

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 978 413**

51 Int. Cl.:

**B01D 15/16** (2006.01)

**G01N 30/86** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.02.2019 PCT/US2019/019050**

87 Fecha y número de publicación internacional: **29.08.2019 WO19165148**

96 Fecha de presentación y número de la solicitud europea: **21.02.2019 E 19710843 (4)**

97 Fecha y número de publicación de la concesión europea: **06.03.2024 EP 3755447**

54 Título: **Condiciones determinantes para la purificación de proteínas**

30 Prioridad:

**21.02.2018 US 201862633584 P**

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

**12.09.2024**

73 Titular/es:

**JUST-EVOTEC BIOLOGICS, INC. (100.0%)  
401 Terry Avenue North  
Seattle, WA 98109, US**

72 Inventor/es:

**SHAVER, JEREMY, MARTIN;  
AMIMEUR, TILELI;  
GILLESPIE, RON;  
KETCHEM, RANDAL, R. y  
GARCIA, FERNANDO**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 978 413 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

Condiciones determinantes para la purificación de proteínas

### Referencia cruzada a solicitudes relacionadas

5 Esta solicitud reclama prioridad para la solicitud provisional en EE. UU. N.º 62/633.584 presentada el 21 de febrero de 2018 y titulada "Determining Conditions for Purification of Proteins".

### Antecedentes

10 Las proteínas están compuestas por una secuencia de aminoácidos que están unidos mediante enlaces químicos. La secuencia de aminoácidos de una proteína particular se basa en una secuencia de nucleótidos en el ácido desoxirribonucleico (ADN) a partir del cual se expresa la proteína. La funcionalidad y estructura de una proteína se puede basar en la secuencia de aminoácidos de la proteína. Las proteínas pueden tener una variedad de funciones dentro de un organismo, como la regulación de la actividad enzimática o la señalización celular. Algunas proteínas también se pueden usar terapéuticamente para tratar una afección biológica. Por ejemplo, proteínas, como un anticuerpo, pueden, en algunos casos, unirse a un patógeno para atacarlo y destruirlo por otros agentes del organismo, como células T o macrófagos. En otro ejemplo, las proteínas se pueden unir a una molécula para transportarla a una ubicación dirigida en un organismo para aliviar los fenotipos de una afección biológica.

15 Las proteínas se pueden expresar a través de una variedad de procesos, pero las proteínas expresadas están a menudo contenidas en una solución junto con impurezas asociadas con la expresión de las proteínas. La purificación de las proteínas expresadas puede tener lugar utilizando un número de técnicas. La optimización del proceso de purificación de proteínas puede ser una tarea costosa y que requiere mucho tiempo ya que pueden ser necesarias muchas rondas de experimentación utilizando diferentes tipos de técnicas y condiciones cromatográficas para identificar las condiciones de purificación que proporcionan los mayores rendimientos y pureza.

20 El documento US 2002/010566 A1 describe un método para optimizar las condiciones para la purificación cromatográfica por HPLC de un soluto, en particular prediciendo el ancho del pico del soluto que tiene en cuenta la compresión del pico. El método descrito no está diseñado explícitamente para proteínas o biomoléculas, simplemente un "soluto".

25 El documento US 2016/367911 A1 describe un método para optimizar las condiciones para la purificación cromatográfica de una proteína diana generando un modelo creado a partir de datos obtenidos sometiendo dicha proteína diana a una variedad de condiciones cromatográficas.

30 El documento US 2017/322190 A1 describe un método para optimizar las condiciones para la purificación cromatográfica de una proteína diana generando un modelo creado a partir de datos obtenidos sometiendo dicha proteína diana a una variedad de condiciones cromatográficas.

35 El documento US 2006/096924 A1 describe un método para optimizar las condiciones para la purificación cromatográfica de una proteína diana generando un modelo creado a partir de datos obtenidos sometiendo dicha proteína diana a una variedad de condiciones cromatográficas.

40 El documento US 2008/066530 A1 describe un método para optimizar las condiciones para la purificación cromatográfica de una proteína diana generando un modelo creado a partir de datos obtenidos sometiendo dicha proteína diana a una variedad de condiciones cromatográficas.

45 El documento US 2004/034477 A1 describe un método para generar condiciones óptimas para la purificación cromatográfica de un compuesto diana comparando la estructura del compuesto diana con otros compuestos en una base de datos y usando condiciones similares a las condiciones ya optimizadas conocidas para compuestos con una estructura similar a los compuestos diana. Es evidente para los expertos en la técnica que este método es mucho más adecuado para compuestos orgánicos más pequeños que biomoléculas o proteínas más grandes y significativamente más complejas.

### Compendio

45 La invención es un método y sistema como se define en las reivindicaciones adjuntas.

### Breve descripción de las figuras

La FIG. 1 es un diagrama de algunas implementaciones de una arquitectura para determinar las condiciones para optimizar la purificación de proteínas.

50 La FIG. 2 ilustra algunas implementaciones de técnicas para determinar las condiciones para optimizar la purificación de proteínas usando datos obtenidos a partir de una placa multipocillo.

La FIG. 3 ilustra algunas implementaciones de técnicas para determinar modelos para predecir la pureza y el

rendimiento de proteínas en diversas condiciones de purificación.

La FIG. 4 ilustra algunas implementaciones de técnicas para utilizar una pluralidad de modelos para determinar las condiciones óptimas para la purificación de proteínas.

5 La FIG. 5 ilustra algunas implementaciones de un sistema para determinar las condiciones de purificación de proteínas usando uno o más procesos cromatográficos.

La FIG. 6 es un diagrama de flujo de un proceso de ejemplo para determinar las condiciones de purificación de proteínas usando uno o más procesos cromatográficos.

La FIG. 7 ilustra una serie de interfases de usuario que indican el rendimiento y la pureza de una proteína con respecto a un número de niveles de pH y concentraciones de sal.

10 La FIG. 8 ilustra los resultados a partir de diferentes modelos que se pueden usar para determinar el rendimiento y la pureza de las proteínas usando cromatografía multi-modo a varios niveles de pH y de concentraciones de sal.

La FIG. 9 muestra una configuración de placa de filtro de 96 pocillos.

La FIG. 10 muestra una comparación de la estabilidad de la solución para 2 mAbs para el intervalo de condiciones del proceso corriente abajo evaluadas.

15 La FIG. 11 ilustra la estabilidad térmica mediante DSF para el intervalo de condiciones del proceso corriente abajo evaluadas.

La FIG. 12, paneles A a D, y la FIG. 13 ilustran un ejemplo de datos generados durante un experimento de unión de un solo lote.

La FIG. 14 ilustra la pureza frente al rendimiento para modos emparejados en todo el proceso corriente abajo.

20 **Descripción detallada**

Los conceptos descritos en la presente memoria están dirigidos a determinar condiciones optimizadas para la purificación de proteínas. En particular, las técnicas y sistemas descritos en la presente memoria están relacionados con la determinación de condiciones para procesos cromatográficos que se pueden utilizar para purificar proteínas después de que las proteínas se hayan expresado. En algunas implementaciones, los sistemas pueden utilizar técnicas de aprendizaje automático para determinar los tipos de procesos cromatográficos y las condiciones para los procesos cromatográficos que dan lugar a rendimientos y pureza óptimos para diversas proteínas.

25 La cromatografía en columna puede utilizar varios principios para separar las proteínas de las impurezas en una solución. La cromatografía en columna puede incluir colocar una cantidad de material en una columna dispuesta verticalmente, donde el material interacciona con las moléculas que fluyen a través de la columna. A medida que diferentes moléculas interaccionan con el material colocado en la columna, las moléculas se mueven a través de la columna a diferentes velocidades. Por lo tanto, algunas moléculas se pueden mover más lentamente a través de la columna que otras, lo que da lugar a una separación de las moléculas que se mueven a través de la columna. Luego se pueden capturar las diferentes moléculas a medida que salen de la columna. Las moléculas se pueden separar en función del tamaño, la forma, la carga, la hidrofobicidad, combinaciones de estas u otras características que pueden producir interacciones entre las moléculas y el material colocado en la columna. En diversas implementaciones, el material colocado en la columna que interacciona con las moléculas que se mueven a través de la columna se puede denominar material de fase estacionaria. Las moléculas que se mueven a través de la columna se pueden denominar fase móvil o eluyente que pasa sobre la fase estacionaria. Algunos procesos cromatográficos pueden incluir cromatografía de intercambio iónico, cromatografía de afinidad, cromatografía líquida de alta presión (HPLC), cromatografía de interacción hidrófoba (HIC), cromatografía de modo mixto (MMC), cromatografía de fase inversa (RPC), cromatografía de exclusión por tamaño y similares.

30 El costo y la eficacia de un proceso de purificación de proteínas que usa cromatografía en columna pueden depender del tipo de proceso cromatográfico utilizado para purificar las proteínas ya que algunas técnicas cromatográficas pueden ser más eficaces para purificar ciertas proteínas que otras técnicas cromatográficas. El tipo de proceso cromatográfico utilizado para purificar una proteína también puede corresponder al uso de ciertos materiales de fase estacionaria que pueden afectar al costo de la purificación de proteínas. En algunos casos, se pueden utilizar múltiples procesos de cromatografía para purificar proteínas; sin embargo, cada etapa cromatográfica utilizada en el proceso de purificación de proteínas puede provocar la pérdida de cierta cantidad de proteína.

35 Normalmente, la selección de procesos cromatográficos y la determinación de las condiciones para purificar proteínas se basa en el conocimiento disponible para los expertos en la técnica, como artículos de revistas, artículos de investigación y libros de texto, y se realiza un proceso de prueba y error basado en este conocimiento mediante el cual una proteína particular se purifica usando uno o más procesos cromatográficos en una variedad de condiciones. El uso del proceso convencional para seleccionar técnicas y condiciones cromatográficas tiene un alcance limitado y no es eficaz. Es decir, para cada proteína que se va a purificar, se utiliza una cantidad de material de fase estacionaria

y una cantidad de proteína con el fin de determinar un posible protocolo para la purificación de la proteína. Dado que se utiliza una cantidad de material de fase estacionaria y una cantidad de proteína para identificar procesos cromatográficos para purificar una proteína, el número de combinaciones de condiciones, como el pH y la concentración de sal de la solución de fase móvil, también es limitado. En particular, el costo de la proteína y del material de la fase estacionaria, así como el tiempo utilizado, durante más de unas pocas ejecuciones de prueba para determinar las condiciones que proporcionen el mejor rendimiento y pureza para la proteína, pueden llegar a ser prohibitivos.

Las técnicas y sistemas descritos en la presente memoria superan los problemas con las técnicas convencionales para determinar las condiciones de purificación de proteínas recopilando datos con respecto a diferentes técnicas de cromatografía en diversas condiciones para un número limitado de proteínas y luego utilizando el aprendizaje automático para optimizar el proceso de purificación para muchas otras proteínas adicionales. En algunas implementaciones, se pueden desarrollar modelos para predecir las condiciones cromatográficas óptimas para una proteína basándose en datos particulares obtenidos sobre la proteína. Por ejemplo, los modelos para determinar el rendimiento y la pureza para un número de técnicas cromatográficas en diversas condiciones se pueden basar en secuencias de proteínas. Además, los modelos para determinar el rendimiento y la pureza para un número de técnicas cromatográficas en diversas condiciones se pueden basar en estructuras secundarias asociadas con las proteínas. En otras situaciones, los modelos para determinar el rendimiento y la pureza para un número de técnicas cromatográficas con respecto a ciertas condiciones se pueden basar en valores de propiedades biofísicas de las proteínas que se miden usando técnicas analíticas.

Los modelos desarrollados para determinar el rendimiento y la pureza de las proteínas se pueden evaluar juntos para determinar una técnica cromatográfica particular o una combinación particular de técnicas que proporcionen el rendimiento y la pureza óptimos para la proteína. El costo de implementar las técnicas cromatográficas también se puede evaluar al determinar una técnica cromatográfica particular o una combinación de técnicas cromatográficas que proporcionen el rendimiento y la pureza óptimos para la proteína. El costo de una técnica cromatográfica se puede basar en el costo del material de fase estacionaria, el costo de la fase móvil, el tamaño de la columna utilizada, la dificultad de trabajar con el material de fase estacionaria, el rendimiento de una etapa de cromatografía individual, el rendimiento combinado de varias etapas de cromatografía, o combinaciones de estas.

