restricción *PacI* y *HindIII*. pML2604 se transforma en la porina mutante cuádruple y los niveles de expresión y el estado oligomérico de la porina de *MspA* se comprueban por transferencia Western de las proteínas nativas y desnaturalizadas. La actividad del túnel de la porina de *MspA* se comprueba mediante experimentos de bicapa lipídica

Tabla 7: Mutantes de *MspA*

Fila 1	Fila 2
<i>MspA</i> D90A	<i>MspA</i> T84C

ES 2 576 114 T3

MspA D91A MspA D90A/D91A MspA D90E MspA D91E MspA D90E/D91E	MspA I87C MspA D91C MspA D93C MspA A96C MspA P97C
MspA D90F MspA D91F MspA D90F/D91F MspA D90G MspA D91G MspA D90G/D91G MspA D90H MspA D91H MspA D90H/D91H MspA D90K MspA D91K MspA D90K/D91K MspA D90L MspA D91L MspA D90L/D91L MspA D90R MspA D91R MspA D90R/D91R MspA D90S MspA D91S MspA D90S/D91S MspA D90W MspA D91W MspA D90W/D91W MspA D90Y MspA D91Y MspA D90Y/D91Y MspA Q126C MspA D90N MspA D91N MspA D93N MspA D90N/D91N	MspA G100C MspA N102C MspA P107C MspA G112C MspA V113C MspA S114C MspA D118C MspA N121C MspA E127C MspA F131C MspA D134C MspA S136C MspA A138C MspA E139C MspA G141C MspA V144C MspA H148C MspA T150C MspA A155C MspA R161C MspA R165C MspA S173C MspA T175C MspA E179C MspA V184C MspA N79C/D90K/D91N/P97C MspA K47S/D90K/D91N/P97C/D134C MspA ΔA96-P98 MspA ΔT95-F99 MspA ΔI94-G100 MspA ΔD93-L101 MspA ΔG92-N102
MspA D90N/D91N/D93N MspA D90Q/D91N/D93N MspA D90Q/D91Q/D93N MspA D90T/D91N/D93N MspA D90T/D91T/D93N MspA D91E MspA D90E MspA D90E/D91E MspA D90N/D91N/D93Q MspA D90N/D91N/G92Q/D93N MspA G1C MspA D3C MspA E5C MspA D10C MspA D13C MspA R14C MspA T17C MspA W21C MspA D22C MspA G27C MspA R33C MspA R38C MspA G44C MspA K47C MspA I49C MspA E57C MspA G60C	MspA N79R/D90N/D91N/D93N MspA N79W/D90N/D91N/D93N MspA D90N/D91N/D93N/Q126R MspA D90N/D91N/D93N/TI30R MspA D90N/D91N/D93N/D134R MspA D90N/D91N/D93N/Q126W MspA D90N/D91N/D93N/T130W MspA D90N/D91N/D93N/D134W MspA D90N/D91N/D93N/D118W/D134R/E139K MspA D90N/D91N/D93N/D118F/D134R/E139K MspA D90N/D91N/D93N/D118H/D134R/E139K MspA D90N/D91N/D93N/D118Y/D134R/E139K MspA N79W/D90N/D91N/D93N/D118R/E139K MspA N79F/D90N/D91N/D93N/D118R/E139K MspA N79H/D90N/D91N/D93N/D118R/E139K MspA N79Y/D90N/D91N/D93N/D118R/E139K MspA D90N/D91K/D93N MspA D90N/D91R/D93N MspA D90N/D91W/D93N MspA D90N/D91W/D93N MspA D90N/D91T/D93N MspA D90N/D91L/D93N MspA D90N/D91H/D93N MspA D90N/D91S/D93N MspA D90N/D91N/D93N/D118R MspA D90N/D91N/D93N/D118R/E139R MspA D90N/D91N/D93N/D118R/E139K

MspA E63C MspA G69C MspA S73C MspA L74C MspA V76C	MspA D90N/D91N/D93N/D118R/D134R/E139K MspA D90Q/D91N/D93N/D118R/D134R/E139K MspA D90Q/D91Q/D93N/D118R/D134R/E139K MspA D90T/D91N/D93N/D118R/D134R/E139K MspA D90T/D91T/D93N/D118R/D04R/E09K
---	---

LISTADO DE SECUENCIAS

[0226]

5 <110> University of Washington
University of Alabama Research Foundation
Gundlach, Jens H.
Niederweis, Michael
10 Butler, Thomas Z.
Pavlenok, Mikhail
Troll, Mark
<120> Nanoporos MSP y procedimientos relacionados
<130> UWOTL-1-33161
<150> Estados Unidos 61/098.938
15 <151> 2008-09-22
<160> 52
<160> 52

<170> PatentIn version 3.5
20
<210> 1
<211> 184
<212> PRT
<213> Mycobacteria Smegmatis
25
<400> 1
Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu
1 5 10 15
Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp
30 20 25 30
Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr
35 35 40 45
Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu
50 55 60
35 Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe
65 70 75 80
Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Gly Asp Ile Thr Ala
85 90 95
40 Pro Pro Phe Gly Leu Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly
100 105 110
Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val
115 120 125
Ala Thr Phe Ser Val Asp Val Ser Gly Ala Glu Gly Gly Val Ala Val
130 135 140
45 Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu
145 150 155 160
Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr
165 170 175
50 Tyr Gly Glu Pro Trp Asn Met Asn
180

<210> 2
<211> 184
55 <212> PRT
<213> Mycobacteria Smegmatis

<400> 2
60 Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu
1 5 10 15
Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp
20 25 30
Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr

ES 2 576 114 T3

Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu
 35 40 45
 50 55 60
 5 Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe
 65 70 75 80
 Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala
 85 90 95
 Pro Pro Phe Gly Leu Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly
 100 105 110
 10 Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val
 115 120 125
 Ala Thr Phe Ser Val Asp Val Ser Gly Pro Ala Gly Gly Val Ala Val
 130 135 140
 15 Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu
 145 150 155 160
 Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr
 165 170 175
 Tyr Gly Glu Pro Trp Asn Met Asn
 180

20

<210> 3
 <211> 15
 <212> PRT
 25 <213> Secuencia artificial

<220>
 <223> Péptido enlazador entre péptidos de MspA

30

<400> 3
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15

35

<210> 4
 <211> 72
 <212> ADN
 <213> Secuencia artificial

40

<220>
 <223> oligonucleótidos

<400> 4
 gctgttgctc tctcgcaaca gcaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
 45 aaaaaaaaaa aa 72

50

<210> 5
 <211> 76
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> oligonucleótidos

55

<400> 5
 gctctgttgc tctctcgcaa cagagcaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
 aaaaaaaaaa aaaaaa 76

60

<210> 6
 <211> 80
 <212> ADN
 <213> Secuencia artificial

65

<220>
 <223> oligonucleótidos

ES 2 576 114 T3

<400> 6
gctgtctgtt gctctctcgc aacagacagc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
aaaaaaaaaa aaaaaaaaaa 80

5
<210> 7
<211> 75
<212> ADN
<213> Secuencia artificial

10
<220>
<223> oligonucleótidos

15
<400> 7
cccccaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
ctctattctt atctc 75

20
<210> 8
<211> 75
<212> ADN
<213> Secuencia artificial

25
<220>
<223> oligonucleótidos

30
<400> 8
cccccaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
cacacacaca cacac 75

35
<210> 9
<211> 15
<212> ADN
<213> Secuencia artificial

40
<220>
<223> oligonucleótidos

45
<400> 9
gagataagaa tagag 15

50
<210> 10
<211> 50
<212> ADN
<213> Secuencia artificial

55
<220>
<223> oligonucleótidos

60
<400> 10
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 50

65
<210> 11
<211> 50
<212> ADN
<213> Secuencia artificial

70
<220>
<223> oligonucleótidos

75
<400> 11
aaaaaacccc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 50

5 <210> 12
 <211> 50
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> oligonucleótidos
 10 <400> 12
 aaaccctaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 50
 15 <210> 13
 <211> 50
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> oligonucleótidos
 20 <400> 13
 aacaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 50
 25 <210> 14
 <211> 50
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> oligonucleótidos
 <400> 14
 35 caaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 50
 40 <210> 15
 <211> 50
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> oligonucleótidos
 45 <400> 15
 ccccaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 50
 50 <210> 16
 <211> 50
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> oligonucleótidos
 55 <400> 16
 ccccccccc ccccccccc ccccccccc ccccccccc ccccccccc 50
 60 <210> 17
 <211> 50
 <212> ADN
 <213> Secuencia artificial
 65 <220>
 <223> oligonucleótidos

ES 2 576 114 T3

<400> 17
 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 50

5 <210> 18
 <211> 50
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Oligonucleótidos

<400> 18
 gggaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 50

15

<210> 19
 <211> 45
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <223> Oligonucleótidos

25 <400> 19
 ggcggtggcg gtagcggcgg tggcggtagc ggcggtggcg gtagc 45

30 <210> 20
 <211> 549
 <212> ADN
 <213> Mycobacteria Smegmatis

35 <400> 20
 ggcctggaca acgagctgag cctcgttgat ggccaggacc gcaccctcac cgtgcagcag 60
 tgggacacct tcctcaatgg tgtgttcccc ctggaccgca accgtcttac ccgtgagtgg 120
 ttccactccg gtcgcgcaaa gtacatcgtg gccggccccg gtgccgacga gttcgagggc 180
 acgctggaac tcggctacca gatcggcttc ccgtggctgc tgggtgtggg catcaacttc 240
 agctacacca ccccgaaat cctgatcgac gacggtgaca tcaccgctcc gccgttcggc 300
 40 ctgaactcgg tcatcaccac gaacctgttc cccgggtgtg cgatctcggc agatctgggc 360
 aacggccccg gcatccagga agtcgcaacg ttctcggctc acgtctccgg cgccgagggg 420
 ggcgtggccg tgtcgaacgc ccacggcacc gtgaccgggt cgccggcgcg tgtgctgctg 480
 cgtccgttcg cccgcctgat cgcctcgacc ggtgactcgg tcaccaccta cggcgaacc 540
 tggaaatg 549

45

<210> 21
 <211> 549
 <212> ADN
 <213> Mycobacteria Smegmatis

50

<220>
 <223> Gen de MspA Mutante

55 <400> 21
 ggcctagaca atgagcttag ccttgttgac ggccaagacc gaaccctgac cctccagcaa 60
 tgggatacct ttctcaacgg tgtattcccc ctggatcgca atcgtctaac ccgcgagtgg 120
 ttccattccg gccgcgctaa gtatatcgta gccggacccg gggccgatga gtttgagggg 180
 acgtcgaac ttggctatca gataggcttt ccgtggctgc taggtgtcgg cattaacttt 240
 60 agctatacca caccgaatat cctcatcgat gacggcgaca taaccgcgcc gcctttcggg 300
 ctgaattcgg taatcacgcc gaatctgttt cccgggggtg ctatctcagc agacctgggg 360
 aacgggtccc gaatccaaga agtggcaact ttctccgtcg atgtctcggg cgccgagggc 420
 ggcgtagccg tctcgaatgc ccatggcacg gtgacagggt cagccggggg tgtactgcta 480
 cgtccattcg cgcgcctaata cgcgtcgaca ggtgattcgg ttaccacata cgggtaacca 540
 65 tggaaatg 549

<210> 22

ES 2 576 114 T3

<211> 549
 <212> ADN
 <213> Mycobacteria Smegmatis

5 <220>
 <223> Gen de MspA mutante

<400> 22
 ggactggata acgaactgag tctggtagat gttcaggatc gcaatctcac agtgcaacag 60
 10 tgggacactt tcctgaatgg cgtgtttccc ctcgaccgga accgacttac gcgtgaatgg 120
 tttcactcag gtcgtgccaa atacattgtg gcgggcccg gtgccgacga attcgaaggc 180
 accctggagc tcgggtacca aatcgggttc ccttgggtccc tgggcgtggg tatcaatttc 240
 agttacacga ccccaacat tctgatagac gatggtgata tcacggctcc cccgtttggc 300
 ctcaactccg tcataacccc caacctcttc cctgggtgtct cgatttcggc cgatctcggc 360
 15 aatggccctg gcattcagga ggtcgcctacg ttttcgggtgg acgtgtccgg tgccgaaggc 420
 ggtgtggcgg tgtccaacgc gcacgggtacc gtcaccggcg cggcgggccc cgtgctcctg 480
 cggccggttg cccggctgat tgccctccacc ggggactccg tcacgacctg tggcgagccc 540
 tggaaacatg 549

20 <210> 23
 <211> 549
 <212> ADN
 <213> Mycobacteria Smegmatis

25 <220>
 <223> Gen de MspA mutante

<400> 23
 30 ggccttgaca acgagctaag cctagttgac ggccaggacc gtaccctaac cctacagcag 60
 tgggatacct ttctcaatag cgtcttccca ctggaccgca atcgtctgac ccgagagtgg 120
 ttccactccg ggcgcgcaaa gtatatcgtc gccgggtccc gtgccgacga gttcagattc 180
 acgcttgaac taggctacca gattggcttc ccgtggctgc ttgggtgtgg cataaacttc 240
 agctaacaca ctccgaacat cctaactcag gacggggaca ttaccgcacc gccattcggg 300
 35 ctgaactcgg tgatcacacc gaacctggtt cccggagtgt caatctccgc agatcagggg 360
 aacggaccgg ggatccagga agttgcaacc ttctcagtcg acgtctcagg cgagagggg 420
 ggcgtcgcgg taticgaacgc ccacggcact gtgacgggtg ccgccgggtg tgtcctgctc 480
 cgtcccttcg cacgcctcat cgcactcgacg ggtgactcgg taaccacgta cggggaaccg 540
 tggaaacatg 549

40 <210> 24
 <211> 549
 <212> ADN
 <213> Mycobacteria Smegmatis

<220>
 <223> Gen de MspA mutante

<400> 24
 50 gggctggata acgaactgag tctcctggat ggacaggatc gcaatctcac ggtgcagcag 60
 tgggacacat tccttaatgg ggtgttcccc cttgaccgta accggcttac acgtgaatgg 120
 ttccactcgg gtcgggcaaa gtacatagtg gctggccctg gtgctgacga gttcgaaggc 180
 aacttgaac tcggttacca aatcgggttc ccctgggtcac tgggggtggg aatcaacttc 240
 55 agctacacaa cccctaacat actgatagac gacggtgaca tcaactgctcc tccgtttggc 300
 cttactcag tcattacccc taacctattc cccgggtggt cgatatcggc ggatcttggc 360
 aacggcccgg gcatacagga agtcgcgacg ttctcgggtg acgtttccgg ggccgagggc 420
 ggggtggcag tgtcaaacgc tcacgggacc gttaccgggg cggcaggcgg ggtgcttctg 480
 cgcccgttcg cccgtctgat agcctcaacc ggcgactcag tcacaacctg cggcgagccc 540
 60 tggaaacatg 549

65 <210> 25
 <211> 549
 <212> ADN
 <213> Mycobacteria Smegmatis

<220>

ES 2 576 114 T3

<223> Gen de MspA mutante

<400> 25

```

5  ggactcgata atgaactcag tctagtagac ggtcaagatc ggaatctgac agtccaacaa 60
   tgggatactt ttctgaacgg cgtctttccg ctcgatcggg atcgactcac gcgcgaaatgg 120
   tttcattcag ggcgtgcaaa atatattgtc gcggggccgg gcgcggatga atttgaaggt 180
   accctcgagc tggggatatca aattgggttt ccttgggtccc ttggcgtcgg tattaatttt 240
   agttatacga cgcccaatat tctcatagat gatggcgata taacggcgcc cccatttggg 300

10  ctcaattccg tgataacgcc caatctcttt cctggcgctt ccatttccgc cgacctcggg 360
   aatggctctg gtattcaaga ggtggctacc ttttccgtgg atgtgtcggg tgcggaaggc 420
   ggtgtagcgg tatccaatgc gcatgggtaca gtcacaggcg ccgcgggtgg cgtcctcctc 480
   cggccatttg cacggctaata tgcgtccacg ggggattccg tgacgacgta tggtcagccg 540
   tggaaacatg                                     549

```

15

<210> 26

<211> 549

<212> ADN

20 <213> Mycobacteria Smegmatis

<220>

<223> Gen de MspA mutante

25 <400> 26

```

   gggctagata atcagcttag tcttgtggac gggcaagatc gaacgctgac gctccaacaa 60
   tgggatacgt ttctaaacgg cgtatttccg ctcgatcggg atcgactaac gcgcgaaatgg 120
   tttcattcgg gccgagctaa atatatagta gcaggaccgg gggcagatga atttgaagga 180
   acactcgagc ttgggtatca aatagggttt ccttgggtccc taggagtcgg gattaatttt 240
30  agttatacaa cacccaatat actcattgat gatggcgata taacggcgcc tccttttggg 300
   ctcaattccg taattacgcc taatctcttt ccgggggtgt ctatatcagc cgacctcggg 360
   aacgggtccg gaattcaaga ggtggcgact ttttccgtgg atgtttcggg agcggaaggc 420
   ggtgtagcag cgtccaatgc tcatggaacg gtcacaggcg cagcaacagg agtactcca 480
   cggccatttg cgcggtcaat tgcgtctaca ggcgattccg ttacgacata tgggtgagcca 540
35  tggaaatg                                     549