Además, las técnicas y sistemas descritos en la presente memoria pueden minimizar la cantidad de datos necesarios para generar los modelos utilizados para predecir el rendimiento y la pureza de las proteínas. En diversas implementaciones, las condiciones de cromatografía utilizadas en las columnas se pueden determinar a partir de un número mínimo de combinaciones de condiciones de cromatografía que se pueden simular usando una placa multipocillo. Se puede identificar el número mínimo de combinaciones de condiciones de cromatografía minimizando las diferencias entre los datos de referencia y un subconjunto de pocillos de la placa multipocillo en diversas condiciones de cromatografía para diferentes materiales de fase estacionaria. En particular, se pueden utilizar subconjuntos de condiciones de cromatografía para un material de fase estacionaria para algunas proteínas a través de una gama más amplia de proteínas para proporcionar datos de rendimiento y pureza para generar modelos que predigan el rendimiento y la pureza de proteínas para materiales de fase estacionaria particulares.

Al utilizar las técnicas y sistemas descritos en la presente memoria, se pueden determinar las condiciones óptimas para la purificación de proteínas para un número de proteínas sin tener que realizar ejecuciones reales de procesos cromatográficos para cada una de las proteínas que se evalúan en varias condiciones diferentes. Por tanto, se reduce el costo de determinar las condiciones óptimas para la purificación de una proteína y se aumenta la eficacia. Además, al minimizar la cantidad de datos obtenidos para generar los modelos para determinar las condiciones óptimas que maximicen el rendimiento y la pureza para la purificación de proteínas, se minimiza la cantidad de recursos informáticos, como recursos de procesamiento y recursos de memoria.

La FIG. 1 es un diagrama de algunas implementaciones de una arquitectura 100 para determinar las condiciones para optimizar la purificación de proteínas. La arquitectura 100 puede evaluar un número de proteínas, como la proteína representativa 102, para producir datos de entrenamiento 104. Los datos de entrenamiento 104 se pueden proporcionar a un sistema de generación de modelos 106 para producir un número de modelos para predecir medidas de desempeño de un número de materiales de fase estacionaria para diversas técnicas cromatográficas. Las medidas del desempeño pueden incluir al menos una de rendimiento o de pureza para un material de fase estacionaria individual. En algunas implementaciones, la proteína 102 puede incluir un anticuerpo. En un ejemplo ilustrativo, la proteína 102 puede incluir un anticuerpo IgG.

Los datos de entrenamiento 104 pueden incluir datos de secuencia 108 que indican al menos una porción de una secuencia de aminoácidos de la proteína 102. Los datos de secuencia 108 pueden indicar el aminoácido en posiciones individuales de la secuencia de la proteína 102. En diversas implementaciones, los datos de secuencia 108 se pueden determinar mediante espectrometría de masas. Se pueden encontrar técnicas particulares para determinar una secuencia de proteínas mediante espectrometría de masas en Hunt, DF y cols. "Protein Sequencing by Tandem Mass Spectrometry", *Proceedings of the National Academy of Sciences of the United States of America* 83.17 (1986): 6233-6237. Impreso. Además, los datos de secuencia 108 se pueden determinar usando la degradación de Edman como se describe en Berg JM, Tymoczko JL, Stryer L. *Biochemistry*. 5ª edición. Nueva York: WH Freeman; 2002. Sección 4.2, Amino Acid Sequences Can Be Determined by Automated Edman Degradation. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK22571/>. Además, los datos de secuencia 108 se pueden determinar a partir

de una secuencia de nucleótidos asociada con la proteína 102, como una secuencia de ácido desoxirribonucleico (ADN) o una secuencia de ácido ribonucleico (ARN) asociada con la proteína 102, como se describe en Smith, A. (2008) Nucleic acids to amino acids: DNA specifies protein. *Nature Education* 1(1): 126.

5 Los datos de entrenamiento 104 también pueden incluir datos de estructura 110 que indican características estructurales de la proteína 102. Los datos de estructura 110 pueden indicar características secundarias de la proteína 102, características terciarias de la proteína 102 o ambas. Los datos de estructura 110 pueden indicar características de la estructura tridimensional de la proteína 102, como giros y/o bucles. En ciertas implementaciones, los datos de estructura 110 pueden indicar hélices  $\alpha$ , giros  $\beta$ , láminas  $\beta$ , bucles  $\Omega$  o combinaciones de estos. Los datos de estructura 110 también pueden indicar regiones hidrófobas de la proteína 102, regiones polares de la proteína 102, regiones de la proteína 102 que tienen ciertas cargas o combinaciones de estas. Los datos de estructura 110 pueden indicar un número de posiciones asociadas con un tipo de estructura secundaria o un tipo de estructura terciaria de la proteína 102. La estructura secundaria de la proteína 102 se puede determinar mediante una serie de métodos que incluyen los descritos en Determination of the secondary structure of proteins by laser Raman spectroscopy; J. L. Lippert, D. Tyminski y P. J. Desmeules; *Journal of the American Chemical Society* 1976 98(22), 7075-7080 DOI: 10.1021/ja00438a057.

15 Los datos de entrenamiento 104 también pueden incluir datos obtenidos a partir del análisis de la proteína 102 mediante uno o más sistemas de purificación 112 para producir datos de purificación 114. Uno o más sistemas de purificación 112 pueden incluir una o más técnicas cromatográficas. En implementaciones particulares, la proteína 102 se puede expresar y luego purificar utilizando al menos una técnica cromatográfica en diversas condiciones, como 20 diversos niveles de pH y de concentraciones de sal. Los datos de purificación 114 también pueden incluir datos de rendimiento y pureza obtenidos al utilizar uno o más sistemas de purificación 112 con respecto a la proteína 102. El uno o más sistemas de purificación 112 pueden incluir procesos cromatográficos, como cromatografía de intercambio iónico (IEX), cromatografía de afinidad, cromatografía de interacción hidrófoba (HIC), cromatografía multi-modo (MMC), cromatografía de fase inversa (RPC), cromatografía de exclusión por tamaño (SEC), o combinaciones de estas. También se contempla el uso de otras técnicas cromatográficas en otras implementaciones. En algunas 25 implementaciones, los datos de purificación 114 se pueden determinar analizando los valores de pureza y rendimiento obtenidos a partir de placas multipocillo con respecto a diversos materiales de fase estacionaria. Los materiales de fase estacionaria analizados con respecto a los sistemas de purificación 112 pueden incluir materiales poliméricos, como resinas. En implementaciones particulares, los materiales de fase estacionaria analizados con respecto a los sistemas de purificación 112 pueden incluir materiales que contienen sílice. En algunos ejemplos ilustrativos, los materiales de fase estacionaria utilizados con respecto a los sistemas de purificación 112 pueden incluir sílice, agarosa entrecruzada, metacrilato, polivinil éter hidrófilo, celulosa, hidrogel, polimetacrilato, hidroxipatita o combinaciones de estos.

30 Los datos de entrenamiento 104 también pueden incluir datos obtenidos sometiendo la proteína 102 a pruebas analíticas 116 para producir propiedades medidas 118. Las pruebas analíticas pueden incluir una serie de ensayos que producen datos sobre la proteína 102. Las propiedades medidas 118 pueden incluir un peso molecular de la proteína 102 que se puede determinar usando cromatografía de exclusión por tamaño. Las propiedades medidas 118 también pueden incluir turbidez medida usando un espectrofotómetro UV-Vis. Además, las propiedades medidas 118 pueden incluir una medida de estabilidad de la proteína 102 que se determina mediante fluorimetría diferencial de barrido (DSF). La estabilidad de la proteína 102 también se puede medir usando un ensayo de desplegamiento químico o un ensayo de viscosidad. Además, las propiedades medidas 118 pueden incluir medidas de interacciones entre regiones de la proteína 102 según lo determinado mediante espectroscopia de auto-interacción de nanopartículas (SINS).

35 Los datos de entrenamiento 104 se pueden proporcionar a un sistema de generación de modelos 106 que puede determinar modelos para predecir medidas de desempeño para diversas técnicas de purificación de proteínas. Los modelos individuales producidos por el sistema de generación de modelos 106 se pueden asociar con un material de fase estacionaria respectivo. Por ejemplo, un primer modelo 120 puede predecir el rendimiento y/o la pureza para un primer material de fase estacionaria. Además, un segundo modelo 122 puede predecir el rendimiento y/o la pureza de un segundo material de fase estacionaria. El sistema de generación de modelos 106 puede generar hasta un enésimo modelo 124 que puede predecir el rendimiento y/o la pureza de un enésimo material de fase estacionaria. En algunas implementaciones, el primer material de la fase estacionaria se puede asociar con una primera técnica cromatográfica y el segundo material de la fase estacionaria se puede asociar con una segunda técnica cromatográfica que es diferente de la primera técnica cromatográfica. En otras implementaciones, el primer material de la fase estacionaria puede ser diferente del segundo material de la fase estacionaria, pero asociado con una misma técnica cromatográfica. En ejemplos ilustrativos, el primer modelo 120 se puede producir con respecto a un material de fase estacionaria asociado con una técnica cromatográfica de intercambio aniónico, el segundo modelo 122 se puede producir con respecto a un material de fase estacionaria asociado con una técnica cromatográfica de interacción hidrófoba, y el enésimo modelo 124 se puede producir con respecto a un material de fase estacionaria asociado con una técnica cromatográfica de modo mixto.

40 En 126, el sistema de generación de modelos 106 puede analizar los datos de entrenamiento 104 para identificar relaciones entre características de un grupo de proteínas que incluye la proteína representativa 102. En algunas implementaciones, el sistema de generación de modelos 106 puede determinar relaciones entre secuencias de las proteínas y una de más de las propiedades medidas. En diversas implementaciones, el sistema de generación de modelos 106 puede determinar relaciones entre el rendimiento y/o la pureza de uno o más sistemas de purificación

112 y los datos de secuencia 108, los datos de estructura 110 y/o las propiedades medidas 118 para un grupo de proteínas que incluye la proteína 102. Por ejemplo, el sistema de generación de modelos 106 puede determinar que las proteínas que tienen una secuencia particular en un intervalo de posiciones pueden tener valores de turbidez dentro de un intervalo específico. En otro ejemplo, el sistema de generación de modelos 106 puede determinar que las proteínas que tienen un intervalo de pesos moleculares se correlacionan con rendimientos relativamente más altos cuando se purifican usando MMC en relación con los rendimientos logrados cuando se purifican usando técnicas cromatográficas de intercambio aniónico. El sistema de generación de modelos 106 también puede determinar relaciones entre las condiciones de purificación, como niveles de pH y/o concentraciones de sal, y las secuencias de las proteínas utilizadas para producir los datos de entrenamiento 104, estructuras de las proteínas usadas para producir los datos de entrenamiento 104, propiedades medidas de las proteínas usadas para producir los datos de entrenamiento 104, o combinaciones de estas. Además, el sistema de generación de modelos 106 puede determinar las relaciones entre las condiciones de purificación y los tipos de materiales de fase estacionaria utilizados en uno o más sistemas de purificación 112.

En 128, el sistema de generación de modelos 106 puede generar un número de modelos para predecir medidas de desempeño, como rendimiento y/o pureza, de uno o más sistemas de purificación 112. En particular, los modelos producidos por el sistema de generación de modelos 106, como los modelos 120, 122, 124, pueden predecir medidas de desempeño para proteínas que son diferentes de las proteínas usadas para producir los datos de entrenamiento 104. Es decir, los modelos 120, 122, 124 pueden predecir medidas de desempeño para proteínas que tienen secuencias diferentes a las proteínas usadas para producir los datos de entrenamiento 104. Los modelos generados por el sistema de generación de modelos 106 pueden incluir variables asociadas con uno o más de los datos de secuencia 108, los datos de estructura 110, condiciones de purificación, o una o más de las propiedades medidas 118. Los modelos también pueden incluir coeficientes para las variables que indican pesos relativos de las variables entre sí. Las variables y coeficientes incluidos en los modelos se pueden basar en las relaciones entre las diversas características de los datos de entrenamiento que se determinan en la operación 126. En un ejemplo ilustrativo, los modelos 120, 122, 124 pueden incluir modelos polinomiales.