```

35

<210> 27

<211> 549

<212> ADN

40 <213> Mycobacteria Smegmatis

<220>

<223> Gen de MspA mutante

45 <400> 27

```

   ggacttgata acgaactaag tctggtagat ggtcaggatc gtaatcttac agttcaacag 60
   tgggacactt tcctgaatgg cgtttttcct ctcgaccgga accgactgac gcgggaaatgg 120
   tttcactcag gccgtgcaaa atacaatgta gcgggtccgg gagcggacga attcgaaggg 180
50  accctagagc ttgggtacca aatgggtttc ccttgggtccc tcggcgtagg tataaatttc 240
   agttacacga cacccaacat tcttatagac gatgggggata ttacggcacc cccctttggg 300
   ctcaactccg taataacacc caacctcttc cctggagtct ctatttcagc cgatctcggg 360
   aatggacctg gaattcagga ggttgctaca ttttctgtgg acgtgtcagg tgcagaagga 420
   ggtgtcgcgg tctccaacgc gcacgggtacg gtcacgggcg ctgcgggggg cgttctcctt 480
55  cggccctttg cgcggtcat tgcctccaca ggggactccg taacgactta tggacagcct 540
   tggaaatg                                     549

```

55

<210> 28

<211> 211

<212> PRT

60 <213> Mycobacteria Smegmatis

<220>

65 <223> Proteína de MspA

<400> 28

Met Lys Ala Ile Ser Arg Val Leu Ile Ala Met Val Ala Ala Ile Ala

ES 2 576 114 T3

1 Ala Leu Phe Thr Ser Thr Gly Thr Ser His Ala Gly Leu Asp Asn Glu
 5 Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu Thr Val Gln Gln Trp
 Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp Arg Asn Arg Leu Thr
 Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr Ile Val Ala Gly Pro
 10 Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu Gly Tyr Gln Ile Gly
 Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe Ser Tyr Thr Thr Pro
 15 Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala Pro Pro Phe Gly Leu
 Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly Val Ser Ile Ser Ala
 Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val Ala Thr Phe Ser Val
 20 Asp Val Ser Gly Ala Glu Gly Gly Val Ala Val Ser Asn Ala His Gly
 Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu Arg Pro Phe Ala Arg
 25 Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr Tyr Gly Glu Pro Trp
 Asn Met Asn
 210

30 <210> 29
 <211> 215
 <212> PRT
 <213> Mycobacteria smegmatis

35 <220>
 <223> Proteína de MspB

<400> 29
 40 Met Thr Ala Phe Lys Arg Val Leu Ile Ala Met Ile Ser Ala Leu Leu
 Ala Gly Thr Thr Gly Met Phe Val Ser Ala Gly Ala Ala His Ala Gly
 Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu Thr
 45 Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp Arg
 Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr Ile
 50 Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu Gly
 Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe Ser
 Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala Pro
 55 Pro Phe Gly Leu Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly Val
 Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val Ala
 60 Thr Phe Ser Val Asp Val Ser Gly Pro Ala Gly Gly Val Ala Val Ser
 Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu Arg
 Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr Tyr
 65 Gly Glu Pro Trp Asn Met Asn
 210 215

ES 2 576 114 T3

<210> 30
 <211> 215
 <212> PRT
 <213> Mycobacteria Smegmatis

5
 <220>
 <223> Proteína de MspC

<400> 30

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

40 <210> 31
 <211> 207
 <212> PRT
 <213> Mycobacteria Smegmatis

45 <220>
 <223> Proteína de MspD

<400> 31

50 Met Arg Tyr Leu Val Met Met Phe Ala Leu Leu Val Ser Val Thr Leu
 1 5 10 15
 Val Ser Pro Arg Pro Ala Asn Ala Val Asp Asn Gln Leu Ser Val Val
 20 25 30
 Asp Gly Gln Gly Arg Thr Leu Thr Val Gln Gln Ala Glu Thr Phe Leu
 35 40 45
 55 Asn Gly Val Phe Pro Leu Asp Arg Asn Arg Leu Thr Arg Glu Trp Phe
 50 55 60
 His Ser Gly Arg Ala Thr Tyr His Val Ala Gly Pro Gly Ala Asp Glu
 65 70 75 80
 Phe Glu Gly Thr Leu Glu Leu Gly Tyr Gln Val Gly Phe Pro Trp Ser
 85 90 95
 60 Leu Gly Val Gly Ile Asn Phe Ser Tyr Thr Thr Pro Asn Ile Leu Ile
 100 105 110
 Asp Gly Gly Asp Ile Thr Gln Pro Pro Phe Gly Leu Asp Thr Ile Ile
 115 120 125
 65 Thr Pro Asn Leu Phe Pro Gly Val Ser Ile Ser Ala Asp Leu Gly Asn
 130 135 140
 Gly Pro Gly Ile Gln Glu Val Ala Thr Phe Ser Val Asp Val Lys Gly
 145 150 155 160

ES 2 576 114 T3

	Ala Lys Gly Ala Val Ala Val Ser Asn Ala His Gly Thr Val Thr Gly	
	Ala Ala Gly Gly Val Leu Leu Arg Pro Phe Ala Arg Leu Ile Ala Ser	
5	Thr Gly Asp Ser Val Thr Thr Tyr Gly Glu Pro Trp Asn Met Asn	
	165 170 180 185 190 195 200 205	
10	<210> 32 <211> 50 <212> ADN <213> Secuencia Artificial	
15	<220> <223> oligonucleótidos	
20	<400> 32 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	50
25	<210> 33 <211> 50 <212> ADN <213> Secuencia Artificial	
30	<220> <223> oligonucleótidos	
30	<400> 33 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaca	50
35	<210> 34 <211> 50 <212> ADN <213> Secuencia Artificial	
40	<220> <223> oligonucleótidos	
40	<400> 34 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaat	50
45	<210> 35 <211> 50 <212> ADN <213> Secuencia Artificial	
50	<220> <223> oligonucleótidos	
55	<400> 35 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaata	50
60	<210> 36 <211> 50 <212> ADN <213> Secuencia Artificial	
65	<220> <223> oligonucleótidos	
65	<400> 36 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaataa	50

ES 2 576 114 T3

5 <210> 37
 <211> 50
 <212> ADN
 <213> Secuencia Artificial

 <220>
 <223> oligonucleótidos

 10 <400> 37
 cccccccccc cccccccccc cccccccccc cccccccccc ccccccccca 50

 15 <210> 38
 <211> 50
 <212> ADN
 <213> Secuencia Artificial

 20 <220>
 <223> oligonucleótidos

 <400> 38
 cccccccccc cccccccccc cccccccccc cccccccccc ccccccccac 50

 25 <210> 39
 <211> 50
 <212> ADN
 <213> Secuencia Artificial

 30 <220>
 <223> oligonucleótidos

 35 <400> 39
 cccccccccc cccccccccc cccccccccc cccccccccc cccccccacc 50

 40 <210> 40
 <211> 58
 <212> ADN
 <213> Secuencia Artificial

 45 <220>
 <223> oligonucleótidos

 <400> 40
 cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc cccccccc 58

 50 <210> 41
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

 55 <220>
 <223> oligonucleótidos

 60 <400> 41
 cgaccagcac ggcatacatc 20

 65 <210> 42
 <211> 20
 <212> ADN
 <213> Secuencia Artificial

 <220>

	<223> oligonucleótidos	
5	<400> 42 cgttctcggc tcgatgatcc	20
	<210> 43 <211> 27 <212> ADN <213> Secuencia Artificial	
10	<220> <223> oligonucleótidos	
15	<400> 43 cctgatcaac aacggtaaca tcaccgc	27
20	<210> 44 <211> 21 <212> ADN <213> Secuencia Artificial	
25	<220> <223> oligonucleótidos	
30	<400> 44 ctgggcacgc ctgggcaacg g	21
35	<210> 45 <211> 21 <212> ADN <213> Secuencia Artificial	
40	<220> <223> oligonucleótidos	
45	<400> 45 tccggcgccc gcggtggcgt g	21
50	<210> 46 <211> 19 <212> ADN <213> Secuencia Artificial	
55	<220> <223> oligonucleótidos	
60	<400> 46 ggcgccaagg gtggcgtgg	19
65	<210> 47 <211> 20 <212> ADN <213> Secuencia Artificial	
70	<220> <223> oligonucleótidos	
75	<400> 47 cgttctcggc cgcgctctcc	20
80	<210> 48 <211> 50	

ES 2 576 114 T3

<212> ADN
 <213> Secuencia Artificial

 5 <220>
 <223> oligonucleótidos

 <400> 48
 cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc 50

 10 <210> 49
 <211> 50
 <212> ADN
 <213> Secuencia Artificial

 15 <220>
 <223> oligonucleótidos

 <400> 49
 20 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 50

 <210> 50
 <211> 52
 25 <212> ADN
 <213> Secuencia Artificial

 <220>
 <223> oligonucleótidos
 30 <400> 50
 aacccaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 52

 35 <210> 51
 <211> 47
 <212> ADN
 <213> Secuencia Artificial

 40 <220>
 <223> oligonucleótidos

 <400> 51
 45 tacgcataca tcctaagaac tcagactacc tccaataaa tccacac 47

 <210> 52
 <211> 47
 <212> ADN
 50 <213> Secuencia Artificial

 <220>
 <223> oligonucleótidos

 <400> 52
 55 tcagactacc tccaataaa tccgcagcaa tcctcacacc taataat 47

REIVINDICACIONES

- 5 1. Porina de *Mycobacterium Smegmatis* (Msp) que tiene un vestíbulo y una zona de constricción que definen un túnel, en la que la Msp comprende una MspA mutante que comprende una mutación en las posiciones 90, 91 y 93, en la que la mutación en la posición 90 es D90N, la mutación en la posición 91 es D91N, y la mutación en la posición 93 es D93N.
- 10 2. Msp, según la reivindicación 1, que comprende al menos un aminoácido con carga positiva adicional en comparación con el vestíbulo de una Msp de tipo salvaje.
- 15 3. Msp, según la reivindicación 2, que comprende la mutación en la posición 118, en la que la mutación en la posición 118 es D118R.
- 20 4. Msp, según la reivindicación 3, que comprende la mutación en la posición 139, en la que la mutación en la posición 139 es E139K o E139R.
- 25 5. Msp, según la reivindicación 4, que comprende las mutaciones en las posiciones 118, 139 y 134, en la que la mutación en la posición 118 es D118R, la mutación en la posición 134 es D134R, y la mutación en la posición 139 es E139K.
- 30 6. Utilización de una Msp, según cualquiera de las reivindicaciones anteriores, para la detección de la presencia de un analito.
- 35 7. Procedimiento para detectar la presencia de un analito, que comprende:
 aplicar un campo eléctrico suficiente para trasladar un analito de un primer medio líquido conductor a un segundo medio líquido conductor en comunicación de líquidos a través de una Msp, según cualquiera de las reivindicaciones 1-5 ; y
 medir una corriente de iones, en el que la aparición de un bloqueo en el patrón de corriente de iones indica la presencia del analito en el primer medio líquido conductor.
- 40 8. Procedimiento, según la reivindicación 7, en el que dicho primer medio líquido conductor se mantiene en tierra, y se aplica una tensión positiva al segundo medio líquido conductor.
- 45 9. Procedimiento, según la reivindicación 7 o la reivindicación 8, en el que el bloqueo o bloqueos en el patrón de corriente de iones se **caracterizan por**:
 (a) reducción de la corriente de iones a entre 80% y 50% del nivel no bloqueado; o
 (b) reducción de la corriente de iones a menos del 50% del nivel no bloqueado.
- 50 10. Procedimiento, según cualquiera de las reivindicaciones 7-9, en el que el procedimiento se lleva a cabo a 180 mV o más.
- 55 11. Aparato para utilizar en el procedimiento, según cualquiera de las reivindicaciones 7-10, que comprende una Msp, según cualquiera de las reivindicaciones 1-5, en el que el túnel de la Msp se encuentra entre un primer medio líquido conductor y un segundo medio líquido conductor; en el que al menos un medio líquido conductor comprende un analito, y en el que el sistema es operativo para detectar el analito.
- 60 12. Utilización, según la reivindicación 6, procedimiento, según cualquiera de las reivindicaciones 7-10, o aparato, según la reivindicación 11, en los que el analito es un polímero, por ejemplo ácido nucleico.
- 65 13. Procedimiento, según la reivindicación 12, en el que el analito es un polímero que comprende más de una unidad, que comprende además:
 identificar una o más unidades de polímero en un procedimiento que comprende medir la corriente de iones o proporcionar un patrón de corriente que comprende un bloqueo para cada unidad de polímero; y
 comparar uno o más bloqueos en cada patrón de corriente con (i) uno o más bloqueos en el patrón de corriente o (ii) uno o más bloqueos en un patrón de corriente obtenido usando un polímero que tiene unidades conocidas.
14. Utilización, procedimiento o aparato, según la reivindicación 12 o la reivindicación 13, en el que el ácido nucleico es ADN, por ejemplo ADN monocatenario, o ARN.
15. Utilización, procedimiento o aparato, según cualquiera de las reivindicaciones 12-14, para la secuenciación de ácidos nucleicos.
16. Ácido nucleico que codifica una Msp, según cualquiera de las reivindicaciones 1-5.
17. Vector que comprende un promotor unido operativamente a la secuencia de ácidos nucleicos, según la reivindicación 16.

18. Bacteria *Mycobacterium Smegmatis* mutante capaz de la expresión de Msp, comprendiendo la bacteria:

- 5 (a) una delección de una MspA de tipo salvaje;
(b) una delección de una MspC de tipo salvaje;
(c) una delección de una MspD de tipo salvaje;
(d) un vector según la reivindicación 17.

19. Bacteria *Mycobacterium Smegmatis* mutante, según la reivindicación 18, en la que la bacteria es la cepa M16 de *Mycobacterium Smegmatis*.

- 10 20. Porina de *Mycobacterium Smegmatis* obtenible de las bacterias *Mycobacterium smegmatis* mutantes de la reivindicación 18 o la reivindicación 19.

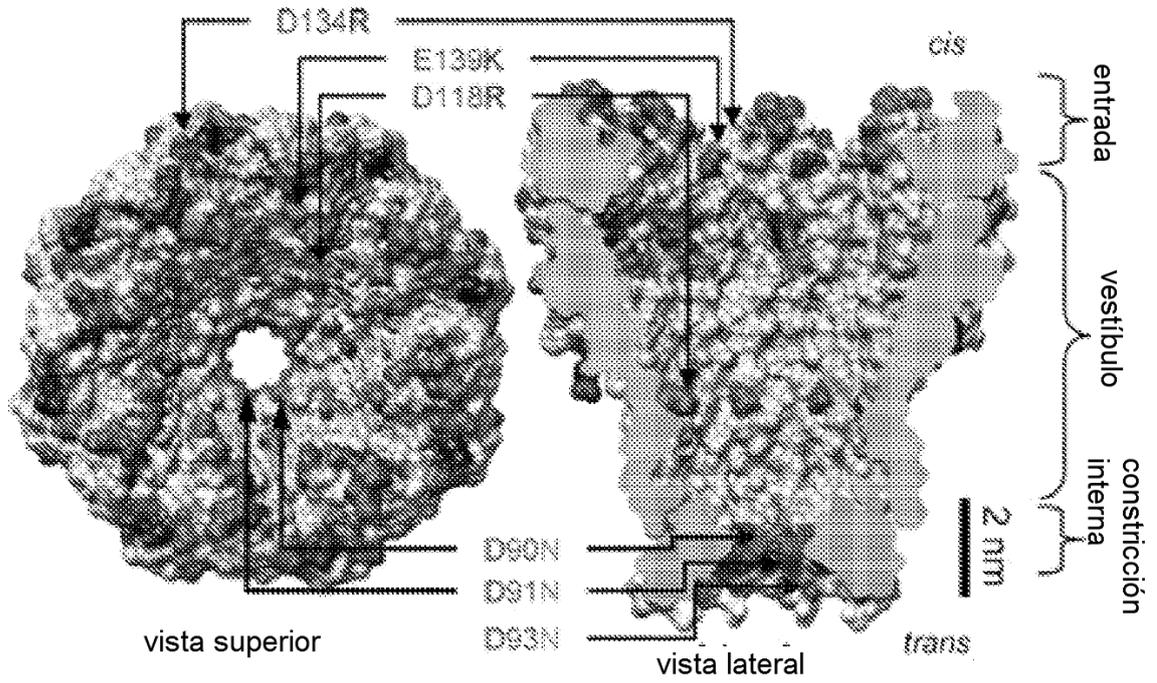


Figura 1

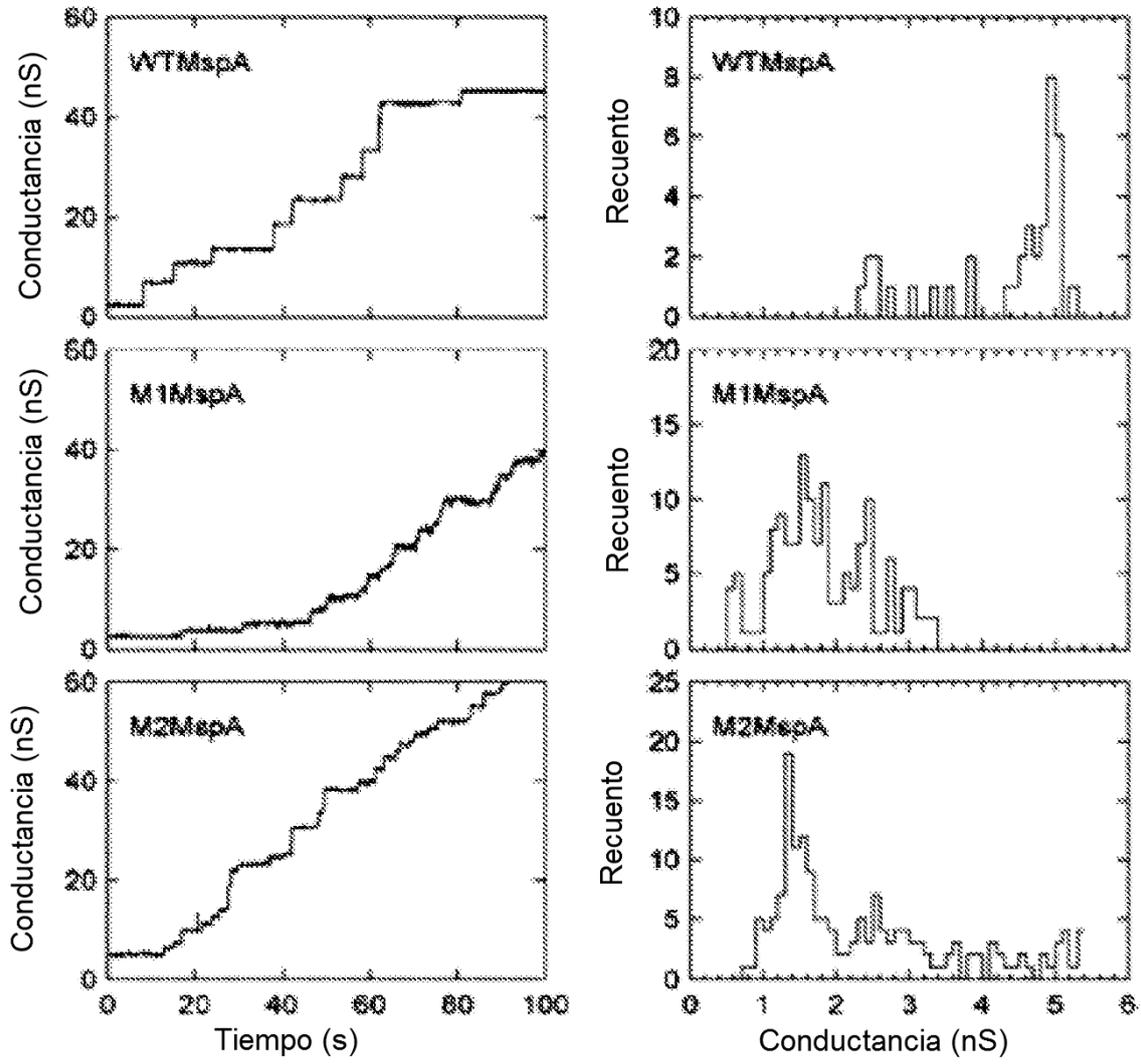
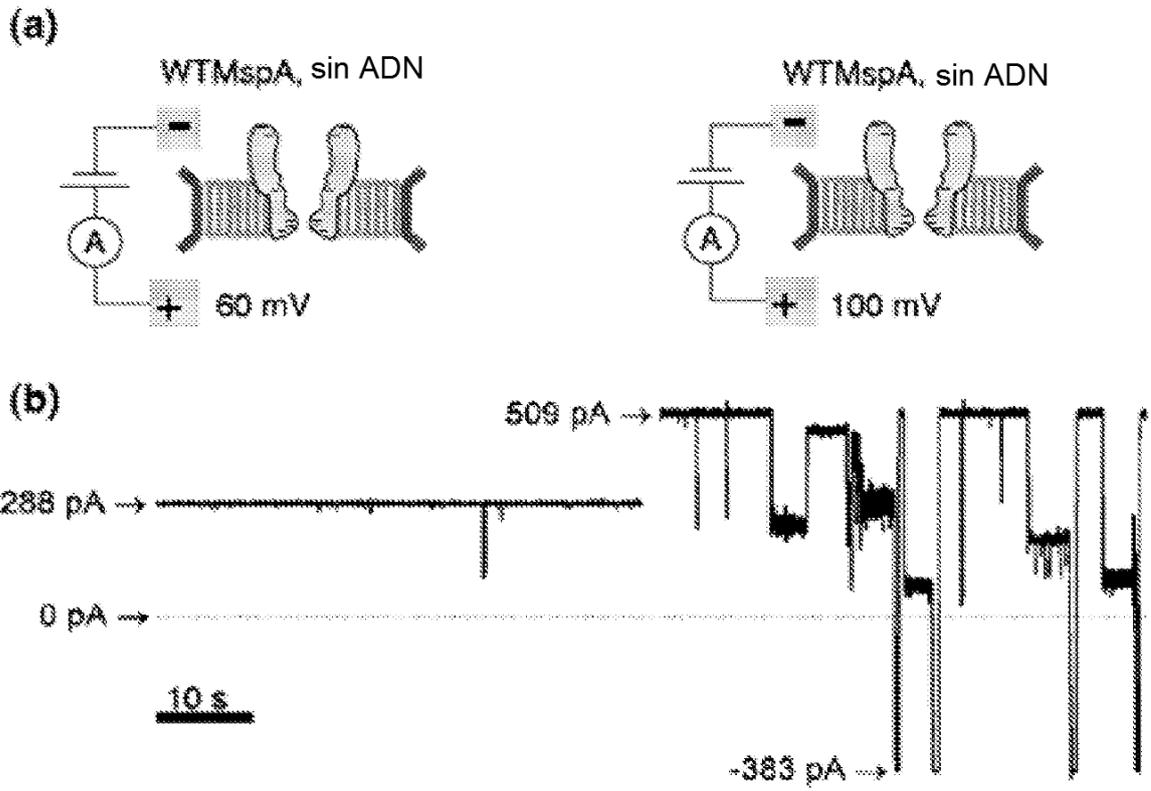