El sistema de generación de modelos 106 puede realizar un proceso iterativo para determinar los modelos 120, 122, 124. Por ejemplo, el sistema de generación de modelos 106 puede determinar una función polinómica para un material de fase estacionaria particular y probar la función polinómica frente a los datos de entrenamiento 104. Si la función polinómica no produce medidas de desempeños consistentes con los datos de entrenamiento 104, el sistema de generación de modelos 106 puede modificar las variables y/o coeficientes de la función polinómica inicial para producir una función polinómica adicional. El sistema de generación de modelos 106 puede entonces evaluar la función polinómica adicional con respecto a los datos de entrenamiento 104. El sistema de generación de modelos 106 puede evaluar un número de iteraciones de un modelo para cada material de fase estacionaria hasta que los resultados producidos por los modelos correspondan con los datos de entrenamiento 104 dentro de un valor umbral. Es decir, el sistema de generación de modelos 106 puede evaluar modelos para materiales de fase estacionaria para minimizar el error entre los resultados previstos y los resultados reales. En un ejemplo ilustrativo, el sistema de generación de modelos 106 puede probar de forma iterativa modelos para materiales de fase estacionaria individuales hasta que se alcance un mínimo local. En ejemplos ilustrativos adicionales, el sistema de generación de modelos 106 puede probar de forma iterativa modelos para materiales de fase estacionaria individuales hasta que se alcance un mínimo global. Además, el sistema de generación de modelos 106 puede probar de forma iterativa modelos para materiales de fase estacionaria individuales hasta que se haya completado un número específico de ciclos de entrenamiento. En situaciones en las que es menor la precisión de un material de fase estacionaria individual que un umbral de precisión, se pueden realizar ciclos de entrenamiento adicionales y, en algunas situaciones, los parámetros del modelo se pueden ajustar durante los ciclos de entrenamiento adicionales. En implementaciones particulares, los hiperparámetros del modelo se pueden ajustar para ciclos de entrenamiento adicionales en escenarios donde la precisión del modelo para un material de fase estacionaria individual es menor que un umbral de precisión.

Los modelos generados por el sistema de generación de modelos 106 se pueden proporcionar a un sistema de optimización de las condiciones de purificación 130. En el ejemplo ilustrativo de la FIG. 1, el sistema de optimización de las condiciones de purificación 130 puede utilizar los modelos 120, 122, 124, o combinaciones de estos, para optimizar las condiciones de purificación para proteínas que son diferentes de las proteínas utilizadas para producir los datos de entrenamiento 104. En un ejemplo ilustrativo, el sistema de optimización de condiciones de purificación 130 puede determinar las condiciones de optimización para la purificación de una proteína adicional 132. El sistema de optimización de condiciones de purificación 130 puede obtener datos de proteínas adicionales 134 que corresponden a la proteína adicional 132. Los datos de proteínas adicionales 134 pueden incluir una secuencia de la proteína adicional 132, una o más estructuras de la proteína adicional 132, propiedades medidas de la proteína adicional 132, o combinaciones de estas. En algunas implementaciones, la proteína adicional 132 puede tener algunas similitudes con respecto a la proteína 102. Por ejemplo, tanto la proteína 102 como la proteína adicional 132 pueden ser anticuerpos. En otro ejemplo, tanto la proteína 102 como la proteína adicional 132 pueden ser anticuerpos IgG. En otros ejemplos, un peso molecular de la proteína adicional 132 puede estar dentro de al menos aproximadamente el 2 % del peso molecular de la proteína 102, al menos aproximadamente el 5 % del peso molecular de la proteína 102, al menos aproximadamente el 10 % del peso molecular de la proteína 102, al menos aproximadamente el 15 % del peso molecular de la proteína 102, al menos aproximadamente el 20 % del peso molecular de la proteína 102, o al menos aproximadamente el 25 % del peso molecular de la proteína 102.

En 136, el sistema de condiciones de purificación 130 puede evaluar los datos de proteínas adicionales 134 con respecto a uno o más de los modelos generados por el sistema de generación de modelos 106. En implementaciones particulares, el sistema de optimización de condiciones de purificación 130 puede determinar valores de rendimiento y pureza para la proteína adicional 132 con respecto a una serie de materiales de fase estacionaria basándose en los datos de proteínas adicionales 134. Por ejemplo, el sistema de optimización de las condiciones de purificación 130 puede evaluar los datos de proteínas adicionales 134 con respecto al primer modelo 120, al segundo modelo 122, y al enésimo modelo 124 para determinar los valores de rendimiento y pureza para la proteína adicional 132 para cada material de fase estacionaria respectivo asociado con los modelos 120, 122, 124.

En 138, el sistema de optimización de las condiciones de purificación 130 puede determinar uno o más modelos para optimizar la purificación de la proteína adicional 132. En implementaciones particulares, el sistema de optimización de las condiciones de purificación 130 puede producir condiciones de purificación 140 para la proteína adicional 132. Las condiciones de purificación 140 pueden indicar una o más técnicas cromatográficas para purificar la proteína adicional 132. Las condiciones de purificación 140 también pueden indicar uno o más valores de pH y una o más concentraciones de sal a las que realizar las técnicas cromatográficas identificadas por el sistema de optimización de las condiciones de purificación 130.

El sistema de optimización de las condiciones de purificación 130 puede comparar los valores de rendimiento y pureza producidos para la proteína adicional 132 mediante uno o más de los modelos generados por el sistema de generación de modelos 106 para determinar un modelo que maximice los valores de rendimiento y pureza. En algunas implementaciones, el sistema de optimización de las condiciones de purificación 130 puede determinar una combinación de modelos que maximice los valores para la pureza y el rendimiento de la proteína adicional 132. En un ejemplo ilustrativo, el sistema de optimización de las condiciones de purificación 130 puede determinar que un material de fase estacionaria asociado con el primer modelo 120 pueda maximizar los valores de rendimiento y pureza para la proteína adicional. En otro ejemplo ilustrativo, el sistema de optimización de las condiciones de purificación 130 puede determinar que purificar la proteína adicional 132 utilizando un primer material de fase estacionaria asociado con el primer modelo 120 seguido de purificar la proteína adicional 132 con un segundo material de fase estacionaria asociado con el segundo modelo 122 puede maximizar los valores de rendimiento y pureza para la proteína adicional 132. Además, el sistema de optimización de las condiciones de purificación 130 puede determinar las condiciones de purificación 140 basándose al menos en parte en el costo de varios materiales de fase estacionaria. A modo de ilustración, el sistema de optimización de las condiciones de purificación 130 puede determinar las condiciones de purificación 140 basándose al menos parcialmente en los valores de rendimiento, los valores de pureza y el costo de los materiales de la fase estacionaria asociados con los modelos generados por el sistema de generación de modelos 106, como los modelos 120, 122, 124.

La FIG. 2 ilustra algunas implementaciones de técnicas para determinar condiciones para optimizar la purificación de proteínas usando datos obtenidos a partir de una placa multipocillo 202. En el ejemplo ilustrativo de la FIG. 2, los pocillos de la placa multipocillo 202, como el pocillo representativo 204, se pueden dividir en cuadrantes. Por ejemplo, la placa multipocillo 202 puede incluir un primer cuadrante 206, un segundo cuadrante 208, un tercer cuadrante 210 y un cuarto cuadrante 212. En el ejemplo ilustrativo de la FIG. 2, cada cuadrante se puede asociar con un material de fase estacionaria diferente. En otras implementaciones, la placa multipocillo 202 se puede dividir en un número diferente de porciones, como dos porciones, tres porciones o seis porciones, para acomodar un número diferente de materiales de fase estacionaria. Además, aunque cada cuadrante en la placa multipocillo 202 incluye una matriz de 4 x 6 correspondiente a cuatro concentraciones de sal diferentes y seis niveles de pH diferentes, las matrices asociadas con cada porción de la placa multipocillo 202 pueden ser diferentes. A modo de ilustración, las matrices de 4 x 6 de la placa multipocillo 202 pueden corresponder a cuatro niveles de pH diferentes y a seis concentraciones de sal diferentes. En un ejemplo adicional, una placa multipocillo se puede dividir en tres porciones, cada una de las cuales tiene una matriz de 8 x 4 correspondiente a ocho niveles de pH y cuatro concentraciones de sal u ocho concentraciones de sal y cuatro niveles de pH. Además, aunque el ejemplo ilustrativo de la FIG. 2 incluye una placa multipocillo 202 con 96 pocillos, otras implementaciones pueden tener placas multipocillo con diferente número de pocillos.

La placa multipocillo 202 se puede utilizar para proporcionar datos de pureza y rendimiento con respecto a un conjunto de proteínas 214 para varios materiales de fase estacionaria. A modo de ilustración, los pocillos individuales de la placa multipocillo 202 pueden incluir una solución que incluye una cantidad de una proteína seleccionada del conjunto de proteínas 214 y una cantidad de un material de fase estacionaria. La solución incluida en los pocillos individuales de la placa multipocillo también puede tener un valor de pH y una concentración de sal. El valor del pH se puede basar en una cantidad de ácido incluida en la solución o en una cantidad de base incluida en la solución. La concentración de sal se puede basar en una cantidad de sal incluida en la solución, como una cantidad de NaCl incluida en la solución. También se pueden utilizar otras sales en la solución, como sales de ácidos débiles (p. ej.,  $\text{CH}_3\text{COONa}$ ) o sales de bases débiles (p. ej.,  $\text{NH}_4\text{Cl}$ ). En implementaciones particulares, los valores de pH de las soluciones incluidas en los pocillos de la placa multipocillo pueden ser de aproximadamente 2,5 a aproximadamente 8,5, de aproximadamente 3 a aproximadamente 8, de aproximadamente 3 a aproximadamente 5, o de aproximadamente 5 a aproximadamente 8. Además, las concentraciones de sal de las soluciones incluidas en los pocillos individuales de la placa multipocillo 202 pueden ser de aproximadamente 0 milimolar (mM) a aproximadamente 750 mM, de aproximadamente 20 mM a aproximadamente 650 mM, de aproximadamente 25 mM a aproximadamente 400 mM, de aproximadamente 30 mM a aproximadamente 200 mM, o de aproximadamente 5 mM a aproximadamente 100 mM.

Las determinaciones de rendimiento y/o pureza con respecto a los materiales de fase estacionaria, los valores de pH y las concentraciones de sal de los pocillos individuales de la placa multipocillo se pueden basar al menos en parte en el análisis de las soluciones de los pocillos individuales después de un período de tiempo. En implementaciones particulares, la estabilidad coloidal de las soluciones se puede determinar basándose en las propiedades de dispersión de la luz de la solución, como la turbidez, y la estabilidad química de la solución se puede determinar usando un ensayo de desplegamiento químico. Véase Guillermo Senisterra, Irene Chau y Masoud Vedadi. *ASSAY and Drug Development Technologies*. Volumen 10, Número 2, abril de 2012. También se puede determinar la estabilidad térmica de las soluciones, como mediante fluorimetría diferencial de barrido. La pureza se puede determinar mediante cromatografía de exclusión molecular. La unión de las proteínas del conjunto de proteínas 214 a los materiales de fase estacionaria incluidos en los pocillos individuales de la placa multipocillo 202 también se puede utilizar para determinar los valores de rendimiento y/o pureza para las proteínas individuales del conjunto de proteínas 212 bajo las condiciones de los pocillos individuales. En algunos casos, las soluciones incluidas en los pocillos individuales pueden pasar por un proceso cromatográfico particular. Además, para determinar una cantidad de proteína unida al material de fase estacionaria incluido en los pocillos individuales de la placa multipocillo 202, se puede comparar la cantidad de proteína en las soluciones de los pocillos individuales con la cantidad inicial de proteína en la solución cuando la solución se añadió primero al pocillo individual. Los valores de características particulares de las proteínas individuales incluidas en el conjunto de proteínas 214 también se pueden analizar después de que las soluciones hayan estado en los pocillos individuales durante un período de tiempo para determinar la estabilidad de las proteínas individuales, el rendimiento de las proteínas individuales y/o la pureza de las proteínas individuales. A modo de ilustración, en situaciones en las que el conjunto de proteínas 214 incluye anticuerpos, se puede analizar el componente de alto peso molecular (HMW) para determinar la estabilidad de proteínas individuales a diversos valores de pH y de concentraciones de sal.

En ciertas implementaciones, cada cuadrante 206, 208, 210, 212 de la placa multipocillo 202 puede representar diversas condiciones experimentales para una o más proteínas incluidas en el conjunto de proteínas 214. Por ejemplo, los pocillos individuales incluidos en un cuadrante particular 206, 208, 210, 212 pueden representar diferentes combinaciones de concentraciones de sal y de valores de pH para un material de fase estacionaria particular. En otras situaciones, se pueden evaluar múltiples materiales de fase estacionaria a diversos valores de pH y de concentraciones de sal en un cuadrante particular 206, 208, 210, 212.

En las implementaciones descritas en la presente memoria, se puede determinar un subconjunto de pocillos para uno o más de los cuadrantes 206, 208, 210, 212 que representan el rendimiento, la pureza y/o la estabilidad de la proteína para los pocillos agregados del cuadrante. A modo de ilustración, las técnicas descritas con respecto a la FIG. 2 pueden identificar cinco pocillos o seis pocillos del tercer cuadrante 210 que representan los datos que se pueden obtener a partir de las condiciones asociadas con los veinticuatro pocillos del cuadrante 210. Al reducir el número de pocillos analizados para determinar los datos que se pueden utilizar para predecir el rendimiento, la pureza y/o la estabilidad de proteínas individuales, se minimiza la cantidad de proteína y la cantidad de material de fase estacionaria utilizada. En el ejemplo ilustrativo de la FIG. 2, se puede analizar un primer conjunto de pocillos 216 hasta un enésimo conjunto de pocillos 218 para identificar un número mínimo de pocillos para determinar los datos de rendimiento, pureza y/o estabilidad para proteínas individuales del conjunto de proteínas 214. Los pocillos individuales incluidos en el primer conjunto de pocillos 216 hasta el enésimo conjunto de pocillos 218 se pueden mostrar con un color gris en el ejemplo ilustrativo de la FIG. 2.