Figura 2



Figuras 3A, 3B

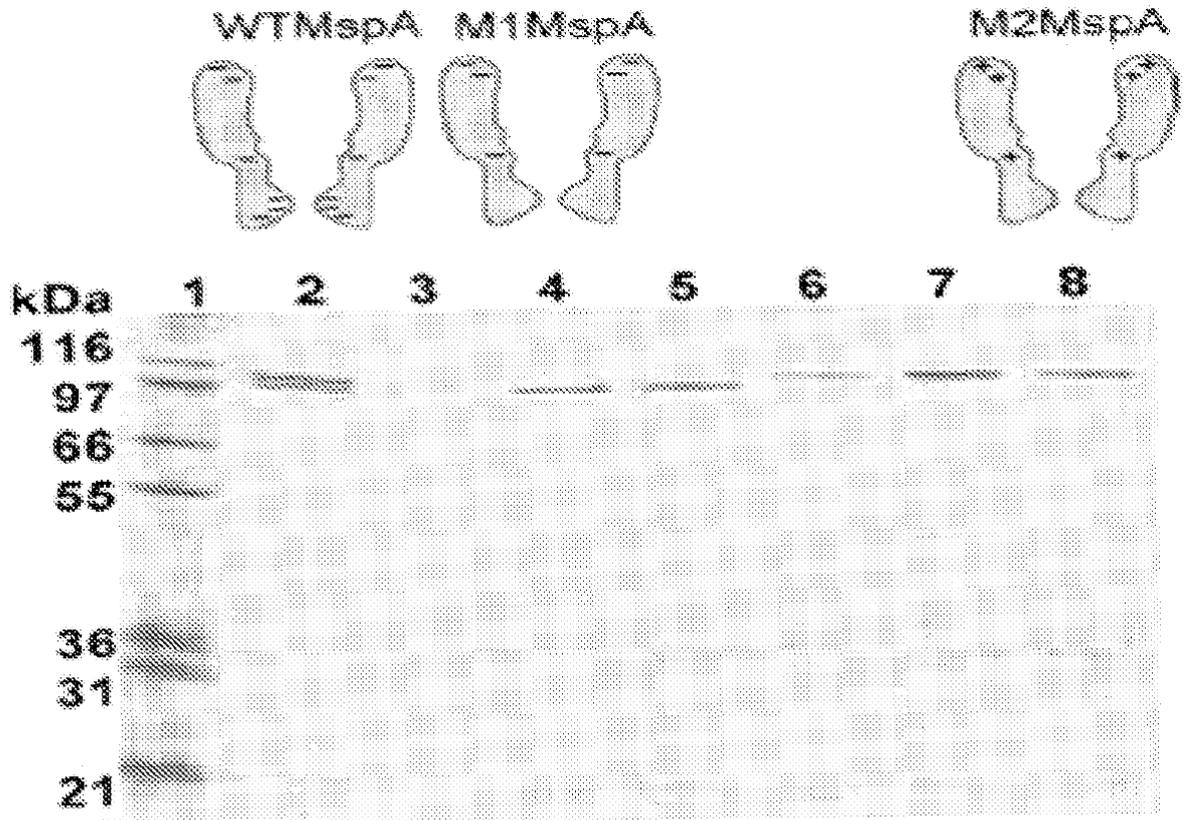


Figura 4

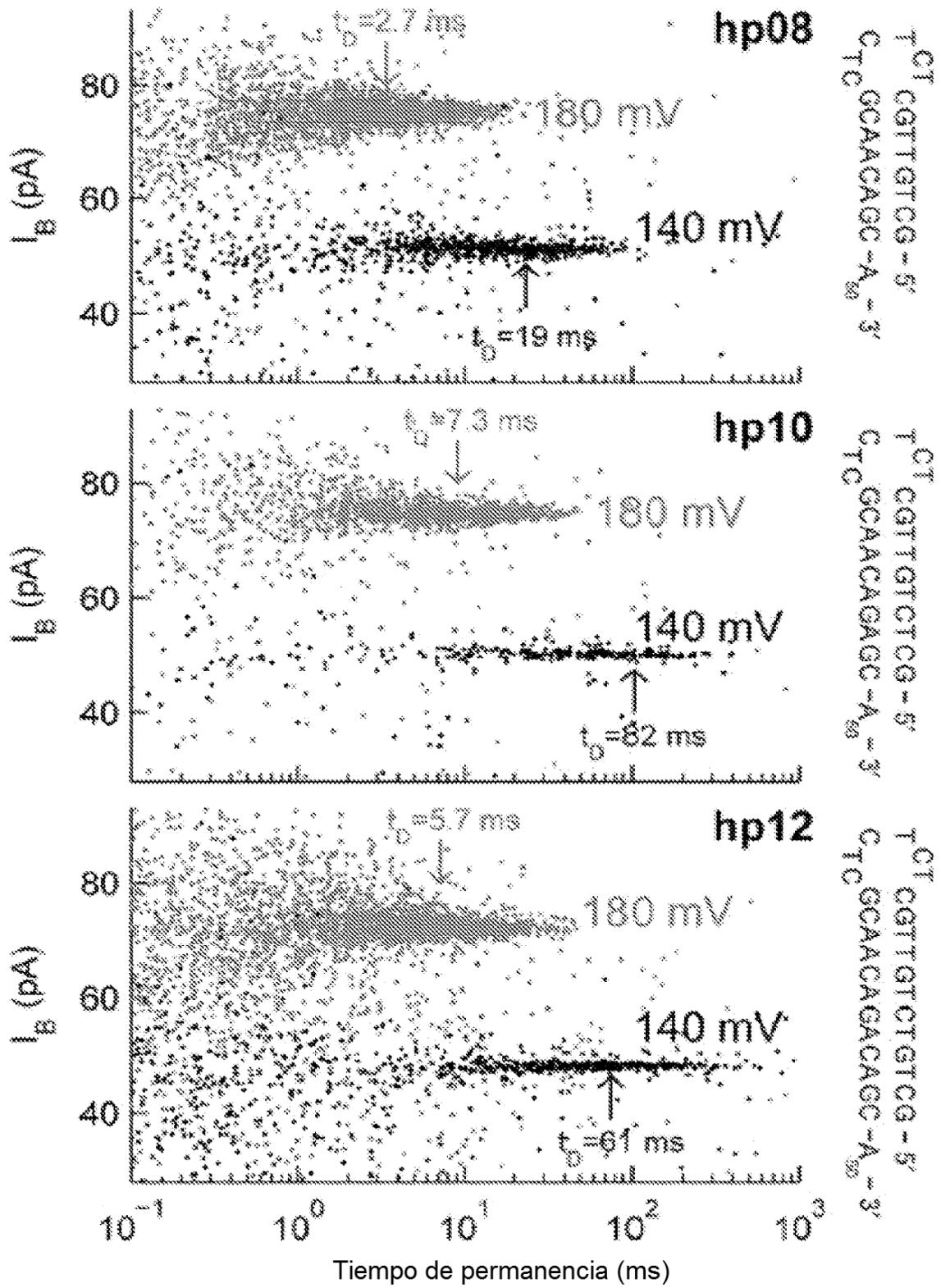


Figura 6

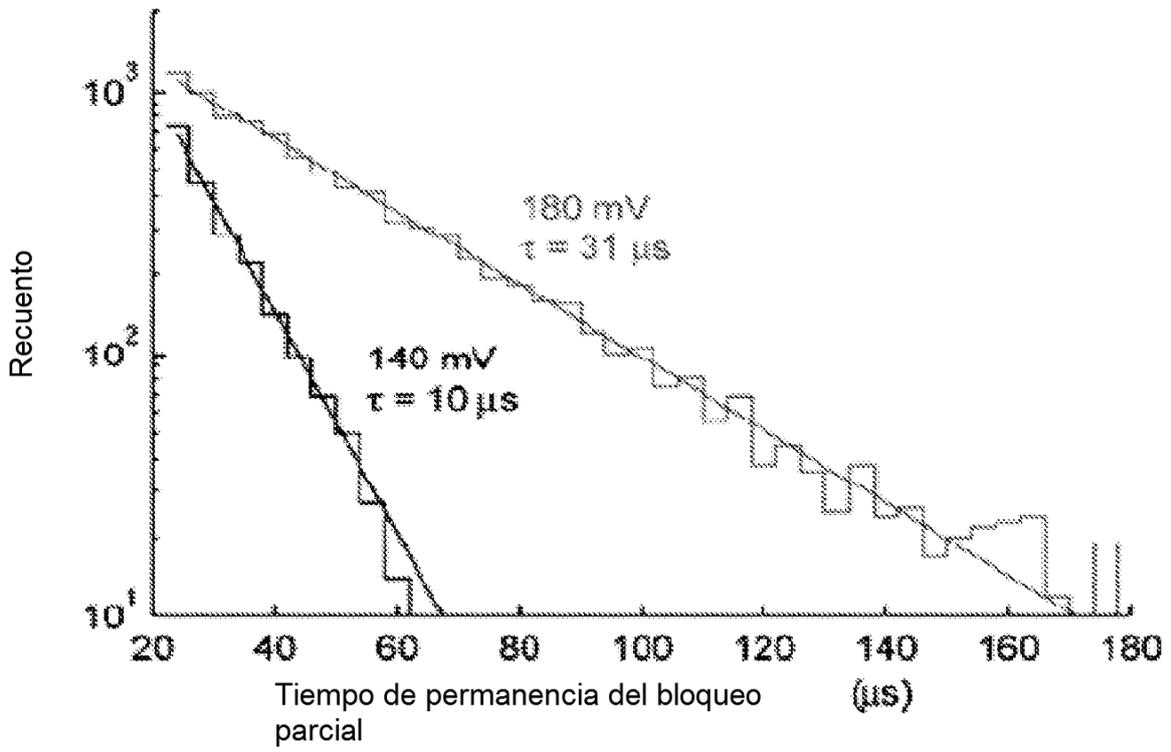


Figura 7

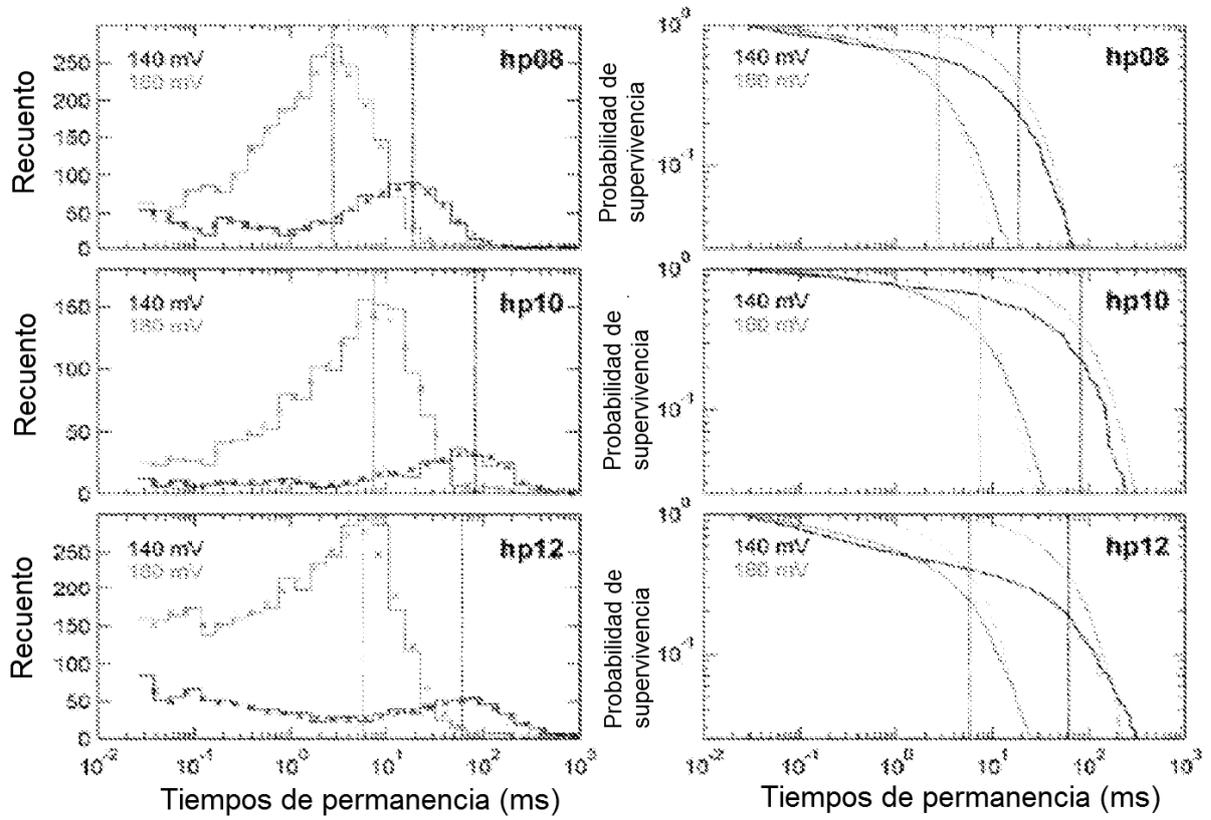
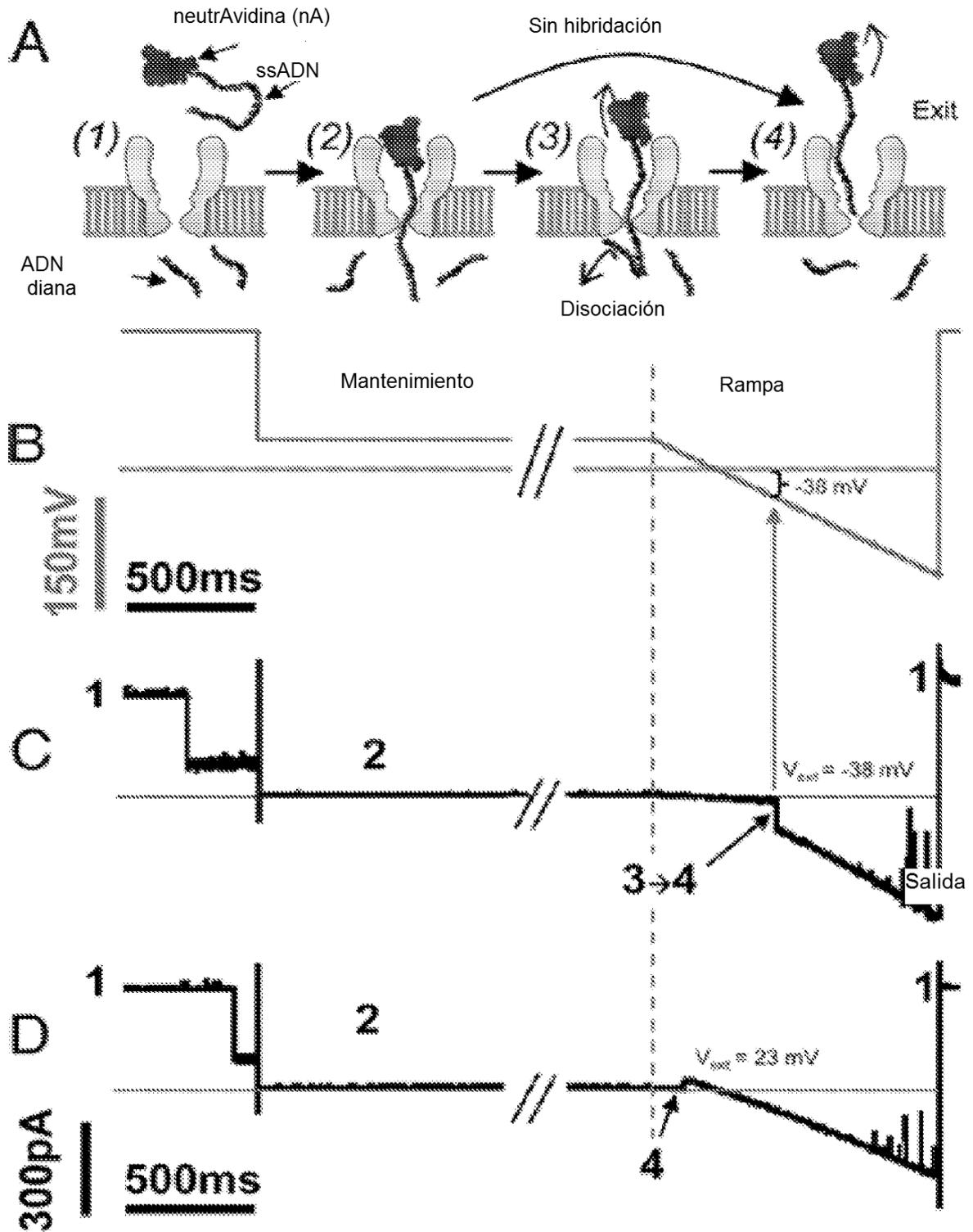
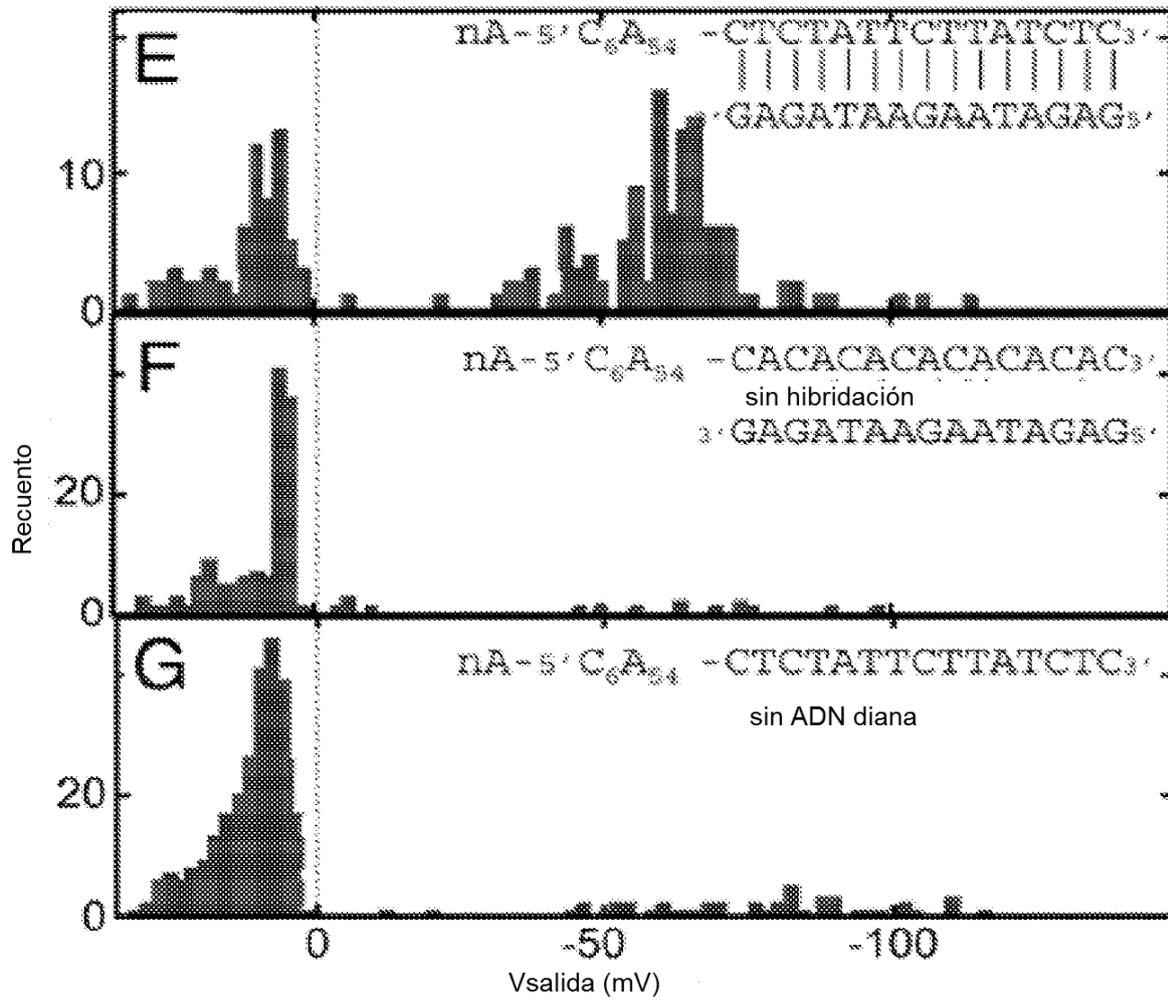


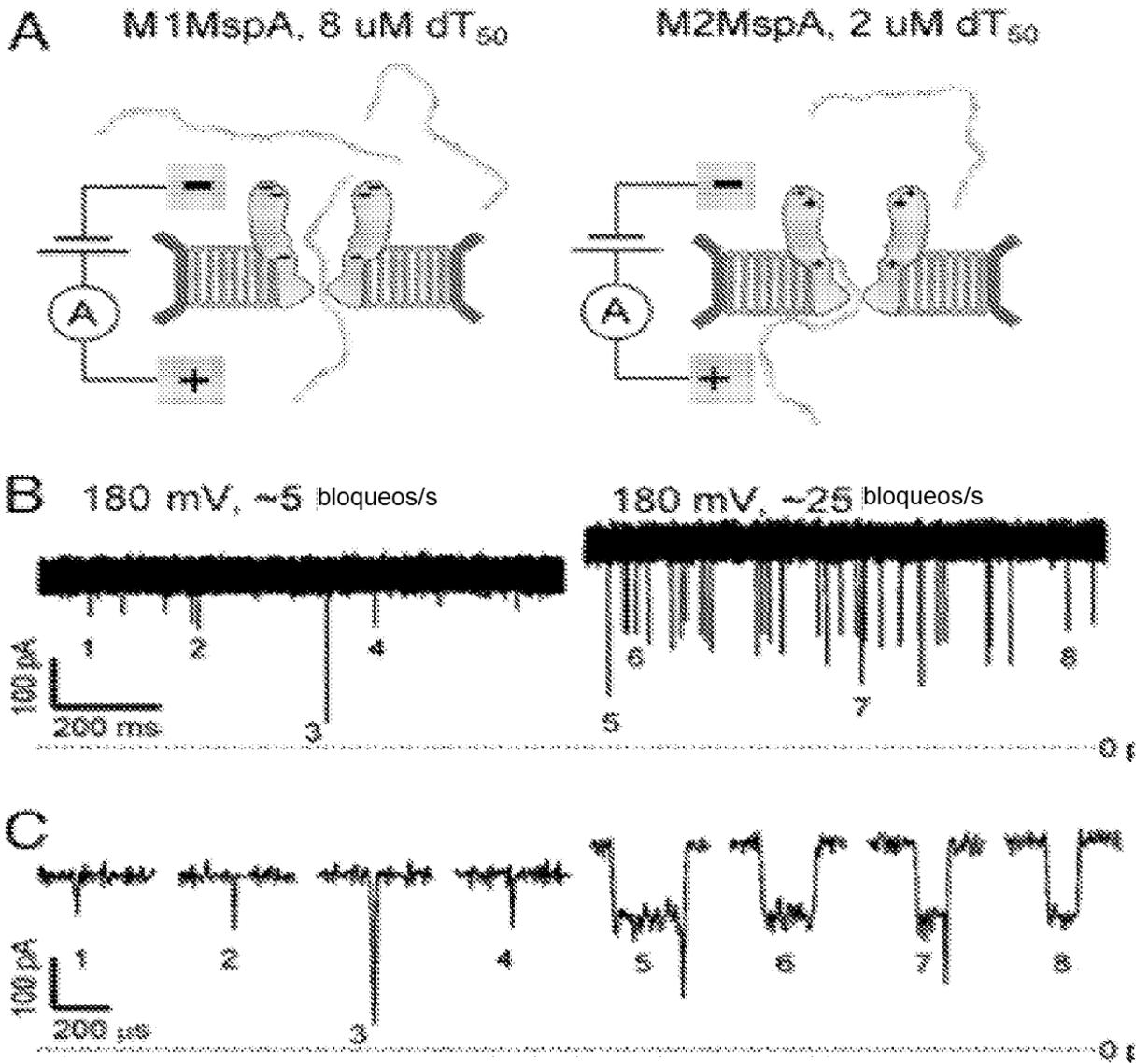
Figura 8



Figuras 9A-9D



Figuras 9E-9G



Figuras 10A-10C

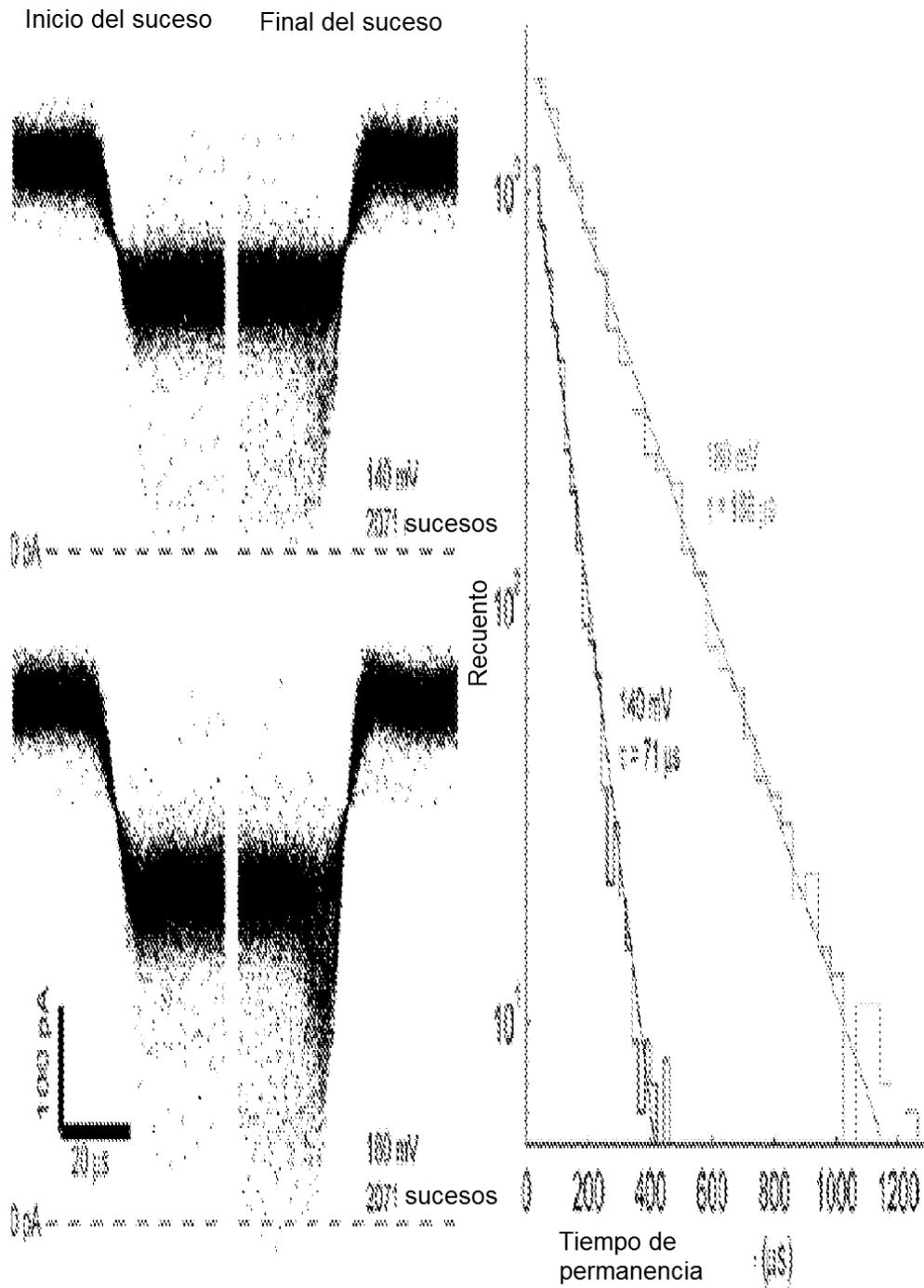


Figura 11

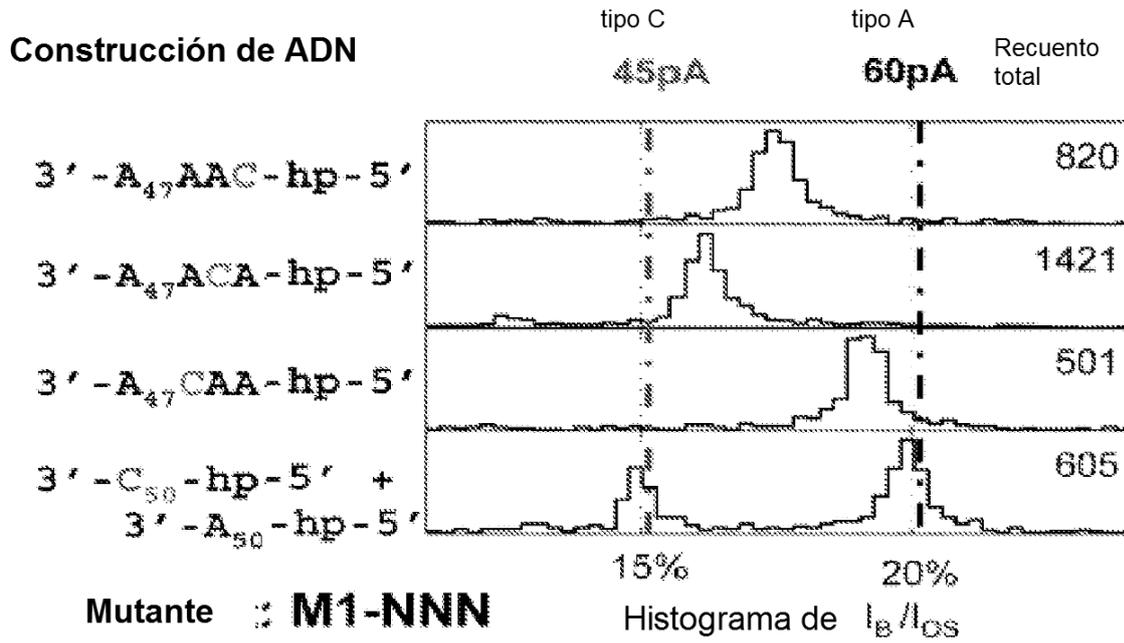


Figura 12A

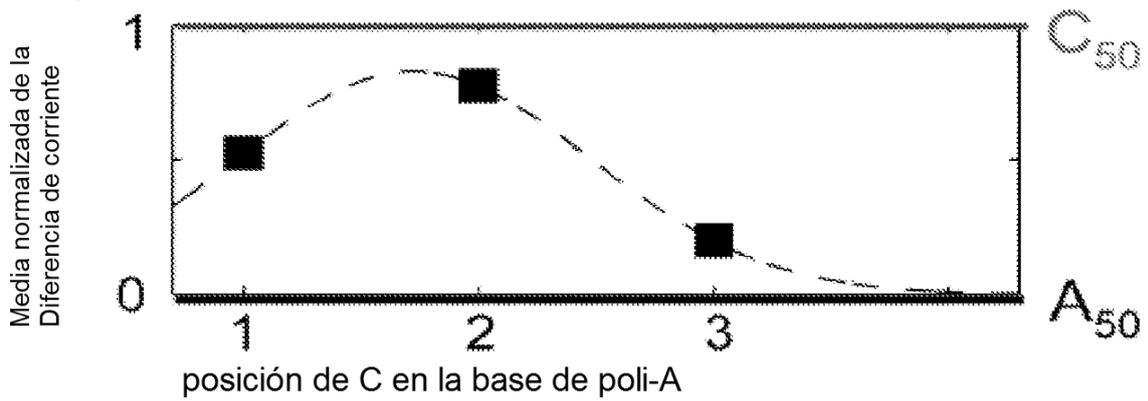


Figura 12B

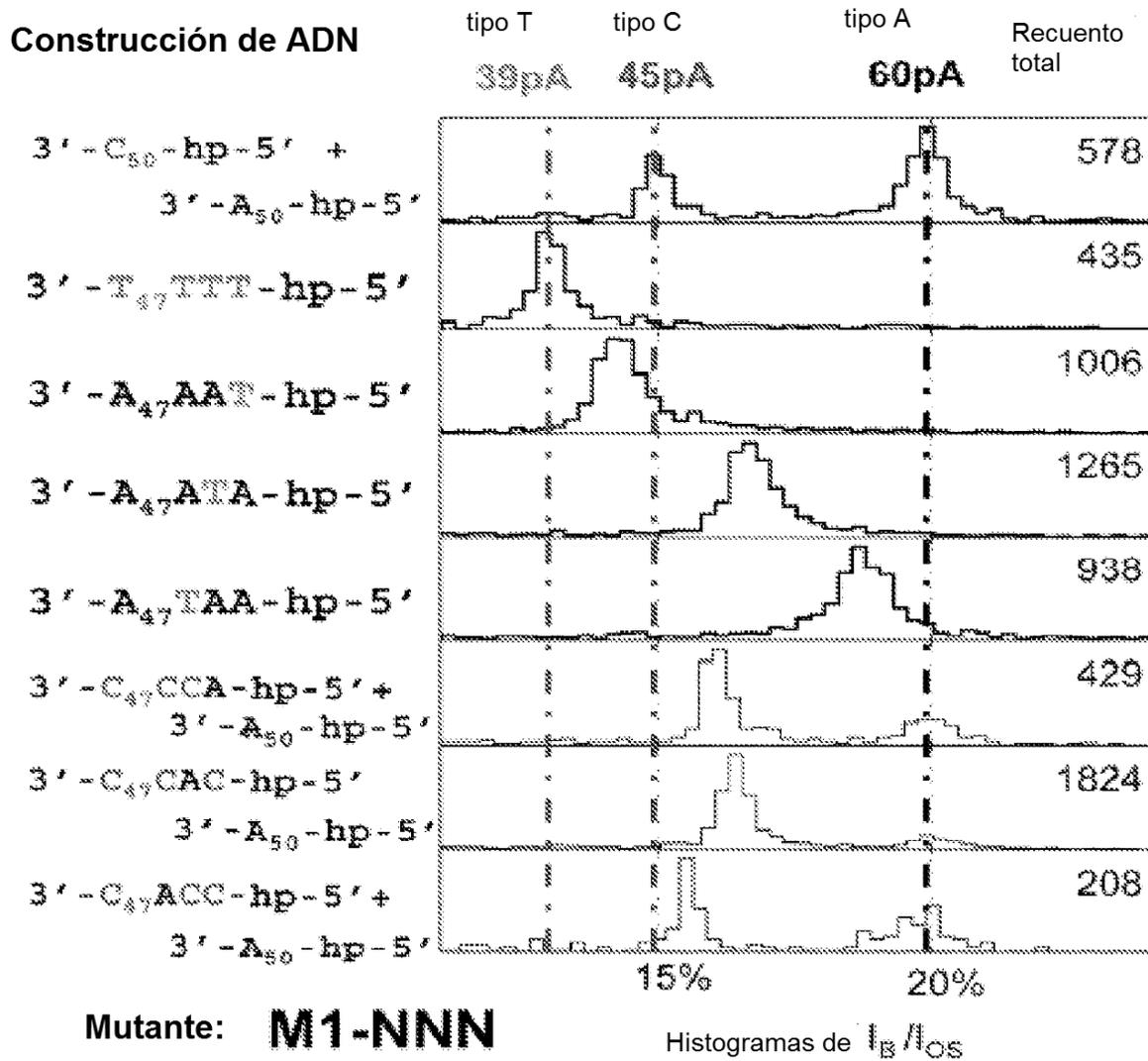


Figura 13

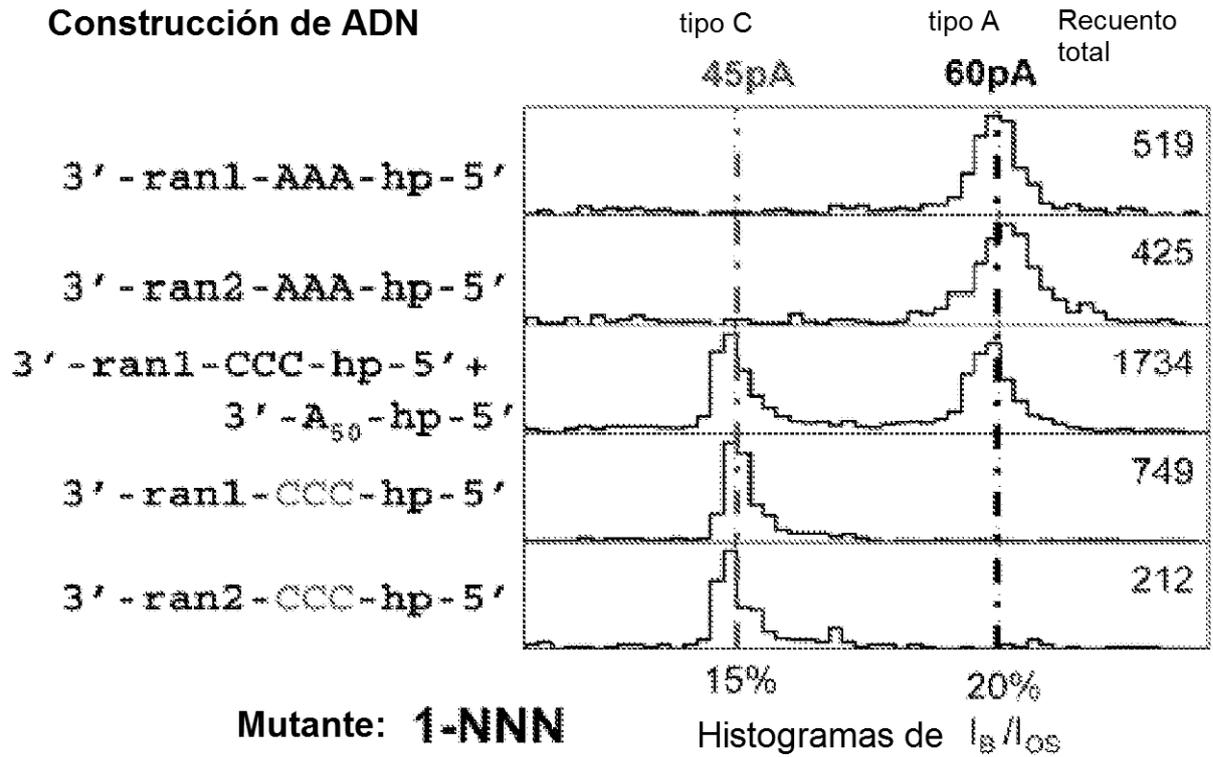
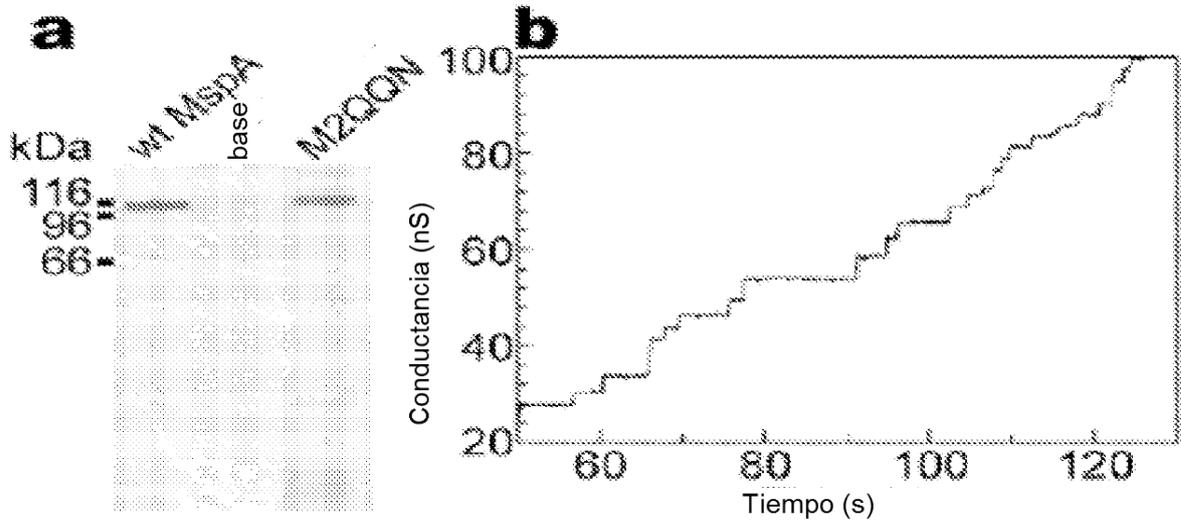