En 220, se puede analizar el primer conjunto de pocillos 216 para producir el primer conjunto de datos 222. El análisis del primer conjunto de pocillos 216 puede incluir determinar el rendimiento y/o la pureza de la proteína incluida en los pocillos individuales del primer conjunto de pocillos 216 bajo los valores de pH y de concentración de sal asociados con los pocillos individuales del primer conjunto de pocillos 216 para producir el primer conjunto de datos 222. El análisis del primer conjunto de pocillos 216 también puede incluir determinar ciertas características de una proteína incluida en los pocillos individuales del primer conjunto de pocillos después de permanecer en los pocillos respectivos durante un período de tiempo, como un componente de alto peso molecular (HMW), para producir el primer conjunto de datos 222. Además, se pueden analizar las características de la solución incluida en los pocillos individuales del primer conjunto de pocillos 216, como la turbidez de la solución, para producir el primer conjunto de datos 222.

Se pueden analizar un número de conjuntos adicionales de pocillos que incluyen diferentes combinaciones de condiciones hasta que se analice un enésimo conjunto de pocillos en 224 para producir un enésimo conjunto de datos 226. En algunos casos, las combinaciones de pocillos que se analizan para cada ronda de análisis pueden incluir el mismo número de pocillos, pero tener diferentes combinaciones de condiciones. En implementaciones particulares, se pueden seleccionar los pocillos del tercer cuadrante 210 para su evaluación de acuerdo con un algoritmo aleatorio o pseudoaleatorio. En un ejemplo ilustrativo, se puede utilizar un algoritmo de Monte Carlo para determinar las combinaciones de pocillos individuales a evaluar para cada ronda de análisis.

Se pueden analizar los conjuntos de datos individuales producidos para cada combinación de pocillos, en 228, con respecto a un conjunto de datos de referencia 230. El conjunto de datos de referencia 230 se puede producir a partir de un mayor número de pocillos del tercer cuadrante que el número de pocillos incluidos en el primer conjunto de pocillos 216 hasta el enésimo conjunto de pocillos 218. Por ejemplo, el conjunto de datos de referencia 230 se puede producir a partir de al menos aproximadamente el 75 % de los pocillos del tercer cuadrante 210, al menos aproximadamente el 80 % de los pocillos del tercer cuadrante 210, al menos aproximadamente el 85 % de los pocillos

del tercer cuadrante 210, al menos aproximadamente el 90 % de los pocillos del tercer cuadrante 210, o al menos aproximadamente el 95 % de los pocillos del tercer cuadrante 210. En diversas implementaciones, el conjunto de datos de referencia 230 se puede derivar de un conjunto de pocillos incluidos en el tercer cuadrante que incluye al menos una porción de los pocillos incluidos en cada uno del primer conjunto de pocillos 216 hasta el enésimo conjunto de pocillos 218. Los datos de referencia 230 pueden incluir datos que se corresponden con los datos incluidos en el primer conjunto de datos 222 hasta el enésimo conjunto de datos 226. A modo de ilustración, el conjunto de datos de referencia 230 puede incluir valores de rendimiento y/o pureza para un número de los pocillos incluidos en el tercer cuadrante 210, datos relacionados con la solución incluida en los pocillos del tercer cuadrante 210 y datos que indican características de la proteína incluida en los pocillos del tercer cuadrante 210.

En implementaciones particulares, se puede comparar al menos una parte del primer conjunto de datos 222 con al menos una parte del conjunto de datos de referencia 230 y se puede determinar un valor de error entre al menos una parte del primer conjunto de datos 222 y al menos una parte del conjunto de datos de referencia 230. Luego se puede seleccionar otro grupo de pocillos y se puede comparar al menos una parte de los datos derivados del segundo grupo de pocillos con al menos una parte del conjunto de datos de referencia 230 y se puede determinar un valor de error entre los datos derivados del segundo grupo de pocillos y los datos de referencia 230. Se pueden analizar un número de grupos adicionales de pocillos y se pueden comparar los datos derivados de los grupos adicionales de pocillos con los datos de referencia 230. Cada grupo sucesivo de pocillos se analiza de tal manera que se reduzca el error entre un grupo de pocillos recién seleccionado y los datos de referencia hasta alcanzar un error mínimo. La combinación de pocillos que está asociada con el error mínimo entre los datos generados a partir de la combinación de pocillos y los datos de referencia está representada por la disposición optimizada de pocillos 232.

La disposición optimizada de pocillos 232 se puede utilizar para generar datos para proteínas adicionales que son diferentes del conjunto de proteínas 214 que se usaron para producir la disposición optimizada de pocillos 232. En el ejemplo ilustrativo de la FIG. 2, se puede colocar una proteína adicional 234 en los pocillos asociados con las condiciones correspondientes a los pocillos de la disposición optimizada de pocillos 232. En particular, se puede colocar una cantidad de proteína adicional 234 en los pocillos de una placa multipocillo a valores de pH, concentraciones de sal, y con una cantidad de material de fase estacionaria que corresponde a las condiciones de la disposición optimizada de pocillos 232. En 236, la proteína adicional se puede analizar con respecto a las condiciones de la disposición optimizada de pocillos 232 para producir un conjunto de datos de proteína adicional 238. El conjunto de datos de proteína adicional 238 puede incluir valores de rendimiento y/o pureza para los pocillos de la disposición optimizada de pocillos 232 con respecto a la proteína adicional 234. El conjunto de datos de proteína adicional 238 también puede incluir datos que indican características de la solución incluida en los pocillos individuales de la disposición optimizada de pocillos para la proteína adicional 234.

La FIG. 3 ilustra algunas implementaciones de técnicas para determinar modelos para predecir la pureza y el rendimiento de proteínas en diversas condiciones de purificación. En particular, una primera proteína 302 se puede asociar con los primeros datos de caracterización 304. Los primeros datos de caracterización 304 pueden incluir una secuencia de la primera proteína 302, una o más estructuras secundarias de la primera proteína 302, una o más estructuras terciarias de la primera proteína 302, o combinaciones de estas. Los primeros datos de caracterización 304 también pueden incluir valores para propiedades biofísicas de la primera proteína 302. El sistema de generación de modelos 106 puede obtener los primeros datos de caracterización 304 y utilizar los primeros datos de caracterización 304 para generar modelos para predecir el rendimiento y la pureza de diversas proteínas. El sistema de generación de modelos 106 también puede obtener datos sobre otras proteínas, hasta una enésima proteína 306, para generar los modelos utilizados para predecir el rendimiento y la pureza. En el ejemplo ilustrativo de la FIG. 3, el sistema de generación de modelos 106 también puede obtener enésimos datos de caracterización 308 que incluyen datos de secuencia para la enésima proteína 306, datos para una o más estructuras secundarias de la enésima proteína, datos para una o más estructuras terciarias de la enésima proteína 306, datos de propiedades biofísicas de la enésima proteína 306, o combinaciones de estos.

Además de los datos de caracterización de proteínas, el sistema de generación de modelos 106 puede utilizar datos de rendimiento y pureza para la primera proteína 302 hasta la enésima proteína 306 para generar modelos para determinar los valores de rendimiento y pureza de proteínas adicionales. Por ejemplo, las condiciones de purificación 310 se pueden aplicar a una columna de cromatografía 312 con respecto a la primera proteína 302 hasta la enésima proteína 306. De esta manera, se pueden recopilar los primeros valores de pureza y rendimiento 314 para la primera proteína 302 hasta los enésimos valores de pureza y rendimiento para la enésima proteína 306 con respecto a múltiples materiales de fase estacionaria en diversas combinaciones de niveles de pH y de concentraciones de sal.

El sistema de generación de modelos 106 puede analizar los primeros datos de caracterización 304 hasta los enésimos datos de caracterización 306 y los primeros valores de rendimiento y pureza 314 hasta los enésimos valores de rendimiento y pureza 316 para determinar las relaciones entre diferentes variables incluidas en los datos de caracterización 304...306 y los valores de rendimiento y pureza 314...316. Por ejemplo, el sistema de generación de modelos 106 puede identificar relaciones entre valores de rendimiento y pureza para proteínas que tienen características de secuencia particulares bajo ciertas condiciones de purificación, como un material de fase estacionaria particular, un intervalo de valores de pH y un intervalo de concentraciones de sal. El sistema de generación de modelos 106 también puede identificar relaciones entre los valores de rendimiento y pureza para proteínas que tienen valores para una o más propiedades biofísicas particulares bajo ciertas condiciones de

purificación. Además, el sistema de generación de modelos 106 puede identificar relaciones entre los valores de rendimiento y pureza para proteínas que tienen estructuras secundarias particulares en diversas posiciones bajo ciertas condiciones de purificación.

5 El sistema de generación de modelos 106 puede generar un primer modelo 318 que tiene primeras variables y primeros pesos 320 hasta un enésimo modelo 322 que tiene enésimas variables y enésimos pesos 324. Los modelos individuales desde el primer modelo 318 hasta el enésimo modelo 322 pueden corresponder a materiales de fase estacionaria individuales que se pueden utilizar para purificar proteínas en una columna de cromatografía, como la columna 312. Por ejemplo, el primer modelo 318 puede predecir valores de rendimiento y pureza para un material de fase estacionaria utilizado en cromatografía de intercambio aniónico. En otro ejemplo, el enésimo modelo 322 puede predecir valores de rendimiento y pureza para un material de fase estacionaria utilizado en cromatografía multimodo. Las variables del primer modelo 318 hasta el enésimo modelo 322 pueden corresponder a características de los datos de caracterización de proteínas. Además, las variables del primer modelo 318 pueden corresponder a varios niveles de pH y de concentraciones de sal. A modo de ilustración, el primer modelo 318 puede incluir variables relacionadas con la secuencia de proteínas y la presencia de diversas estructuras secundarias con respecto a un material de fase estacionaria particular. El primer modelo 318 también puede incluir variables relacionadas con niveles de pH y de concentraciones de sal con respecto al material de fase estacionaria particular. Los pesos asociados al primer modelo 318 hasta el enésimo modelo 322 pueden indicar la cantidad de impacto de una variable en los valores de rendimiento y pureza.

En implementaciones ilustrativas, se pueden evaluar los datos de caracterización para proteínas adicionales con respecto a uno o más de los modelos generados por el sistema de generación de modelos 106. En un ejemplo particular, un modelo generado por el sistema generador de modelos 106 se puede asociar con un material de fase estacionaria y tener variables relacionadas con una cantidad de aminoácidos hidrófobos incluidos en una secuencia de proteína, un número de láminas beta, valores de fluorimetría diferencial de barrido, niveles de pH y concentraciones de sal. Continuando con este ejemplo, se pueden evaluar datos de caracterización para una proteína adicional que indique un número de aminoácidos hidrófobos incluidos en la secuencia de la proteína adicional, un número de láminas beta incluidas en la proteína adicional y valores de fluorimetría diferencial de barrido de la proteína adicional junto con intervalos de niveles de pH y de concentraciones de sal para predecir valores de rendimiento y pureza para los diferentes materiales de la fase estacionaria de acuerdo con uno o más de los modelos generados por el sistema de generación de modelos 106.

En diversas implementaciones, el sistema de generación de modelos 106 puede generar modelos usando técnicas de análisis de componentes principales (PCA). Por ejemplo, el sistema de generación de modelos 106 puede caracterizar los valores de rendimiento y pureza, como los primeros valores de rendimiento y pureza 314 hasta los enésimos valores de rendimiento y pureza 316 como simplemente una variable PCA o como dos variables PCA. El sistema de generación de modelos 106 también puede determinar los pesos respectivos para las variables. De esta manera, las primeras variables y los primeros pesos 320 incluidos en el primer modelo 318 y las enésimas variables y enésimos pesos 324 incluidos en el enésimo modelo 322 pueden incluir una o más variables PCA y sus respectivos pesos. Al implementar técnicas de PCA para determinar variables y pesos a incluir en los modelos, el sistema de generación de modelos 106 puede reducir la potencia de cálculo, como ciclos del procesador, utilizados para predecir el rendimiento y/o la pureza usando los modelos.