Figura 14



Figuras 15A, 15B

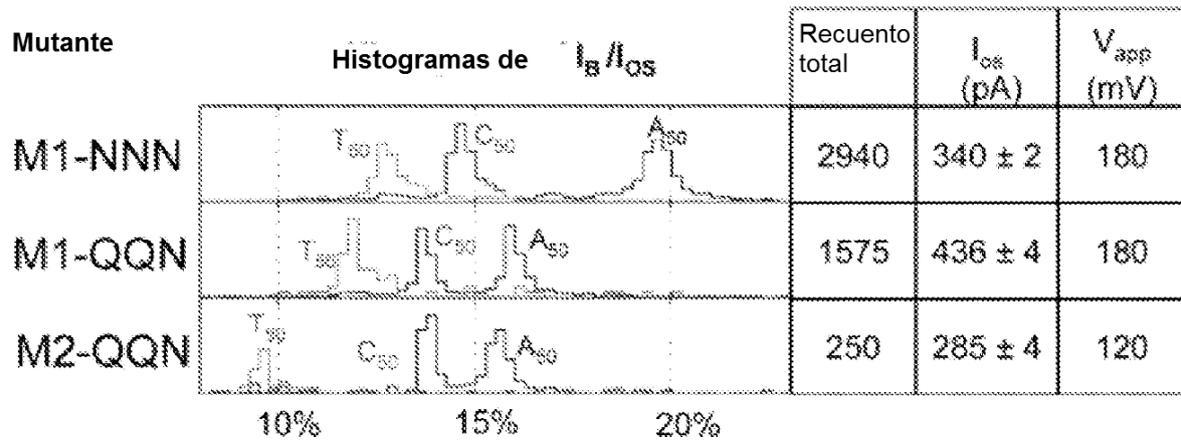


Figura 16

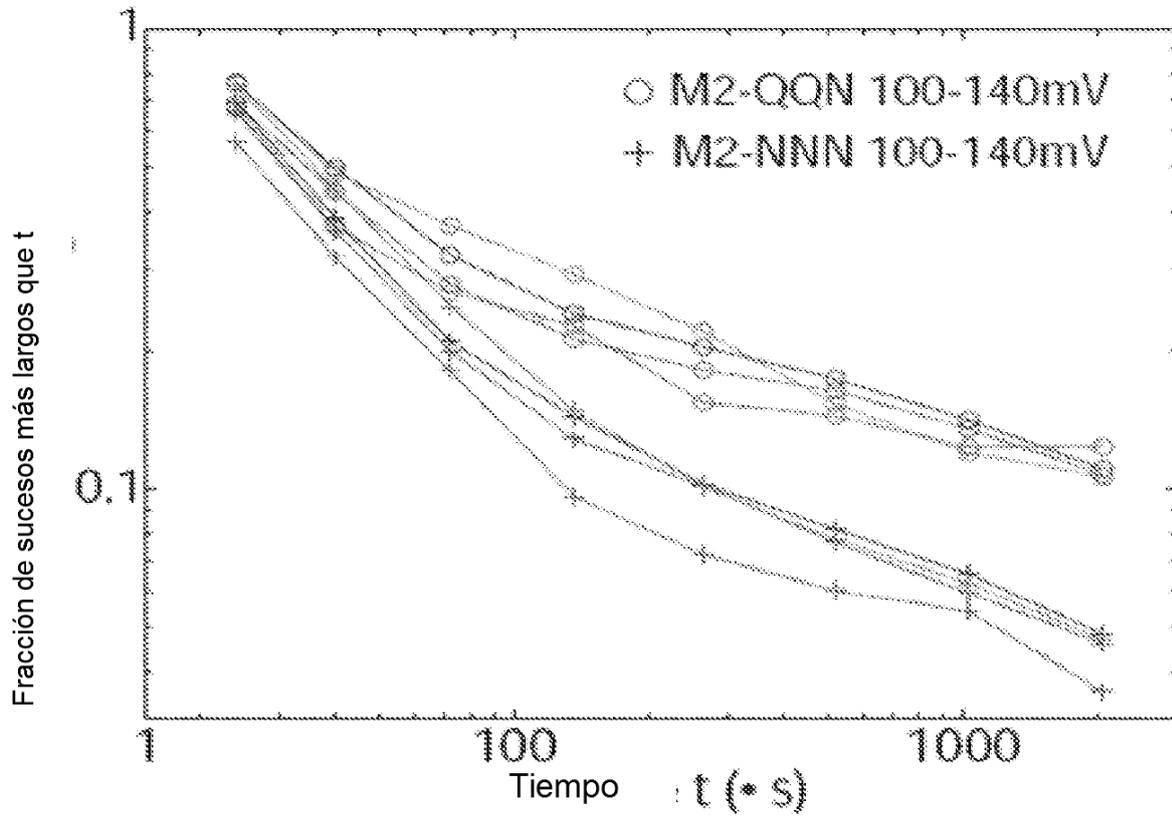


Figura 17

		1	*	*	*	*	*	
MspA	MKAI	SRVLIAMV	---	AAIAALFTSTGTSHAGLDNELSLVDGQDR	TLTVQQWDTFLNGVFP	PLDRNRLTREW		40
MspB	MIAFKR	VLIAM	SALLACTTCMFV	SAGAAHAGLDNELSLVDGQDR	TLTVQQWDTFLNGVFP	PLDRNRLTREW		40
MspC	MKAI	SRVLIAM	SALAAAVAGLFV	SAGTSHAGLDNELSLVDGQDR	TLTVQQWDTFLNGVFP	PLDRNRLTREW		40
MspD	---	MRYLVMME	---	ALLVSVTLVSPRPANA	VDNQLSVVDGQ	ERTLTVQQAETFLNGVFP	PLDRNRLTREW	39
		*	*	*	*	*	*	
MspA	FHS	GRAKYIVAGPGADEFE	GTLELGYQIGFPWSL	GVGINFSY	TPNLI	DDGDI	TAPPFGLNSVITPNLF	110
MspB	FHS	GRAKYIVAGPGADEFE	GTLELGYQIGFPWSL	GVGINFSY	TPNLI	DDGDI	TAPPFGLNSVITPNLF	110
MspC	FHS	GRAKYIVAGPGADEFE	GTLELGYQIGFPWSL	GVGINFSY	TPNLI	DDGDI	TPPFGLNSVITPNLF	110
MspD	FHS	GRATYEVAGPGADEFE	GTLELGYQIGFPWSL	GVGINFSY	TPNLI	DDGDI	TPPFGLNSVITPNLF	109
		*	*	*	*	*	*	
MspA	PGV	SISADLNGPGIQEVAT	FSDVSGAEGG	VAVSNAHGT	VTGAAGGV	LLRPFARLI	ASTGDSVTTYGEPWNMN	184
MspB	PGV	SISADLNGPGIQEVAT	FSDVSGPAGG	VAVSNAHGT	VTGAAGGV	LLRPFARLI	ASTGDSVTTYGEPWNMN	184
MspC	PGV	SISADLNGPGIQEVAT	FSDVSGPAGG	VAVSNAHGT	VTGAAGGV	LLRPFARLI	ASTGDSVTTYGEPWNMN	184
MspD	PGV	SISADLNGPGIQEVAT	FSDVSGAKGAG	VAVSNAHGT	VTGAAGGV	LLRPFARLI	ASTGDSVTTYGEPWNMN	183

Figura 18

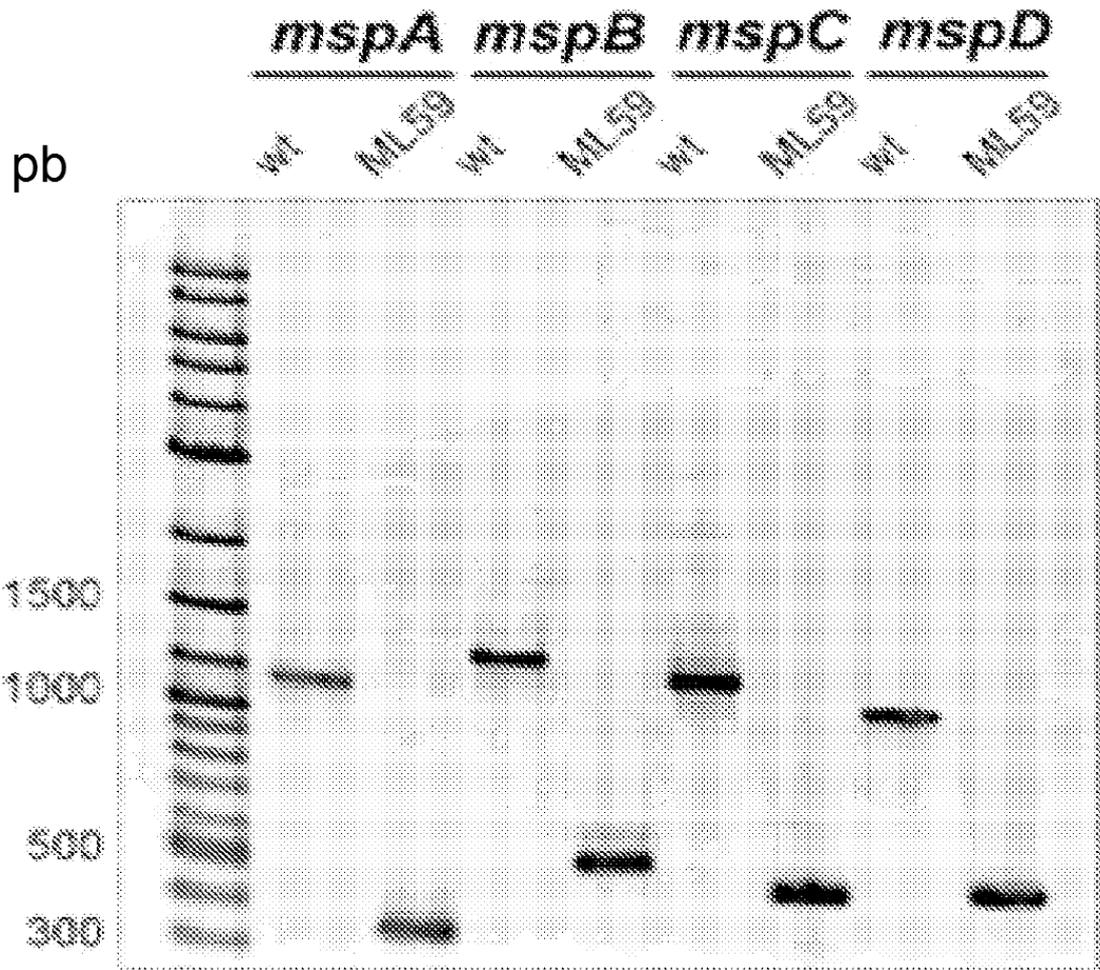


Figura 19

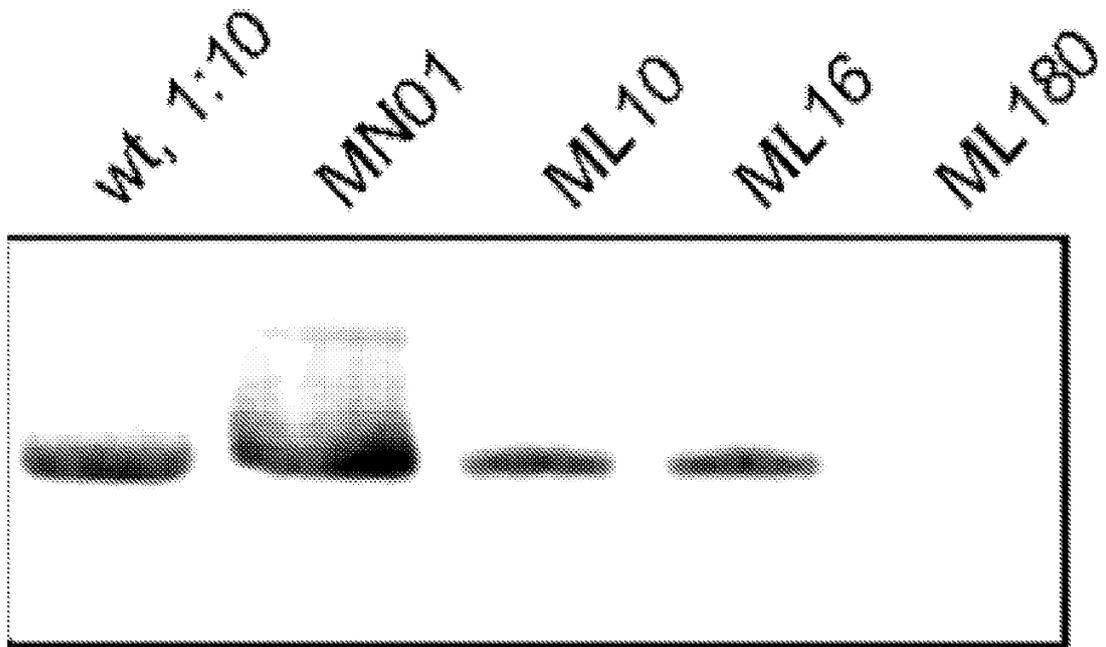


Figura 20

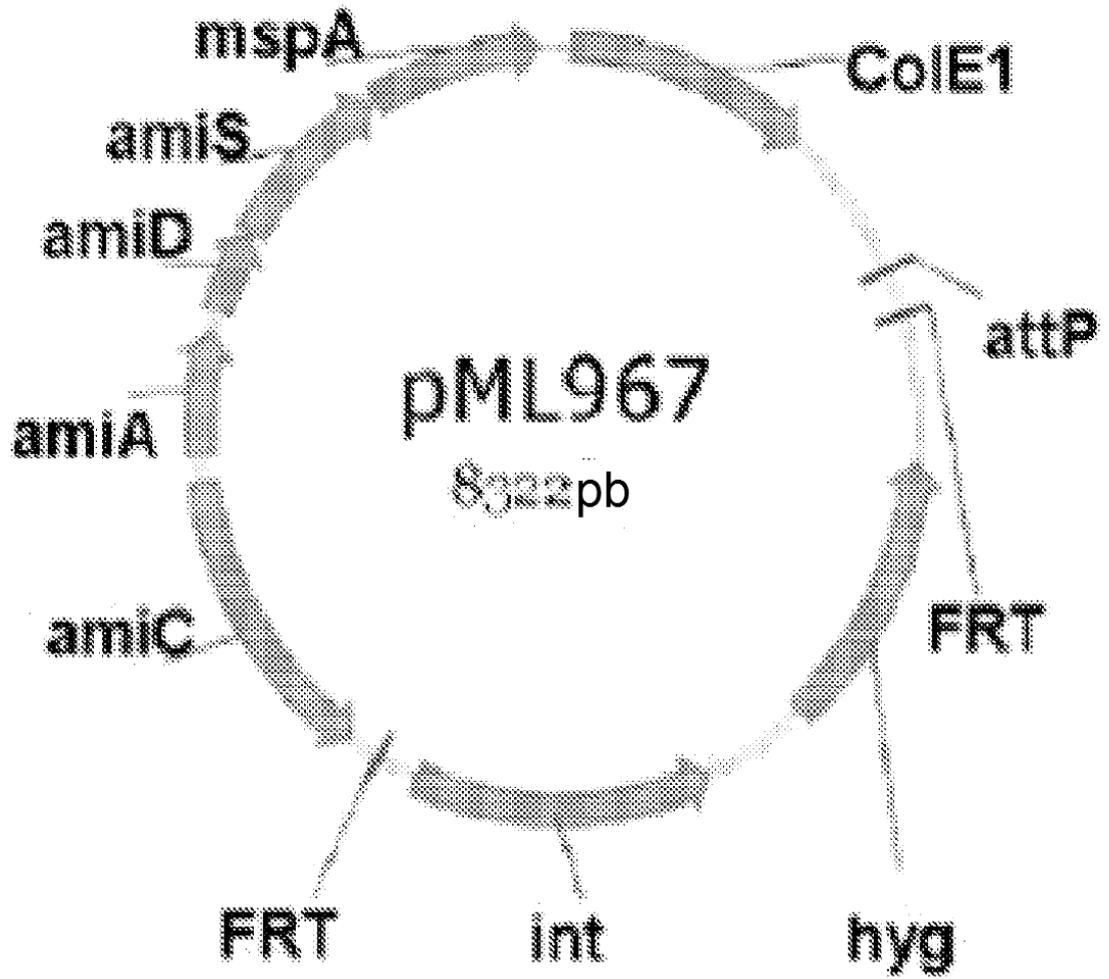


Figura 21A

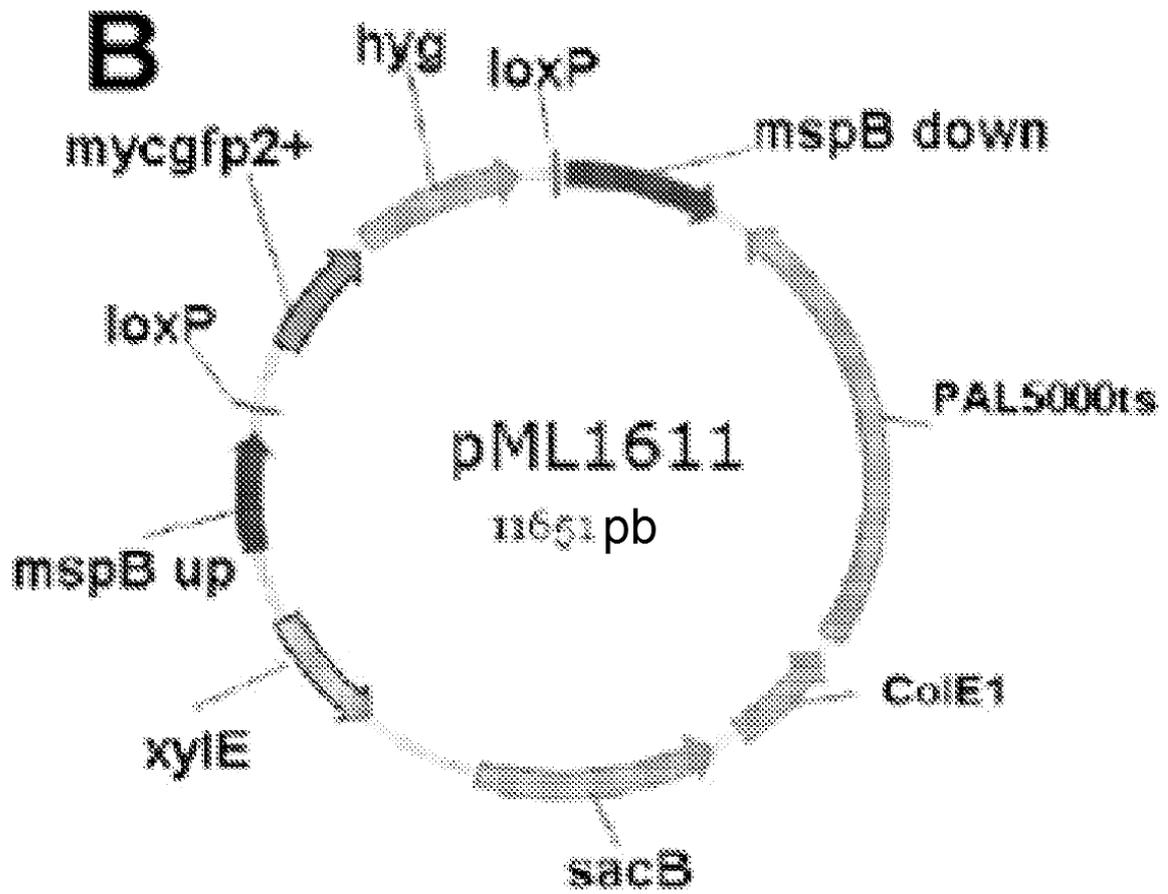


Figura 21B

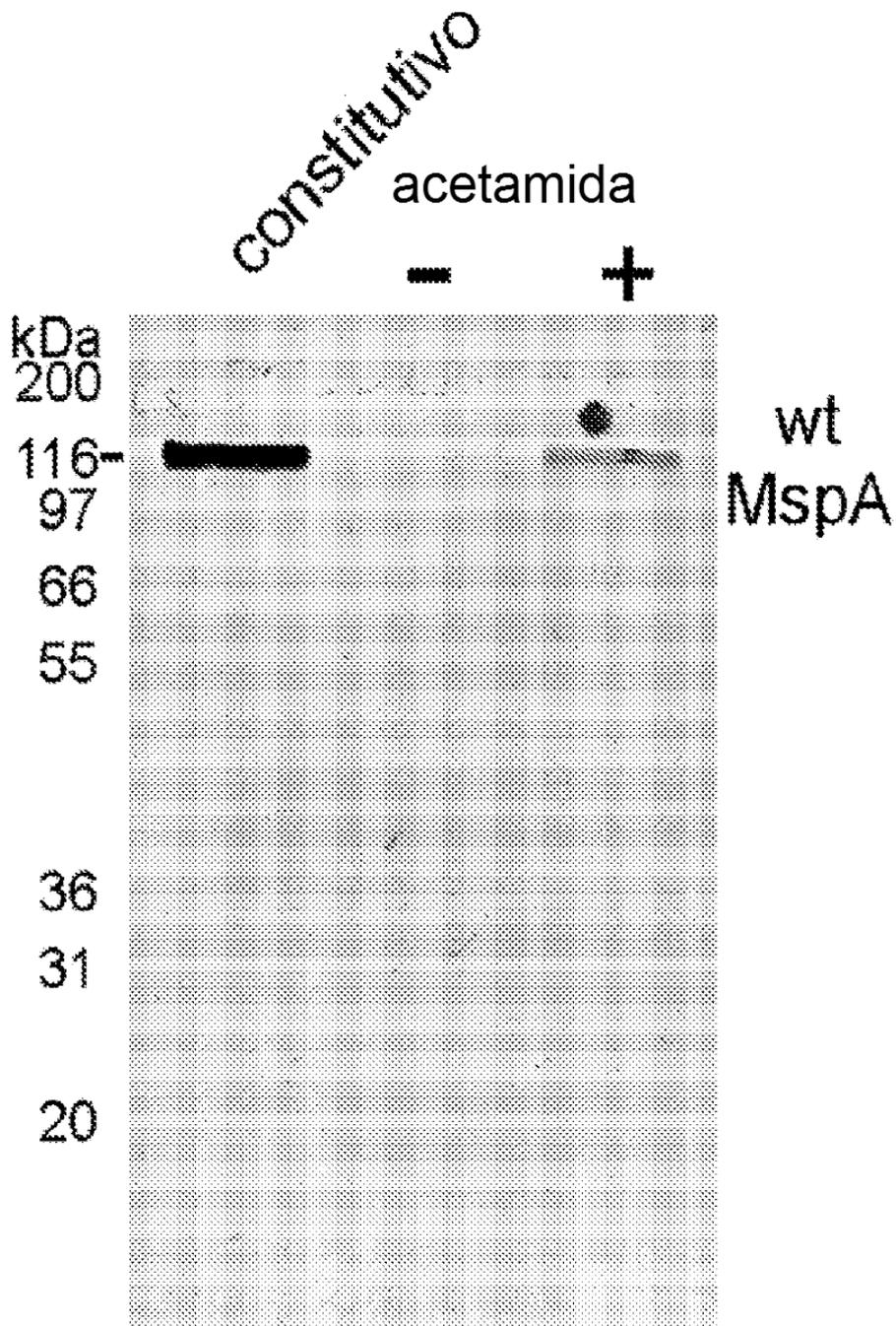


Figura 22

wt *M. smegmatis*

ML705

ML705 + MspA

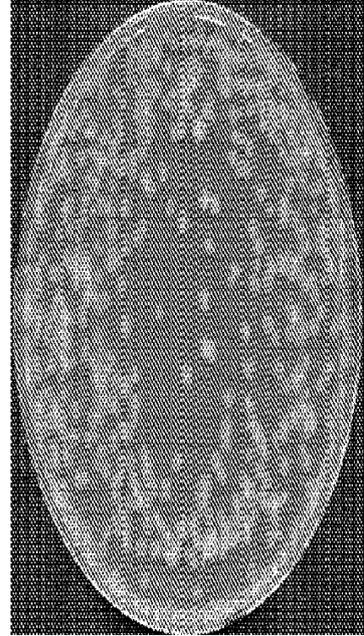
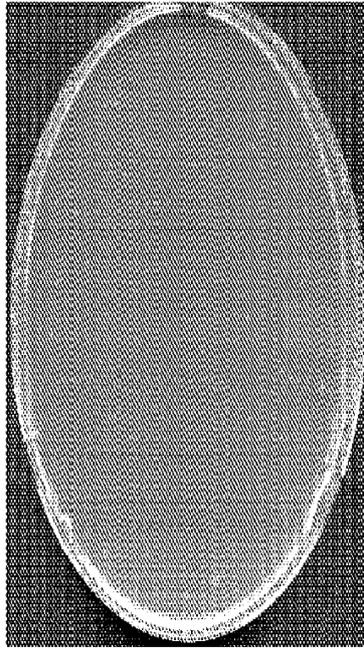
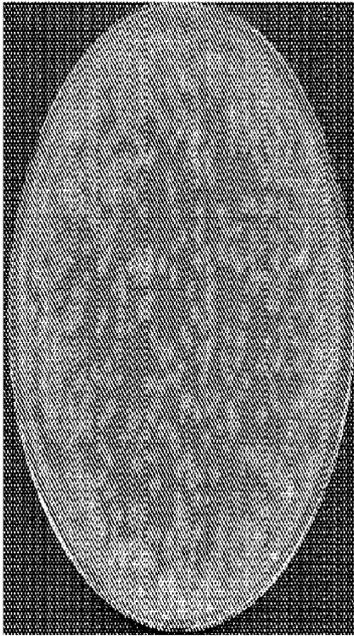