En implementaciones particulares, el sistema de generación de modelos 106 puede generar modelos que se pueden implementar con un grupo de moléculas que tienen un conjunto de características basadas en los datos de caracterización de las moléculas y de la técnica de cromatografía usada para purificar las moléculas. Por ejemplo, moléculas que tienen secuencias particulares con variantes en ciertas posiciones, moléculas que tienen un rango específico de propiedades biofísicas y/o moléculas que tienen rangos específicos de propiedades biofísicas. En ciertas implementaciones, el sistema de generación de modelos 106 puede generar modelos separados para subconjuntos de moléculas incluidas en un grupo de moléculas. En un ejemplo ilustrativo, el sistema de generación de modelos 106 puede generar un modelo único para un grupo de proteínas que tiene un conjunto específico de características que se purifican usando cromatografía de exclusión por tamaño. En otro ejemplo ilustrativo, el sistema de generación de modelos 106 puede generar múltiples modelos para el mismo grupo de proteínas que se purifican usando cromatografía de interacción hidrófoba. Específicamente, el sistema de generación de modelos 106 puede generar un primer modelo para predecir el rendimiento y/o la pureza de moléculas incluidas en el grupo de moléculas que tienen una cantidad de retención relativamente alta con respecto al material de fase estacionaria incluido en la cromatografía y un segundo modelo para predecir el rendimiento y/o la pureza de las moléculas incluidas en el grupo que tienen una cantidad de retención relativamente menor con respecto al material de la fase estacionaria incluido en la columna de cromatografía. El sistema de generación de modelos 106 puede utilizar una cantidad umbral de retención para determinar las moléculas incluidas en un primer grupo que tienen una cantidad de retención relativamente alta en relación con las moléculas incluidas en un segundo grupo que tiene una cantidad de retención relativamente baja. En algunos ejemplos ilustrativos particulares, la retención de moléculas con respecto a un material de fase estacionaria se puede medir de acuerdo con un Factor de Retención,  $k$ . La retención de moléculas con respecto a un material de fase estacionaria también se puede medir de acuerdo con factores de tiempo y/o factores de volumen, como un tiempo de retención para un volumen dado de material de fase móvil.

La FIG. 4 ilustra algunas implementaciones de técnicas para utilizar una pluralidad de modelos para determinar las condiciones óptimas para la purificación de proteínas. Los modelos se pueden generar mediante el sistema de

generación de modelos 106 de la FIG. 1 y de la FIG. 3 y utilizado por el sistema de optimización de las condiciones de purificación 130 para determinar las condiciones óptimas para la purificación de diversas proteínas. En el ejemplo ilustrativo de la FIG. 4, las condiciones óptimas para la purificación de la proteína 402 se pueden determinar mediante el sistema de optimización de las condiciones de purificación 130. A modo de ilustración, se pueden proporcionar al menos una parte de los datos de caracterización de proteínas 404 que están asociados con la proteína 402 al sistema de optimización de las condiciones de purificación 130. Las porciones de los datos de caracterización de proteínas 404 obtenidos por el sistema de optimización de las condiciones de purificación 130 pueden corresponder a los datos asociados con las variables incluidas en los diversos modelos 406. Por ejemplo, en situaciones donde uno o más de los modelos 406 incluyen una variable relacionada con un número de aminoácidos polares incluidos en una secuencia de proteína, los datos de caracterización de proteínas 404 pueden incluir el número de aminoácidos polares incluidos en la secuencia de la proteína 402. En algunos ejemplos ilustrativos, el desplegamiento químico puede ser un factor incluido en los datos de caracterización 404 que predicen los valores de pureza y/o rendimiento. En otros ejemplos ilustrativos, las regiones hidrófobas pueden predecir los valores de pureza y/o rendimiento. Aún en otros ejemplos ilustrativos, tanto el desplegamiento químico como la hidrofobicidad pueden predecir los valores de pureza y/o rendimiento.

Analizando al menos parte de los datos de caracterización de proteínas 404 con respecto a los modelos 406, como un primer modelo 408 hasta un enésimo modelo 410, el sistema de optimización de las condiciones de purificación 130 puede producir valores de rendimiento y pureza para la proteína 402 con respecto a diversos niveles de pH y de concentraciones de sal. En ejemplos particulares, se pueden evaluar los datos de caracterización de proteínas 404 con respecto a los modelos 406 para determinar los respectivos valores de rendimiento y de proteína para combinaciones individuales de niveles de pH y de concentraciones de sal. A modo de ilustración, el sistema de optimización de las condiciones de purificación 130 puede producir primeros valores de rendimiento, pureza, pH y de concentración de sal 412 con respecto al primer modelo 408 hasta los enésimos valores de rendimiento, pureza, pH y de concentración de sal 414 con respecto al enésimo modelo 410. En un ejemplo ilustrativo, el sistema de optimización de las condiciones de purificación 130 puede determinar un rendimiento del 90% y una pureza del 85% para la proteína 402 con respecto al primer modelo 408 a un pH de 5,5 y a una concentración de sal de 300 mM.

El sistema de análisis de resultados del modelo 416 puede obtener los resultados producidos por los modelos 406 y determinar las condiciones óptimas para la purificación de proteínas. En algunos casos, las condiciones óptimas para la purificación de proteínas pueden incluir un material de fase estacionaria único a un nivel de pH particular y a una concentración de sal particular. En otras situaciones, las condiciones óptimas para la purificación de proteínas pueden incluir una combinación de materiales de fase estacionaria a uno o más niveles de pH y a una o más concentraciones de sal. En escenarios particulares, las condiciones óptimas para la purificación de la proteína 402 pueden incluir purificar la proteína 402 usando una técnica cromatográfica que tiene un primer material de fase estacionaria a un primer nivel de pH y a una primera concentración de sal seguido del uso de una técnica cromatográfica adicional que tiene un segundo material de fase estacionaria a un segundo nivel de pH y a una segunda concentración de sal. El sistema de análisis de resultados del modelo 416 también puede utilizar datos de costos para los diversos materiales de fase estacionaria para determinar las condiciones óptimas para la purificación de proteínas. En diversas implementaciones, el sistema de análisis de resultados del modelo 416 puede utilizar datos relacionados con el tamaño de una columna de cromatografía y/o los datos de rendimiento para una columna de cromatografía para determinar las condiciones óptimas para la purificación de proteínas.

El sistema de análisis de resultados del modelo 416 puede producir una o más interfases de usuario que indican datos de rendimiento y de pureza para un número de combinaciones de niveles de pH y de concentraciones de sal para diferentes materiales de fase estacionaria. Por ejemplo, el sistema de análisis de resultados del modelo 416 puede generar una primera interfase de usuario 416 que tiene valores de rendimiento en el eje x y valores de pureza en el eje y. Cada par de rendimiento y pureza incluido en la primera interfase de usuario 416 se puede asociar con un material de fase estacionaria particular a un nivel de pH y a una concentración de sal específicos. Además, el sistema de análisis de resultados del modelo 416 puede generar una segunda interfase de usuario 420 que tiene valores de pH en el eje x y valores de concentración de sal en el eje y. Cada par de pH y concentración de sal incluido en la segunda interfase de usuario 420 se puede asociar con valores de rendimiento y pureza particulares.

La FIG. 5 ilustra algunas implementaciones de un sistema para determinar las condiciones de purificación de proteínas usando uno o más procesos cromatográficos. El sistema 500 incluye un sistema de condiciones de purificación de proteínas 502 que se puede implementar mediante uno o más dispositivos informáticos 504. En algunas implementaciones, se pueden incluir uno o más dispositivos informáticos 504 en una arquitectura informática en la nube que opera uno o más dispositivos informáticos 504 en nombre de una entidad que implementa el sistema de condiciones de purificación de proteínas 502, como un usuario del sistema de condiciones de purificación de proteínas 502. En estos escenarios, la arquitectura informática en la nube puede crear instancias de una o más instancias de máquina virtual en nombre de la entidad que implementa el sistema de condiciones de purificación de proteínas 502 usando uno o más dispositivos informáticos 504. La arquitectura informática en la nube se puede ubicar en remoto de la entidad que implementa el sistema de condiciones de purificación de proteínas 502. En implementaciones adicionales, uno o más dispositivos informáticos 504 pueden estar bajo el control directo de la entidad que implementa el sistema de condiciones de purificación de proteínas 502. Por ejemplo, la entidad que implementa el sistema de condiciones de purificación de proteínas 502 puede mantener uno o más dispositivos informáticos 504 para realizar operaciones relacionadas con la generación de uno o más modelos y analizar los datos de proteínas con respecto a los modelos para determinar las condiciones para purificar proteínas. En diversas implementaciones, uno o más

dispositivos informáticos 504 pueden incluir uno o más ordenadores servidores.

El sistema de condiciones de purificación de proteínas 502 puede incluir uno o más procesadores, como el procesador 506. El uno o más procesadores 506 pueden incluir al menos un procesador de hardware, como un microprocesador. En algunos casos, uno o más procesadores 506 pueden incluir una unidad central de procesamiento (CPU), una unidad de procesamiento de gráficos (GPU), o tanto una CPU como una GPU, u otras unidades de procesamiento. Además, uno o más procesadores 506 pueden incluir una memoria local que puede almacenar módulos de programa, datos de programa y/o uno o más sistemas operativos.

Además, el sistema de condiciones de purificación de proteínas 502 puede incluir uno o más medios de almacenamiento legibles por ordenador, como medios de almacenamiento legibles por ordenador 508. Los medios de almacenamiento legibles por ordenador 508 pueden incluir memoria volátil y no volátil y/o medios extraíbles y no extraíbles implementados en cualquier tipo de tecnología para el almacenamiento de información, como instrucciones legibles por ordenador, estructuras de datos, módulos de programas u otros datos. Dichos medios de almacenamiento legibles por ordenador 508 pueden incluir, pero no se limitan a, RAM, ROM, EEPROM, memoria flash u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento óptico, casetes magnéticos, cintas magnéticas, almacenamiento de estado sólido, almacenamiento en disco magnético, sistemas de almacenamiento RAID, matrices de almacenamiento, almacenamiento conectado a la red, redes de área de almacenamiento, almacenamiento en la nube, medios de almacenamiento extraíbles o cualquier otro medio que se pueda usar para almacenar la información deseada y al que pueda acceder un dispositivo informático. Dependiendo de la configuración del sistema de condiciones de purificación de proteínas 502, el medio de almacenamiento legible por ordenador 508 puede ser un tipo de medio de almacenamiento tangible legible por ordenador y puede ser un medio de almacenamiento no transitorio.

El sistema de condiciones de purificación de proteínas 502 puede incluir una o más interfases de comunicación 510 para comunicarse con otros dispositivos informáticos a través de una o más redes (no mostradas), como una o más de Internet, una red de cable, una red satélite, una red de área amplia de comunicación inalámbrica, una red de área local cableada, una red de área local inalámbrica o una red telefónica pública conmutada (PSTN).

Los medios de almacenamiento legibles por ordenador 508 se pueden usar para almacenar cualquier número de componentes funcionales que son ejecutables por uno o más procesadores 506. En muchas implementaciones, estos componentes funcionales comprenden instrucciones o programas que son ejecutables por uno o más procesadores 506 y que, cuando se ejecutan, implementan la lógica operativa para realizar las operaciones atribuidas al sistema de condiciones de purificación de proteínas 502. Los componentes funcionales del sistema de condiciones de purificación de proteínas 502 que se pueden ejecutar en uno o más procesadores 506 para implementar las diversas funciones y características relacionadas con las condiciones determinantes para la purificación de proteínas, como se describen en la presente memoria, incluyen instrucciones de recopilación y almacenamiento de datos de proteínas 512, instrucciones de generación de modelos 514, instrucciones de optimización de condiciones de purificación 516 e instrucciones de muestreo de placas multipocillo 518.

Además, uno o más dispositivos informáticos 504 pueden incluir uno o más dispositivos de entrada/salida (no mostrados). El uno o más dispositivos de entrada/salida pueden incluir un dispositivo de visualización, teclado, un control remoto, un ratón, una impresora, dispositivos de entrada/salida de audio, un altavoz, un micrófono, una cámara, etc.

El sistema de condiciones de purificación de proteínas 502 también puede incluir, o estar acoplado a, un almacén de datos 520 que puede incluir, pero no se limita a, RAM, ROM, EEPROM, memoria flash, uno o más discos duros, unidades de estado sólido, memoria óptica (p. ej., CD, DVD) u otras tecnologías de memoria no transitorias. El almacén de datos 520 puede mantener información que es utilizada por el sistema de condiciones de purificación de proteínas 502 para realizar operaciones relacionadas con la generación de modelos que se pueden utilizar para determinar las condiciones para la purificación de proteínas. Por ejemplo, el almacén de datos 520 puede almacenar datos de entrenamiento 522 que se pueden analizar para generar los modelos.