Figura 23

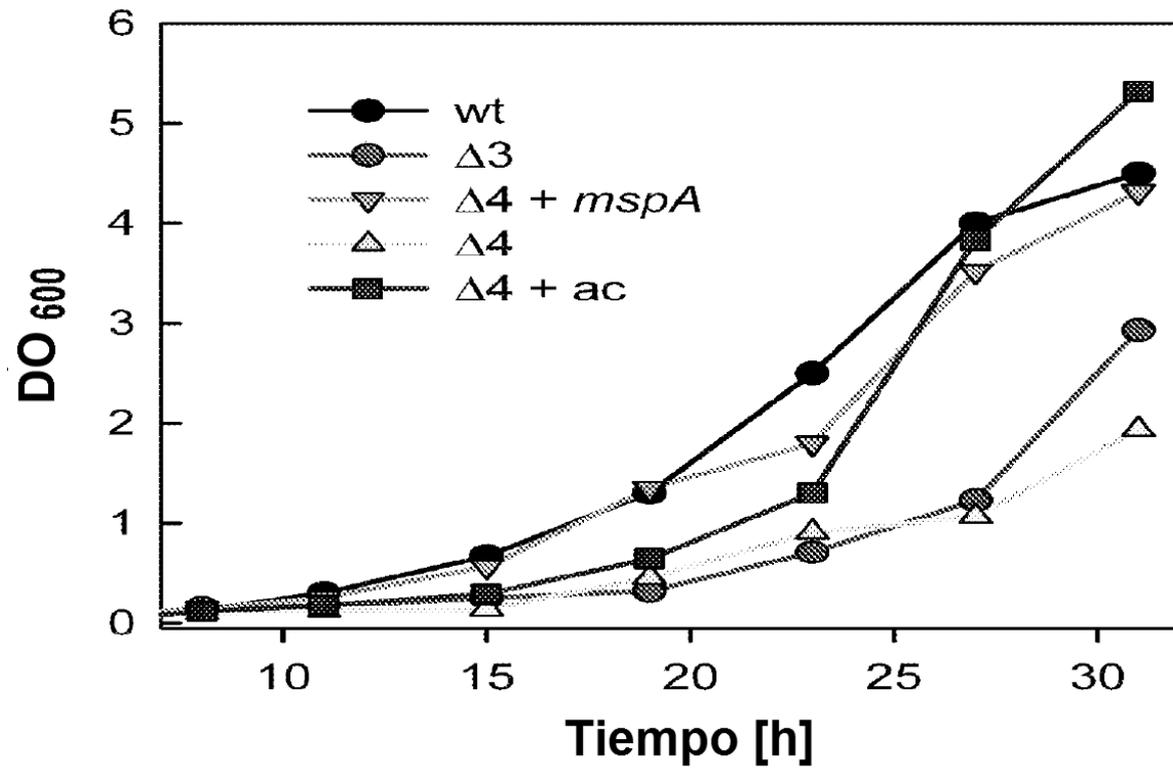


Figura 24

wt **$\Delta 4+ac$** **$\Delta 4$** **$\Delta 3$**

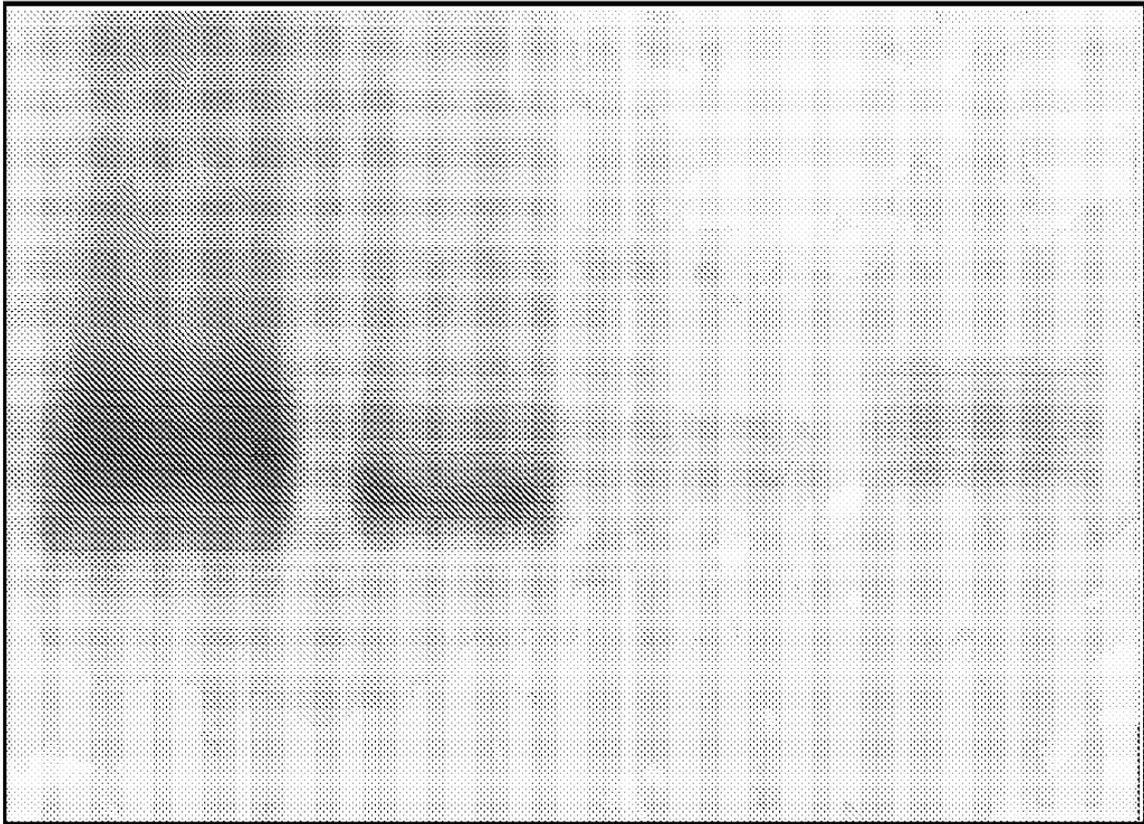


Figura 25

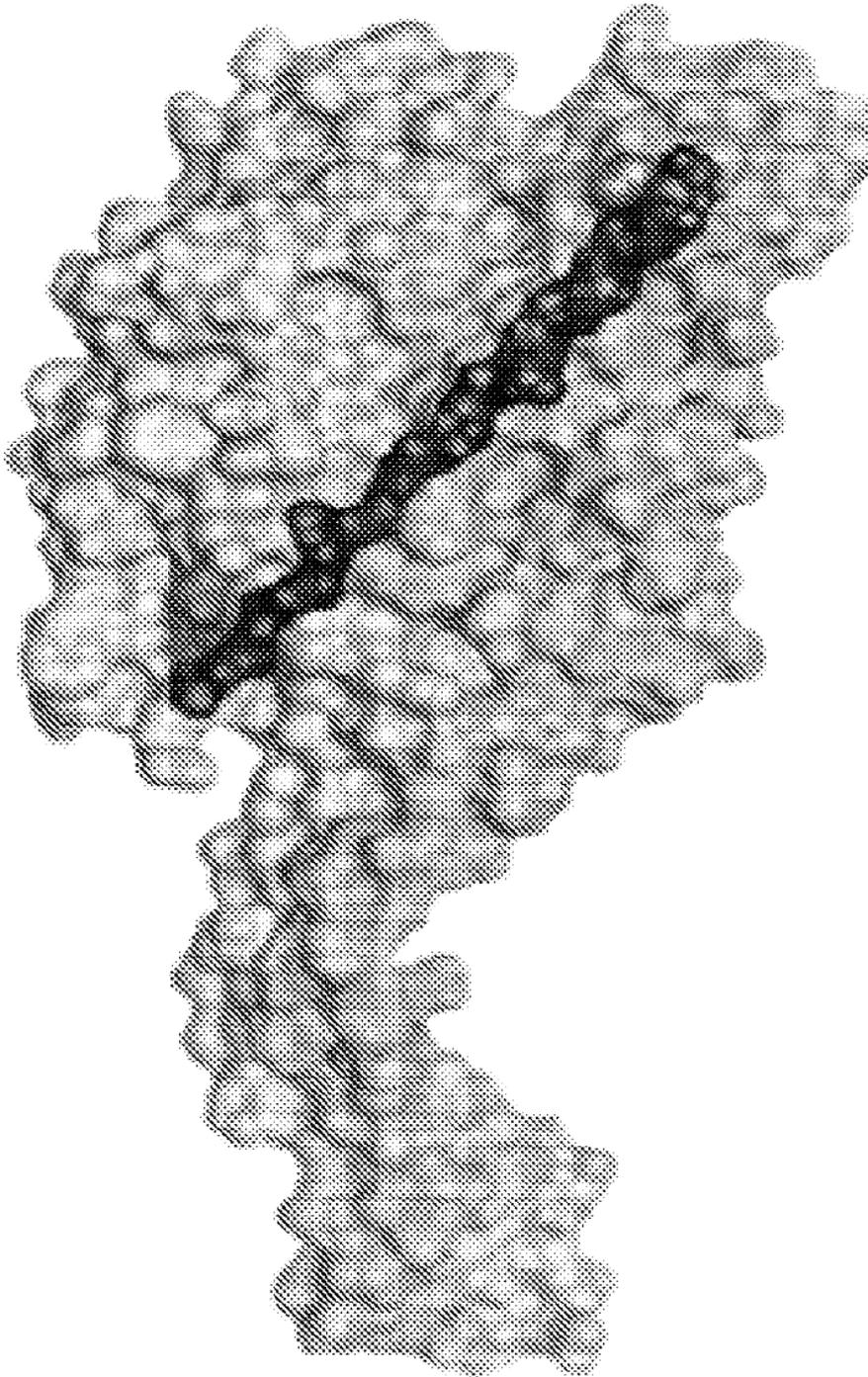


Figura 26A

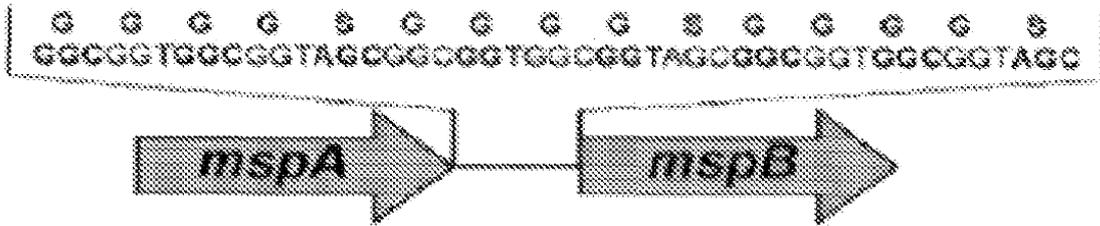


Figura 26B

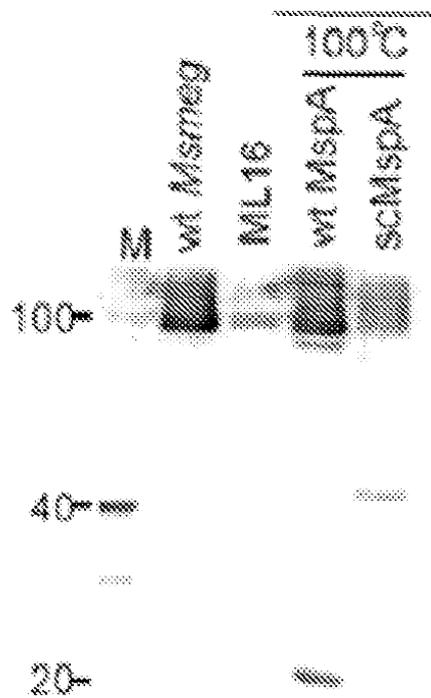


Figura 26C

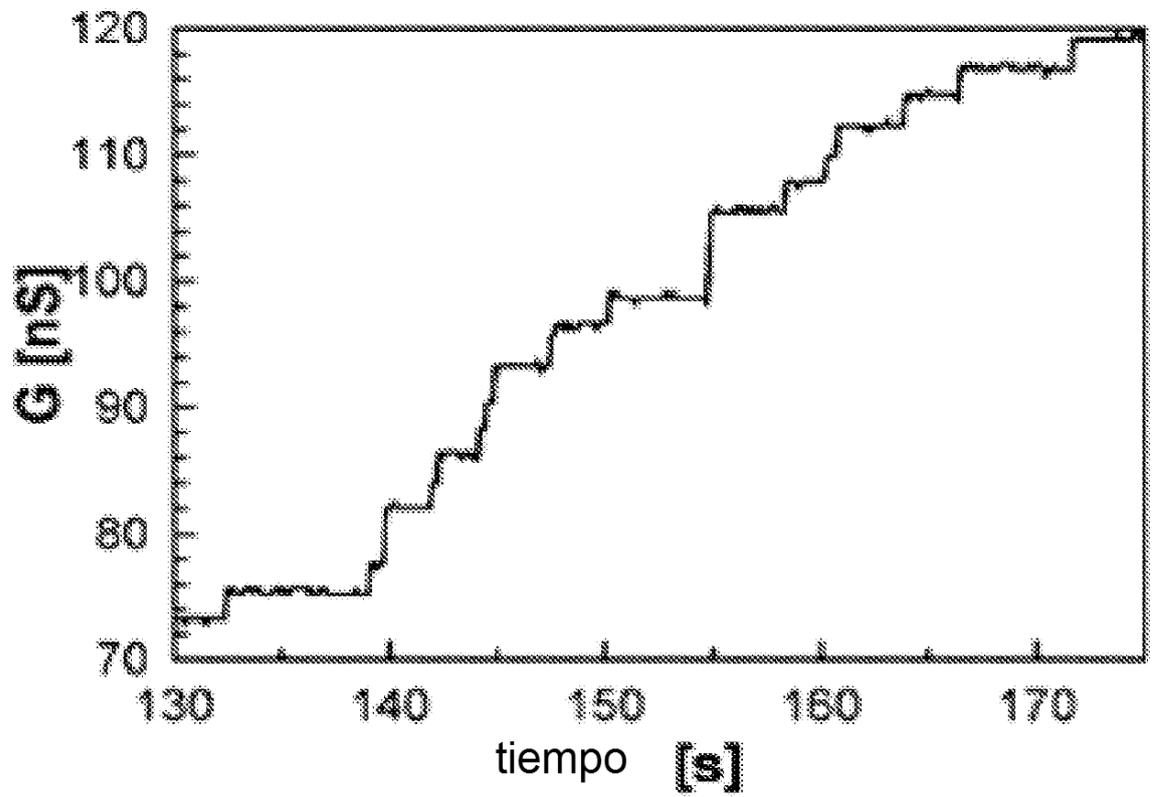


Figura 26D

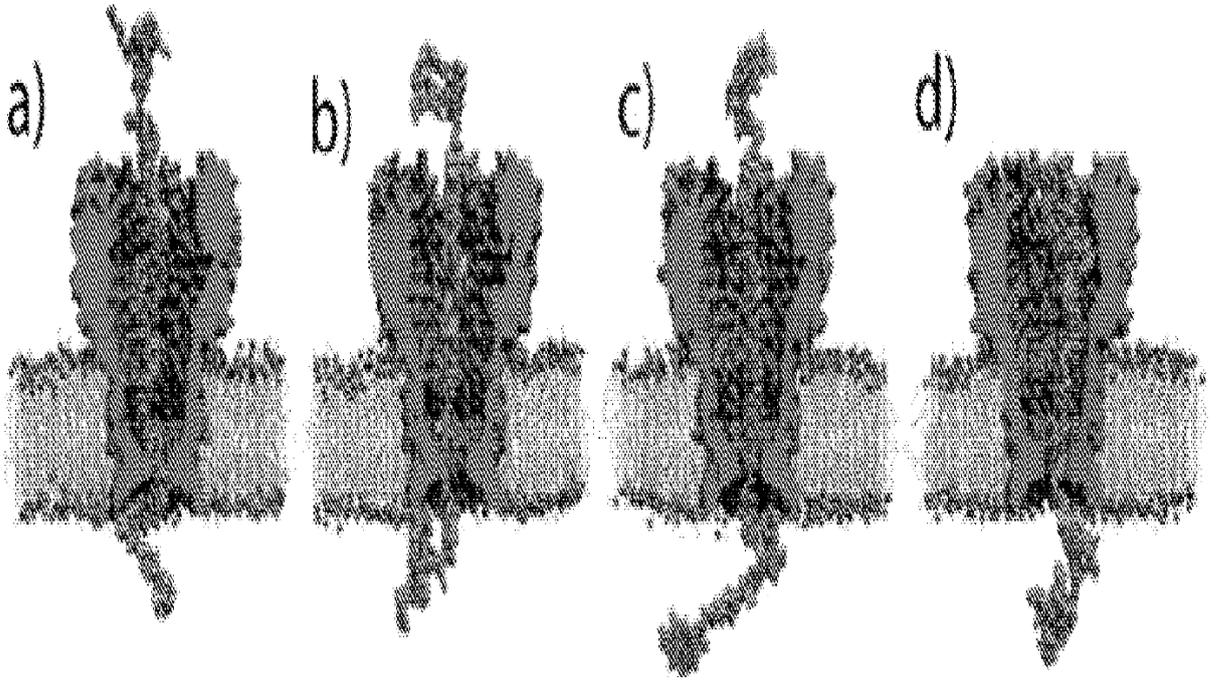


Figura 27

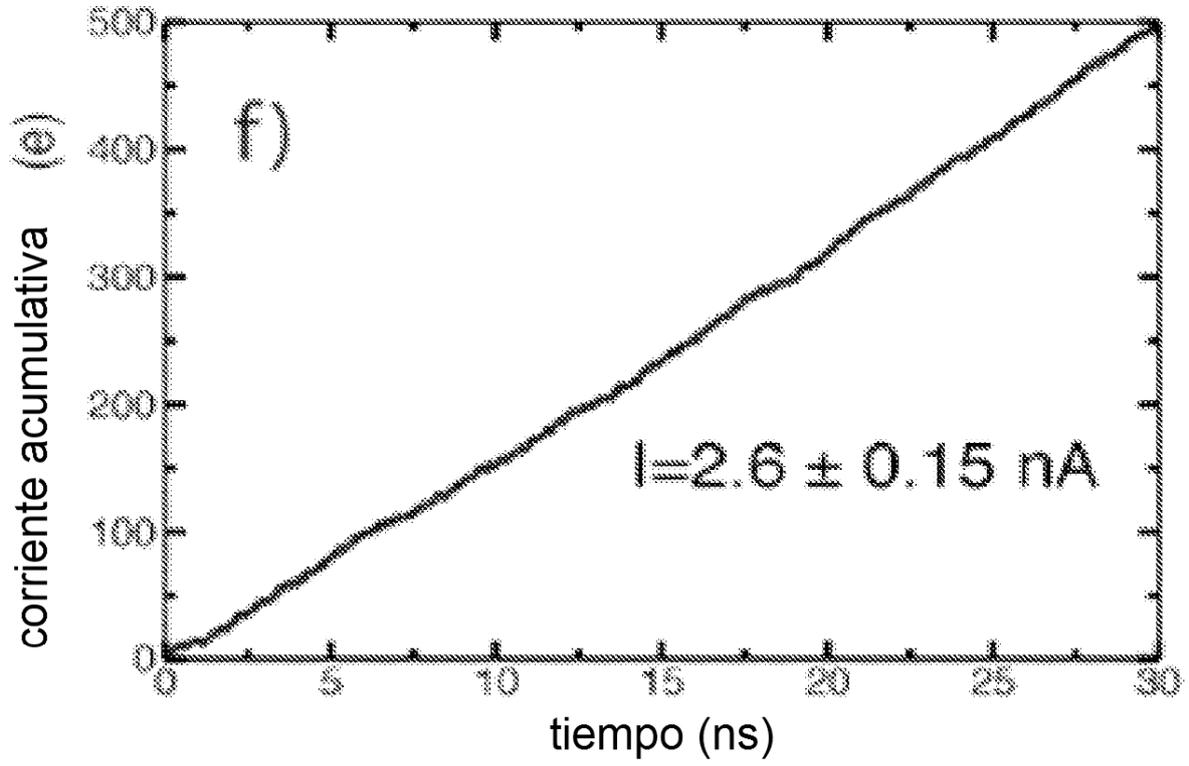


Figura 28

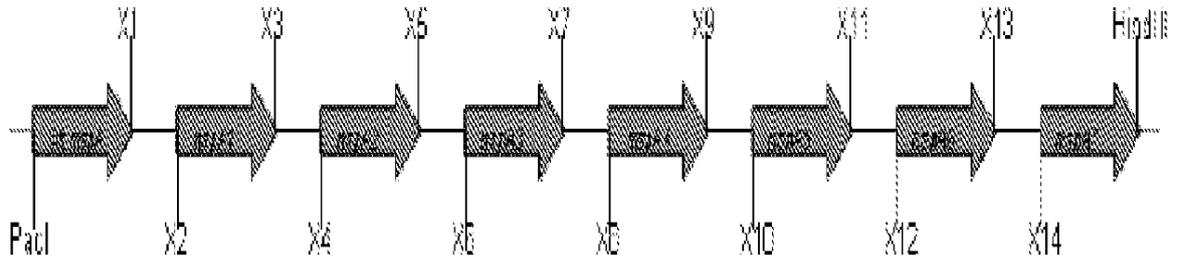


Figura 29

```

MspA      LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTAPPFG-----LNSVITPN-----LFPGVSI SADLGNQ
PorM1     LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
PorM1_2   LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
MspB      LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTAPPFG-----LNSVITPN-----LFPGVSI SADLGNQ
MspC      LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
PorM2     LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mmcs_4296 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mmcs_4382 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mmcs_4297 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mmcs_4383 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mjls_4674 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mjls_4675 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mjls_4677 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mvan_1836 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mvan_4839 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mvan_4840 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
Mflv_1891 LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTGPPFG-----LESVITPN-----LFPGVSI SADLGNQ
MspD      LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTAPPFG-----LDTIITPN-----LFPGVSI SADLGNQ
MCH_4691c LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTAPPFG-----LDTIITPN-----LFPGVSI SADLGNQ
MAB_1080  LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTAPPFG-----LDTIITPN-----LFPGVSI SADLGNQ
MCH_4690c LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTAPPFG-----LDTIITPN-----LFPGVSI SADLGNQ
MCH_4689c LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTAPPFG-----LDTIITPN-----LFPGVSI SADLGNQ
MAB_1081  LELGYQIGFPWLSLGVGINFSYTTNLIIDDDGDTAPPFG-----LDTIITPN-----LFPGVSI SADLGNQ
ro_08561  LTVGYQIGYPIALS-GATIVLNSPGLGFVIGSSNGIDLGLVPTPTLDLHAGSNVGLAGDII PSQELDIDLAPG
ro_03127  LTVGYQIGYPIALS-GATIVLNSPGLGFVIGSSNGINLGLPDPSSLNAGTNVGLAGDII PSQELDIDLAPG
n.d.      LTVGYQIGYPIALS-GATIVLNSPGLGFVIGSSNGIDLGLLPELSLGLNAGTNVELAGDII PSQELDIDLAPG
ro_04074  LTVGYQIGYPIALS-GATIVLNSPGLGFVIGSSNGINLGLVDPDFTLDLNAGTNAELAGDII PSQELDIDLAPG
Mmcs_3857 LTVGYQIGYPIALS-GATIVLNSPGLGFVIGSSNGINLGLVDPDFTLDLNAGTNAELAGDII PSQELDIDLAPG
Mjls_3857 LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
Mjls_3931 LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
Mjls_3843 LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
Mvan_4117 LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
Mflv_2295 LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
MAP3_123c LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
Mav_3943  LEVGYEIGCGIDMSTSDGVTIGGTAGITPGVTEPVTGVPGD----VLPLVVAP-----IAGVLNVGLKPG
MUL_2391  LEVGYEIGCGIDMSTSDGVTIGGTAGITPGVTEPVTGVPGD----VLPLVVAP-----IAGVLNVGLKPG
Mvan_5016 LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
Mflv_1735 LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
MAB_2800  LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
Mvan_5017 LEVGYEIGCGIDMSTSNVSLTGTAGLNPSLGVIGTDVISPDPGILPGIGGN-----IGGGVTVGLKPG
Mflv_1734 LTVGYQIGYPIALS-GATIVLNSPGLGFVIGSSNGIDLGLLPELSLGLNAGTNVELAGDII PSQELDIDLAPG
Mvan_5768 LTVGYQIGYPIALS-GATIVLNSPGLGFVIGSSNGIDLGLLPELSLGLNAGTNVELAGDII PSQELDIDLAPG

```

Figura 30

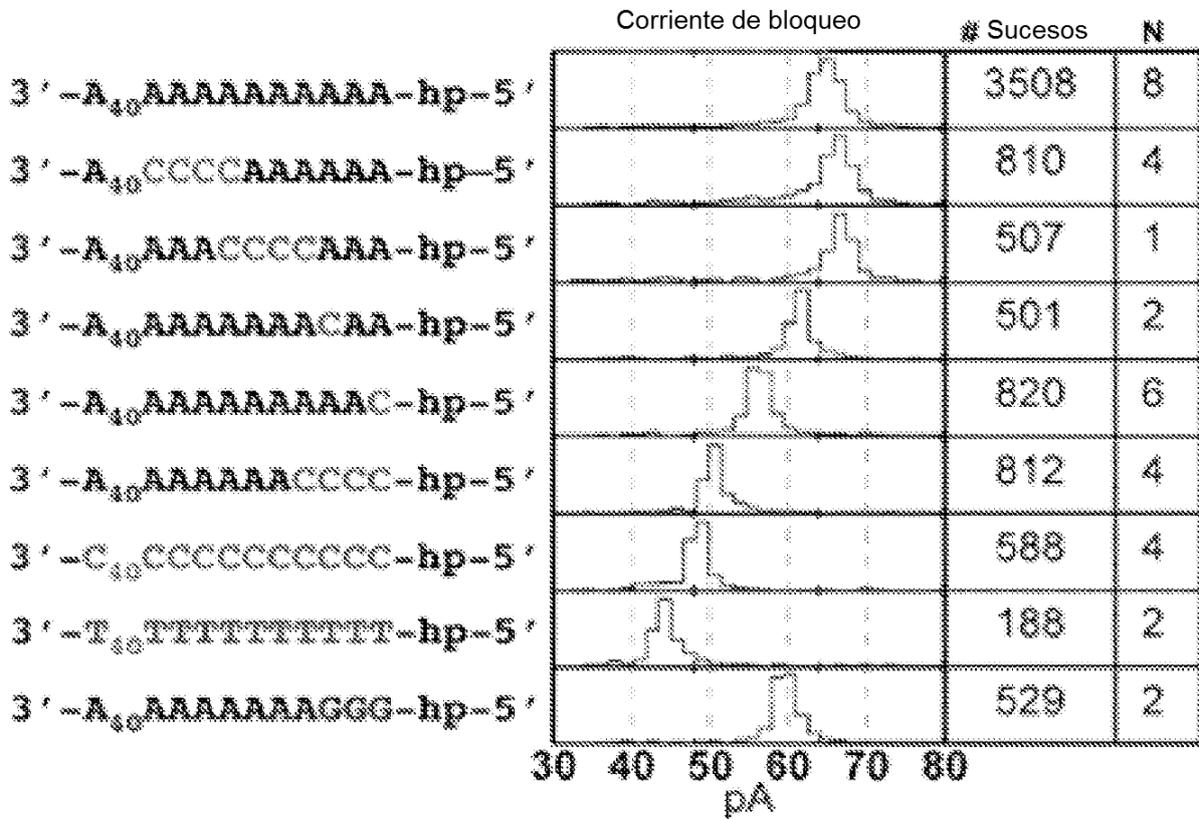


Figura 31