En el ejemplo ilustrativo de la FIG. 5, los datos de entrenamiento 522 pueden incluir datos de secuencia para proteínas donde las secuencias de aminoácidos se pueden utilizar para generar modelos para determinar las condiciones para la purificación de proteínas. Los datos de secuencia pueden indicar aminoácidos incluidos en diversas posiciones de las proteínas que se analizan para generar los modelos de condiciones de purificación de proteínas. Los datos de entrenamiento 522 también pueden incluir datos de estructura para proteínas donde los datos de estructura se pueden usar para generar modelos para determinar las condiciones para purificar proteínas. Los datos de estructura pueden indicar estructuras secundarias de las proteínas, estructuras terciarias de las proteínas o tanto estructuras secundarias como estructuras terciarias de las proteínas. Los datos de entrenamiento 522 también pueden incluir datos de propiedades biofísicas que incluyen valores de diversas propiedades biofísicas de proteínas que se pueden usar para generar modelos para determinar las condiciones de purificación para diversas proteínas. Además, los datos de entrenamiento 522 pueden incluir valores de pureza y de rendimiento para proteínas que se han obtenido con respecto a un número de materiales de fase estacionaria en diversas combinaciones de niveles de pH y de concentraciones de sal.

Las instrucciones de recopilación y almacenamiento de datos de proteínas 508 pueden ser ejecutadas por uno o más

- procesadores 506 para obtener y almacenar datos relacionados con proteínas que se evalúan para generar modelos para determinar las condiciones para la purificación de proteínas. En algunas implementaciones, las instrucciones de recopilación y almacenamiento de datos de proteínas 508 pueden obtener los datos de entrenamiento 522 de varias fuentes. En ciertas implementaciones, las instrucciones de recopilación y almacenamiento de datos de proteínas 512 pueden obtener datos de secuencia, datos de estructura, datos de propiedades biofísicas, datos de condiciones de purificación, medidas de desempeño para procesos cromatográficos, o combinaciones de estos a partir de uno o más sitios web y/o una o más bases de datos que son depósitos de datos relacionados con las proteínas. En implementaciones adicionales, las instrucciones de recopilación y almacenamiento de datos de proteínas 512 pueden producir una o más interfases de usuario que incluyen uno o más elementos de interfase de usuario para capturar datos de secuencia, datos de estructura, datos de propiedades biofísicas, datos de condiciones de purificación y/o medidas de desempeño para procesos cromatográficos. En implementaciones adicionales, las instrucciones de recopilación y almacenamiento de datos de proteínas 512 pueden obtener datos de secuencia, datos de estructura, datos de propiedades biofísicas, datos de condiciones de purificación, medidas de desempeño para procesos cromatográficos o combinaciones de estos de uno o más dispositivos de almacenamiento de datos. El uno o más dispositivos de almacenamiento de datos pueden incluir dispositivos de almacenamiento de datos extraíbles, como tarjetas de memoria, unidades flash o memorias USB. El uno o más dispositivos de almacenamiento de datos también pueden incluir almacenes de datos acoplados al sistema de condiciones de purificación de proteínas 502 a través de una o más redes, como redes de área local cableadas, redes de área local inalámbricas, redes de área amplia inalámbricas o combinaciones de estas.
- Las instrucciones de generación de modelos 514 pueden ser ejecutadas por uno o más procesadores 506 para generar uno o más modelos para determinar las condiciones para la purificación de proteínas. Las instrucciones de generación de modelos 514 puede utilizar los datos de entrenamiento 522 para determinar las relaciones entre las características de las proteínas y el desempeño de los materiales de fase estacionaria de las técnicas cromatográficas. En diversas implementaciones, las instrucciones de generación de modelos 514 pueden identificar relaciones entre secuencias de proteínas, estructuras de proteínas, propiedades medidas de proteínas o combinaciones de estas y el rendimiento y/o la pureza de una o más técnicas cromatográficas. Las instrucciones de generación de modelos 514 también pueden determinar las relaciones entre las condiciones de purificación, como valores de pH y concentraciones de sal, y el desempeño de los materiales de fase estacionaria. Por ejemplo, las instrucciones de generación de modelos 514 pueden determinar relaciones entre diversas combinaciones de concentraciones de sal y valores de pH y de rendimiento y/o pureza para una o más técnicas cromatográficas.
- Las instrucciones de generación de modelos 514 también pueden determinar el grado de influencia que tienen características particulares sobre el desempeño de los materiales de fase estacionaria. En implementaciones particulares, las instrucciones de generación de modelos 514 pueden determinar un grado de influencia que ciertas porciones de secuencias de proteínas tienen sobre el rendimiento y la pureza de las proteínas con respecto a uno o más materiales de fase estacionaria. Además, las instrucciones de generación de modelos 514 puede determinar el impacto que tienen diversas estructuras de proteínas sobre el rendimiento y la pureza con respecto a uno o más materiales de fase estacionaria. Además, las instrucciones de generación de modelos 514 pueden determinar el impacto que las propiedades medidas, como turbidez, peso molecular y estabilidad de proteínas, tienen sobre la pureza y/o el rendimiento en relación con diversos materiales de fase estacionaria. En ciertas implementaciones, las instrucciones de generación de modelos 514 pueden determinar un grado de influencia que las condiciones de purificación, como niveles de pH y de concentraciones de sal, tienen sobre el desempeño de uno o más materiales de fase estacionaria. En diversas implementaciones, se puede tener en cuenta en los modelos la influencia de una característica particular de la proteína en el rendimiento y/o la pureza de un proceso de purificación asignando ciertos pesos a las variables respectivas asociadas con las características de la proteína incluidas en el modelo.
- Después de determinar las relaciones entre una serie de características relacionadas con las proteínas y el desempeño de los materiales de fase estacionaria, las instrucciones de generación de modelos 514 pueden generar modelos individuales para cada material de fase estacionaria. Los modelos pueden determinar el rendimiento y/o la pureza de las proteínas en función de una o más entradas correspondientes a las características de las proteínas. En algunas situaciones, se puede obtener una cantidad mínima de información para una proteína con el fin de implementar el modelo para la proteína. Por ejemplo, un modelo puede utilizar información relacionada con al menos una porción de una secuencia de una proteína (p. ej., ciertas posiciones de la proteína), un número de láminas beta asociadas con la proteína y una estabilidad de la proteína como se indica mediante fluorimetría diferencial de barrido (DSF) para predecir el rendimiento y la pureza de la proteína en intervalos específicos de niveles de pH y de concentraciones de sal. Continuando con este ejemplo, para implementar este modelo, el sistema de condiciones de purificación de proteínas 502 puede obtener información relacionada con la porción particular de la secuencia de la proteína que corresponde a la porción de secuencia asociada con el modelo, el número de láminas beta y los datos de DSF de la proteína para determinar el rendimiento y la pureza de la proteína cuando se purifica con el material de fase estacionaria asociado con el modelo al pH y concentraciones de sal especificados. En determinadas implementaciones, los modelos pueden incluir modelos polinomiales.
- Las instrucciones de optimización de las condiciones de purificación 516 pueden ser ejecutadas por uno o más procesadores 506 para determinar las condiciones para la purificación de proteínas que maximizan el rendimiento y la pureza mientras minimizan el costo del proceso de purificación. El costo del proceso de purificación puede estar relacionado con el valor monetario de implementar el proceso de purificación con un material de fase estacionaria y

también puede incluir, en algunos escenarios, una dificultad para trabajar con el material de fase estacionaria respectivo y/o una durabilidad del material de fase estacionaria. En algunas implementaciones, las instrucciones de optimización de las condiciones de purificación 516 pueden determinar un conjunto optimizado de condiciones para purificar proteínas de acuerdo con uno o más modelos. Los modelos usados para determinar las condiciones para purificar las proteínas se pueden generar usando al menos porciones de los datos de entrenamiento 522.

Las instrucciones de optimización de las condiciones de purificación 516 pueden determinar uno o más materiales de fase estacionaria para utilizar para purificar una proteína particular. En algunas implementaciones, las instrucciones de optimización de las condiciones de purificación 516 pueden determinar que se pueda utilizar una combinación de materiales de fase estacionaria para purificar una proteína. A modo de ilustración, las instrucciones de optimización de las condiciones de purificación 516 pueden determinar que una proteína se pueda purificar utilizando un primer material de fase estacionaria seguido de la purificación usando un segundo material de fase estacionaria. Además, las instrucciones de optimización de las condiciones de purificación 516 pueden determinar condiciones optimizadas para el proceso de purificación. Por ejemplo, las instrucciones de optimización de las condiciones de purificación 516 pueden determinar un intervalo de valores de pH y un intervalo de concentraciones de sal que maximicen el rendimiento y la pureza del proceso de purificación usando uno o más materiales de fase estacionaria, mientras que minimizan el costo. En algunos casos, las instrucciones de optimización de las condiciones de purificación 516 pueden determinar una combinación única de un valor de pH y de una concentración de sal que maximice el rendimiento y la pureza del proceso de purificación usando uno o más materiales de fase estacionaria, mientras que minimizan el costo.

Las instrucciones de muestreo de la placa multipocillo 518 pueden ser ejecutadas por uno o más procesadores 506 para determinar una muestra de pocillos seleccionados a partir de un número mayor de pocillos en una placa multipocillo que proporciona condiciones para purificar proteínas usando un material de fase estacionaria. En diversas implementaciones, una placa multipocillo puede incluir un número de pocillos, cada uno de los cuales está asociado con un conjunto de condiciones de purificación. Por ejemplo, los pocillos individuales de una placa multipocillo se pueden asociar con un material de fase estacionaria, un nivel de pH y una concentración de sal. Las instrucciones de muestreo de la placa multipocillo 518 pueden determinar un subconjunto de los pocillos individuales que modelan un número mayor de pocillos de la placa multipocillo que también están asociados con el material de la fase estacionaria. A modo de ilustración, las instrucciones de muestreo de la placa de multipocillo 518 pueden identificar 5 pocillos de 24 que están asociados con un material de fase estacionaria particular que se puede usar para reproducir los resultados obtenidos de todos los 24 pocillos. Es decir, las instrucciones de muestreo de la placa multipocillo 518 pueden identificar 5 conjuntos de condiciones de purificación que pueden reproducir los resultados de 24 conjuntos de condiciones de purificación con un grado mínimo de error para un material de fase estacionaria particular. En diversas implementaciones, las instrucciones de muestreo de la placa multipocillo 518 pueden realizar un proceso iterativo para comparar los resultados de rendimiento y pureza obtenidos a partir de un subconjunto de pocillos de una placa multipocillo con los resultados de rendimiento y pureza obtenidos de un mayor número de pocillos de la placa multipocillo. Las instrucciones de muestreo de la placa multipocillo 518 pueden realizar comparaciones entre los datos de referencia que se obtienen de un conjunto de pocillos de una placa multipocillo y varios subconjuntos de los pocillos de la placa multipocillo hasta que se minimice un grado de error entre los valores de rendimiento y de pureza para un subconjunto particular de pocillos con respecto a los datos de referencia. Después de determinar un subconjunto de condiciones de purificación para un material de fase estacionaria particular que minimice el error con respecto a los datos de referencia, se puede utilizar el mismo subconjunto de condiciones de purificación en la purificación de proteínas adicionales para el material de fase estacionaria.

La FIG. 6 ilustra un proceso de ejemplo para determinar las condiciones para la purificación de proteínas. Este proceso (así como cada proceso descrito en la presente memoria) se ilustra como diagramas de flujo lógicos, cada operación de este representa una secuencia de operaciones que pueden, al menos en parte, implementarse en hardware, software o una combinación de estos. En el contexto del software, las operaciones representan instrucciones ejecutables por ordenador almacenadas en uno o más medios de almacenamiento legibles por ordenador que, cuando son ejecutadas por uno o más procesadores, realizan las operaciones citadas. Generalmente, las instrucciones ejecutables por ordenador incluyen rutinas, programas, objetos, componentes, estructuras de datos y similares que realizan funciones particulares o implementan tipos de datos abstractos particulares. El orden en el que se describen las operaciones no pretende ser interpretado como una limitación, y cualquier número de las operaciones descritas se puede combinar en cualquier orden y/o en paralelo para implementar el proceso.

La FIG. 6 es un diagrama de flujo de un proceso de ejemplo 600 para determinar las condiciones de purificación de proteínas usando uno o más procesos cromatográficos. En 602, el proceso 600 incluye proporcionar una placa que tiene una pluralidad de pocillos, incluyendo cada pocillo (i) un material de fase estacionaria de una pluralidad de materiales de fase estacionaria para sistemas de cromatografía en columna, (ii) una cantidad de una proteína, y (iii) una cantidad de una solución que tiene un pH y una concentración de una sal. Cada pocillo de la pluralidad de pocillos puede tener una combinación diferente del material de fase estacionaria, de pH y de concentración de la sal.

En 604, el proceso 600 puede incluir determinar una medida de desempeño para cada pocillo de la pluralidad de pocillos. La medida del desempeño de un pocillo individual de la pluralidad de pocillos puede indicar la adsorción de la proteína con respecto al material de fase estacionaria incluido en el pocillo individual. En implementaciones particulares, la medida del desempeño puede indicar un rendimiento y/o una pureza de una técnica cromatográfica que utiliza el material de fase estacionaria incluido en el pocillo.

5 En 606, el proceso 600 puede incluir determinar, basándose al menos parcialmente en medidas individuales de desempeño para la pluralidad de pocillos, las condiciones de cromatografía para un subconjunto de los pocillos. Las condiciones de cromatografía pueden incluir un valor de pH y una concentración de sal para cada pocillo del subconjunto de pocillos. El subconjunto de pocillos se puede determinar seleccionando iterativamente varios subconjuntos de pocillos e identificando el subconjunto de pocillos que minimiza el error con respecto a un conjunto de datos de referencia.

10 En 608, el proceso 600 puede incluir generar una pluralidad de modelos para la pluralidad de materiales de fase estacionaria para predecir el rendimiento y la pureza de uno o más procesos de cromatografía. La pluralidad de modelos se puede generar basándose al menos parcialmente en al menos una de las secuencias de aminoácidos de proteínas, una o más estructuras de las proteínas o datos de caracterización de las proteínas. Los datos de caracterización pueden incluir valores obtenidos a partir de instrumentos analíticos que indican propiedades de las proteínas.

15 En 610, el proceso 600 puede incluir generar, basándose al menos parcialmente en las condiciones de cromatografía y la pluralidad de modelos, al menos un valor de pH del intervalo de valores de pH, al menos una concentración de sal del intervalo de concentraciones de sal, y uno o más materiales de fase estacionaria de la pluralidad de materiales de fase estacionaria para maximizar el rendimiento y la pureza de la proteína.

### Ejemplos experimentales

La FIG. 7 ilustra una serie de interfases de usuario que indican el rendimiento y la pureza de una proteína con respecto a un número de niveles de pH y de concentraciones de sal. Se indican las condiciones óptimas para la purificación de la proteína.

20 La FIG. 8 ilustra los resultados de diferentes modelos que se pueden usar para determinar el rendimiento y la pureza de las proteínas usando cromatografía multimodo a varios niveles de pH y de concentraciones de sal. 802 indica un modelo basado en datos a partir de cada pocillo de un conjunto de pocillos incluidos en una placa multipocillo. 804 indica un modelo basado en datos a partir de un primer subconjunto de pocillos seleccionados del mismo conjunto de pocillos usados para producir el modelo indicado por 802. El modelo indicado en 804 se basa en aproximadamente el 80% de los datos utilizados para producir el modelo indicado por 802. 806 indica un modelo derivado a partir de datos de un segundo subconjunto de pocillos seleccionados del mismo conjunto de pocillos usados para producir el modelo indicado por 802. El modelo indicado en 806 se basa en aproximadamente el 20% de los datos utilizados para producir el modelo indicado por 802.

### Introducción

30 La detección de alto rendimiento (HTS) y la miniaturización son estrategias bien establecidas para optimizar el desarrollo de procesos corriente abajo. Durante el desarrollo de rutina, la utilización de tecnologías HTS en todo el proceso corriente abajo puede mejorar la comprensión general del proceso y evitar mejor las brechas en el diseño del proceso. Durante la fase inicial de desarrollo, los métodos HTS pueden facilitar el desarrollo de múltiples moléculas en paralelo e informar flujos de trabajo eficaces y apropiados para cada fase. El presente trabajo describe un método para el desarrollo integrado de procesos corriente abajo en un único experimento con placa de filtro de 96 pocillos. El diseño de la placa permite la recopilación de datos sobre la estabilidad coloidal, química y térmica de un producto en toda la gama de condiciones de la solución corriente abajo, incluida la inactivación viral con pH bajo. Además, en la misma placa, se realizaron estudios de unión por lotes para investigar las interacciones proteína-adsorbente en cinco tipos de medios cromatográficos.

### 40 Materiales y métodos

El método se implementó usando una placa de filtro de 96 pocillos con 50 µl de resina en cada pocillo. Todo el método, incluida la preparación de la carga y del tampón, se automatizó en un sistema robótico de manipulación de líquidos. Después de la preparación de la carga de solución, la placa se cargó y se incubó durante 60 minutos, seguido de la recuperación del material no unido. A continuación, se incubó la resina de los pocillos con una solución de extracción o con un tampón de elución, dependiendo del modo de cromatografía, seguido de la recogida del material retirado o eluido. Las muestras de carga, no unidas, eluidas y retiradas se analizaron para determinar la concentración de proteínas y de impurezas diana (p. ej., HMW, HCP o clips). Las resinas evaluadas en la placa de 96 pocillos incluyen cromatografías de afinidad de proteína A y de intercambio catiónico (CEX) en modo de unión/elución, y cromatografías de intercambio aniónico (AEX), de interacción hidrofóbica (HIC) y de modo mixto (MMC) en modo de flujo continuo (FT). La Figura 9 presenta la configuración de la placa y la Tabla 1 muestra las condiciones operativas probadas.

La Figura 9 muestra una configuración de placa de filtro de 96 pocillos. Los colores indican áreas de la placa dedicadas a resinas específicas (afinidad con proteína A: rojo, AEX: amarillo, HIC: verde, MMC: violeta, resina CEX 1: azul claro y gris oscuro, y resina CEX 2: azul oscuro y gris claro).

Tabla 1. Condiciones operativas y resultados del experimento con una sola placa de filtro de 96 pocillos

Resina	pH	Sal (mM)	Modo	Carga (g/l)	Pocillos	Resultado
Proteína A	3,3-3,8	0	BE <sup>1</sup>	20	6	Recuperación, HMW, estabilidad a pH bajo
CEX 1	5	0-100	Unión	≥ 80	3	Capacidad estática
	5	0-650 <sup>3</sup>	BE	20	9	Recuperación, pureza
CEX 2	5	0-100	Unión	≥ 80	3	Capacidad estática
	5	0-500 <sup>3</sup>	BE	20	9	Recuperación Pureza
AEX	6-8	33-200	FT <sup>2</sup>	5	18	Recuperación, pureza, eliminación de contaminantes, Kp, estabilidad de la solución (carga)
HIC	5-8	25-400	FT	5	24	
MMC	5-8	25-400	FT	5	24	

<sup>1</sup> BE = Unido y eluido  
<sup>2</sup> FT = Flujo continuo  
<sup>3</sup> Unión a 0 mM de sal seguida de elución gradual a 650 mM y 500 mM de sal en 9 pocillos para la resina CEX 1 y la resina CEX 2, respectivamente

#### Estabilidad de la solución

Se puede obtener un conocimiento significativo sobre las posibles condiciones operativas evaluando la estabilidad de la solución de moléculas. Se usó la estabilidad coloidal, química y térmica en todas las condiciones potenciales de la solución en el proceso corriente abajo para identificar el espacio operativo donde el producto es más estable y para identificar las condiciones que minimicen la pérdida de rendimiento y la agregación. La Figura 10 compara la estabilidad coloidal mediante dispersión de luz y la estabilidad química por cromatografía de exclusión molecular (SEC) para 2 anticuerpos monoclonales (mAb) diferentes. Como se muestra en este ejemplo, el mAb 1 es relativamente estable con respecto a la precipitación y la agregación en todo el intervalo de condiciones operativas corriente abajo evaluadas. Alternativamente, el mAb 2 precipita a pH alto y fuerza iónica baja y muestra una agregación significativa a pH alto y fuerza salina alta, lo que indica que se deben evitar estas condiciones de solución durante el diseño del proceso corriente abajo. La Figura 11 resalta la diferencia en la estabilidad térmica mediante fluorimetría diferencial de barrido (DSF) para un mAb IgG1 y una proteína de fusión-Fc. Como se muestra en este ejemplo, el mAb 3 exhibe una estabilidad térmica típica para una IgG1 junto con una mayor estabilidad térmica a medida que aumenta el pH. La proteína de fusión-Fc, además de una Tm global más baja, muestra una disminución en la estabilidad térmica con un aumento en el pH y la concentración de la sal.

La Figura 10 muestra una comparación de la estabilidad de la solución para 2 mAbs para el intervalo de condiciones del proceso corriente abajo evaluadas. Los paneles A y B presentan turbidez y HMW, respectivamente, para el mAb 1. Los paneles C y D presentan turbidez y HMW para el mAb 2.

La Figura 11 ilustra la estabilidad térmica mediante DSF para el intervalo de condiciones del proceso corriente abajo evaluadas. Panel A: mAb 3; panel B: proteína de fusión-Fc 1.

#### Unión por lotes

La placa de 96 pocillos se usó para realizar estudios de unión por lotes para investigar las interacciones proteína-adsorbente en cinco tipos de medios cromatográficos en una amplia gama de condiciones de pH y sal. Este formato se usó para seleccionar las condiciones de elución de la proteína A; evaluar la capacidad estática de la resina, la fuerza de elución del gradiente y la selectividad para 2 resinas de intercambio catiónico; y evaluar la selectividad del producto y los contaminantes para tres resinas diferentes operadas en modo de flujo continuo para una sola molécula. Las Figuras 12 y 13 presentan un ejemplo de los datos obtenidos en el experimento de placa única. Estos datos se pueden usar para estimar la recuperación de la proteína A y la sensibilidad de la molécula a la inactivación viral con pH bajo; seleccionar una resina CEX en función de su fuerza de elución, recuperación y selectividad para contaminantes; y, finalmente, identificar el flujo continuo óptimo de la resina y las condiciones de operación. El método integrado basado en placas presentado en la presente memoria proporciona información sobre los puntos de ajuste para las operaciones unitarias individuales, así como de la sensibilidad a los intervalos operativos. Además, se puede emplear el modo de emparejamiento para seleccionar resinas y condiciones operativas durante todo el proceso (Fig. 14).

La Figura 12 ilustra ejemplos de datos generados durante un experimento de unión de un solo lote. Panel A: Recuperación de proteína A frente a pH de elución y pool de HMW después de una retención a pH bajo durante 60 minutos. Panel B: Pseudo-cromatogramas para 2 resinas CEX. Panel C: Capacidad estática y balance de masa HMW para 2 resinas CEX. Panel D: Gráfico de pureza (HMW) para 2 resinas CEX.

La Figura 13 ilustra gráficos de contorno de recuperación y de contaminantes para resinas AEX, HIC y MMC operadas en modo de flujo continuo.

5 La Figura 14 ilustra la pureza frente al rendimiento para modos emparejados en todo el proceso corriente abajo. Cada punto de datos representa la pureza mediante SEC (% de monómero) y la recuperación para diferentes opciones de proceso. Los puntos morados son la proteína A emparejada con CEX. Los otros colores indican proteína A y CEX emparejadas con las diferentes opciones de refinamiento (AEX, HIC o MMC). panel: resultados para toda la gama de opciones de proceso. Panel derecho: vista ampliada del cuadro de llamada rojo, que indica las condiciones que producen >70 % de recuperación general y > 98 % de pureza.

10 El método de enfoque presentado permite un ciclo de desarrollo ultrarrápido en donde los científicos de desarrollo de procesos determinan la capacidad de fabricación de las moléculas y los intervalos operativos del proceso para todo un proceso corriente abajo en un solo experimento basado en placas. Los recursos necesarios para este método son aproximadamente 100 mg de proteína y 1 día de ejecución. Por el contrario, para generar datos comparables usando sistemas de cromatografía tradicionales a escala de banco y columnas de tamaño reducido podría requerir de 5 a 8 semanas de tiempo de ejecución y más de 50 g de proteína, dependiendo de las estimaciones.

15 Cuando se combina con una cromatografía de columna robotizada automatizada y miniaturizada para verificar las condiciones operativas, el método presentado se integra en una estrategia de desarrollo apropiada para la fase que puede aumentar drásticamente la eficacia del ciclo de desarrollo corriente abajo sin sacrificar la comprensión y la solidez del proceso.

20 El tema descrito anteriormente se proporciona únicamente a modo de ilustración y no se debe interpretar como limitante. Además, el tema reivindicado no se limita a implementaciones que solucionen cualquiera o todas las desventajas señaladas en cualquier parte de esta descripción. Se pueden realizar diversas modificaciones y cambios al tema descrito en la presente memoria sin seguir las configuraciones y aplicaciones de ejemplo ilustradas y descritas, y sin apartarse del verdadero alcance de la presente invención, que se establece en las siguientes reivindicaciones.

## REIVINDICACIONES

1. Un método que comprende:

5 generar datos de rendimiento y datos de pureza para varias proteínas (102, 302), indicando los datos de rendimiento y los datos de pureza el rendimiento y la pureza para cada una de una pluralidad de proteínas (102, 302) en un intervalo de valores de pH, en un intervalo de concentraciones de sal, y para un material de fase estacionaria de una técnica de cromatografía en columna;

obtener datos de secuencia (108) que indican al menos una porción de secuencias de aminoácidos de proteínas individuales (102, 302) de la pluralidad de proteínas;

10 obtener datos de estructura (110) que indican una o más estructuras exhibidas por proteínas individuales (102, 302) de la pluralidad de proteínas;

generar datos de caracterización (304, 404) para la pluralidad de proteínas (102, 302), incluyendo los datos de caracterización (304, 404) valores obtenidos a partir de pruebas analíticas que indican valores de propiedades de la pluralidad de proteínas (102, 302);

15 generar un modelo (406) para predecir, a uno o más valores de pH y uno o más valores de concentración de sal (414), el rendimiento y la pureza de una proteína adicional (234) para el material de fase estacionaria, en donde el modelo (406) incluye un componente de secuencia que indica una similitud entre las secuencias de aminoácidos de la pluralidad de proteínas (102, 302) y una secuencia de aminoácidos adicional de la proteína adicional (234) y un componente de caracterización que indica similitudes entre valores de las propiedades para la pluralidad de proteínas (102, 302) y valores adicionales de las propiedades de la proteína adicional (234).

20 2. El método de la reivindicación 1, en donde los datos de rendimiento y los datos de pureza se determinan ejecutando un número de columnas de cromatografía para un número de proteínas diferentes usando el material de fase estacionaria en el intervalo de valores de pH y en el intervalo de concentraciones de sal.

25 3. El método de la reivindicación 1 o la reivindicación 2, en donde los datos de rendimiento y los datos de pureza se determinan analizando soluciones obtenidas de los pocillos de una placa multipocillo (202), pocillos individuales de la placa multipocillo (202) que incluyen una cantidad de una solución que incluye la proteína (102, 302) y que tiene un nivel de pH y una concentración de sal y la placa multipocillo (202) que incluye una cantidad del material de fase estacionaria.

4. El método de cualquiera de las reivindicaciones 1 a 3, en donde la una o más estructuras incluyen regiones hidrófobas, regiones polares, pliegues, giros, bucles o combinaciones de estos.

30 5. El método de cualquiera de las reivindicaciones 1 a 4, en donde:

la proteína adicional (234) y el número de proteínas (102, 302) incluyen cada una una estructura secundaria común;

al menos una porción de la secuencia de aminoácidos de la proteína adicional (234) y al menos una porción de las secuencias de aminoácidos del número de proteínas (102, 302) tienen al menos un 75% de identidad;

35 las pruebas analíticas producen valores de una o más propiedades biofísicas del número de proteínas (102, 302); y

un intervalo de valores para una propiedad biofísica de una o más propiedades biofísicas para la proteína adicional (234) está dentro de al menos aproximadamente el 90 % de un intervalo de valores para la propiedad biofísica para el número de proteínas (102, 302).

40 6. El método de cualquiera de las reivindicaciones 1 a 5, en donde los datos de caracterización (304, 404) están relacionados con al menos una de las mediciones de fluorimetría diferencial de barrido para el número de proteínas (102, 302), el peso molecular del número de proteínas (102, 302), turbidez de soluciones que incluyen el número de proteínas (102, 302), o combinaciones de estas.

45 7. El método de cualquiera de las reivindicaciones 1-6, en donde generar un modelo (406) para predecir el rendimiento y la pureza de una proteína adicional (234), a uno o más valores de pH y uno o más valores de concentración de sal, para el material de fase estacionaria, comprende:

determinar una relación entre una estructura de una o más estructuras exhibidas por proteínas individuales (102, 302) de la pluralidad de proteínas y al menos uno del rendimiento y la pureza para cada una de la pluralidad de proteínas (102, 302) a lo largo del intervalo de valores de pH, en el intervalo de concentraciones de sal y para el material de fase estacionaria de la técnica de cromatografía en columna; y

50 en donde el modelo (406) incluye al menos una variable que indica una relación entre la estructura y el rendimiento y la pureza para el material de fase estacionaria.

8. Un sistema (502) que comprende:

uno o más procesadores (506); y

5 uno o más medios de almacenamiento no transitorios legibles por ordenador (508) que almacenan instrucciones legibles por ordenador que, cuando son ejecutadas por uno o más procesadores (506), realizan operaciones que comprenden:

generar datos de rendimiento y datos de pureza para un número de proteínas (102, 302), indicando los datos de rendimiento y los datos de pureza el rendimiento y la pureza para cada una de una pluralidad de proteínas en un intervalo de valores de pH, en un intervalo de concentraciones de sal, y para un material de fase estacionaria de una técnica de cromatografía en columna;

10 obtener datos de secuencia (108) que indican al menos una porción de secuencias de aminoácidos de proteínas individuales (102, 302) de la pluralidad de proteínas;

obtener datos de estructura (110) que indican una o más estructuras exhibidas por proteínas individuales (102, 302) de la pluralidad de proteínas;

15 generar datos de caracterización (304) para la pluralidad de proteínas, incluyendo los datos de caracterización (304) valores obtenidos a partir de pruebas analíticas que indican valores de propiedades de la pluralidad de proteínas; y

20 generar un modelo (406) para predecir, a uno o más valores de pH y uno o más valores de concentración de sal, el rendimiento y la pureza de una proteína adicional (234) para el material de fase estacionaria, en donde el modelo (406) incluye un componente de secuencia que indica una similitud entre las secuencias de aminoácidos de la pluralidad de proteínas y una secuencia de aminoácidos adicional de la proteína adicional (234) y un componente de caracterización que indica similitudes entre valores de las propiedades para la pluralidad de proteínas y valores adicionales de las propiedades para la proteína adicional (234).

9. El sistema (502) de la reivindicación 8, en donde:

25 el modelo (406) es uno de una pluralidad de modelos (120, 122, 124) para predecir el rendimiento y la pureza para cada una de la pluralidad de proteínas (102, 302) en el intervalo de valores de pH y en el intervalo de concentraciones de sal;

cada modelo (120, 122, 124) de la pluralidad de modelos está asociado con un material de fase estacionaria individual de una pluralidad de materiales de fase estacionaria; y

30 cada modelo (120, 122, 124) de la pluralidad de modelos está asociado con un material de fase estacionaria diferente de la pluralidad de materiales de fase estacionaria.

10. El sistema (502) de la reivindicación 9, en donde uno o más medios de almacenamiento no transitorios legibles por ordenador (508) almacenan instrucciones adicionales legibles por ordenador que, cuando son ejecutadas por uno o más procesadores (506), realizan operaciones adicionales que comprenden:

35 generar, usando la pluralidad de modelos (120, 122, 124) condiciones de purificación que incluyen al menos un valor de pH del intervalo de valores de pH, al menos una concentración de sal del intervalo de concentraciones de sal, y uno o más materiales de fase estacionaria de la pluralidad de materiales de fase estacionaria para la proteína adicional (234).

11. El sistema (502) de la reivindicación 10, en donde uno o más medios de almacenamiento no transitorios legibles por ordenador (508) almacenan instrucciones adicionales legibles por ordenador que, cuando son ejecutadas por uno o más procesadores (506), realizan operaciones adicionales que comprenden:

determinar uno o más materiales de fase estacionaria basándose al menos parcialmente en la durabilidad de los materiales de fase estacionaria individuales de la pluralidad de materiales de fase estacionaria;

determinar una cantidad de materiales de fase estacionaria individuales a utilizar en una columna de cromatografía para purificar la proteína; y

45 determinar un tamaño de la columna de cromatografía.

12. El sistema (502) de la reivindicación 11, en donde las condiciones de purificación indican que la proteína se va a purificar utilizando un primer material de fase estacionaria a un primer valor de pH y una primera concentración de sal seguido de la purificación de la proteína utilizando un segundo material de fase estacionaria diferente del primer material de fase estacionaria, a un segundo valor de pH que es diferente del primer valor de pH y a una segunda concentración de sal que es diferente de la primera concentración de sal.

50

13. El sistema (502) de cualquiera de las reivindicaciones 8-12, en donde uno o más medios de almacenamiento no transitorios legibles por ordenador (508) almacenan instrucciones adicionales legibles por ordenador que, cuando son ejecutadas por uno o más procesadores (506), realizan operaciones adicionales que comprenden:

5            determinar una medida del desempeño para cada pocillo de una pluralidad de pocillos de una placa (202), indicando la medida del desempeño de un pocillo individual de la pluralidad de pocillos la adsorción de la proteína con respecto al material de fase estacionaria incluido en el pocillo individual;

            en donde los pocillos individuales de la pluralidad de pocillos incluyen un material de fase estacionaria de un sistema de cromatografía en columna, una cantidad de una o más proteínas y una cantidad de una solución que tiene un pH y una concentración de una sal; y

10            en donde cada pocillo de la pluralidad de pocillos tiene una combinación diferente del material de fase estacionaria, de pH y de concentración de la sal.

14. El sistema (502) de la reivindicación 13, en donde la medida del desempeño incluye un rendimiento, una pureza, una característica de la solución después de un período de tiempo, una propiedad de una cantidad restante de la proteína incluida en la solución después del período de tiempo, o combinaciones de estos.

15            15. El sistema (502) de la reivindicación 13, en donde uno o más medios de almacenamiento no transitorios legibles por ordenador (508) almacenan instrucciones adicionales legibles por ordenador que, cuando son ejecutadas por uno o más procesadores (506), realizan operaciones adicionales que comprenden generar la pluralidad de modelos (120, 122, 124) basados al menos parcialmente en la medida del desempeño para cada pocillo de la pluralidad de pocillos.

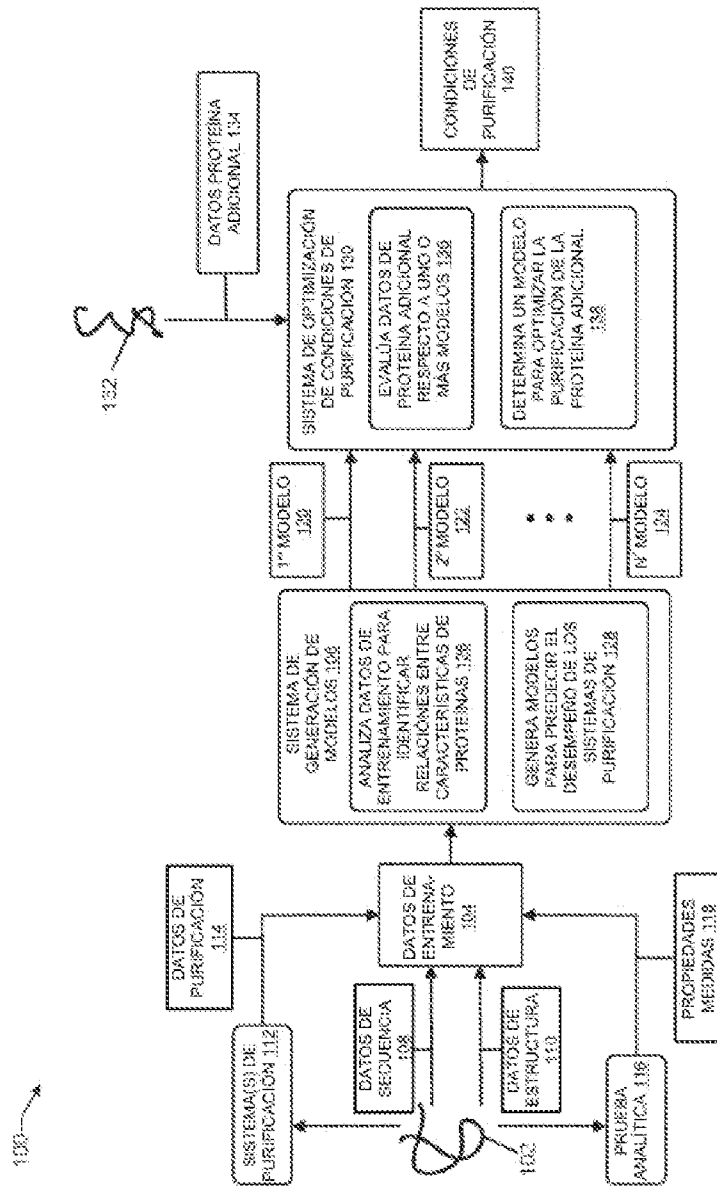


FIG. 1

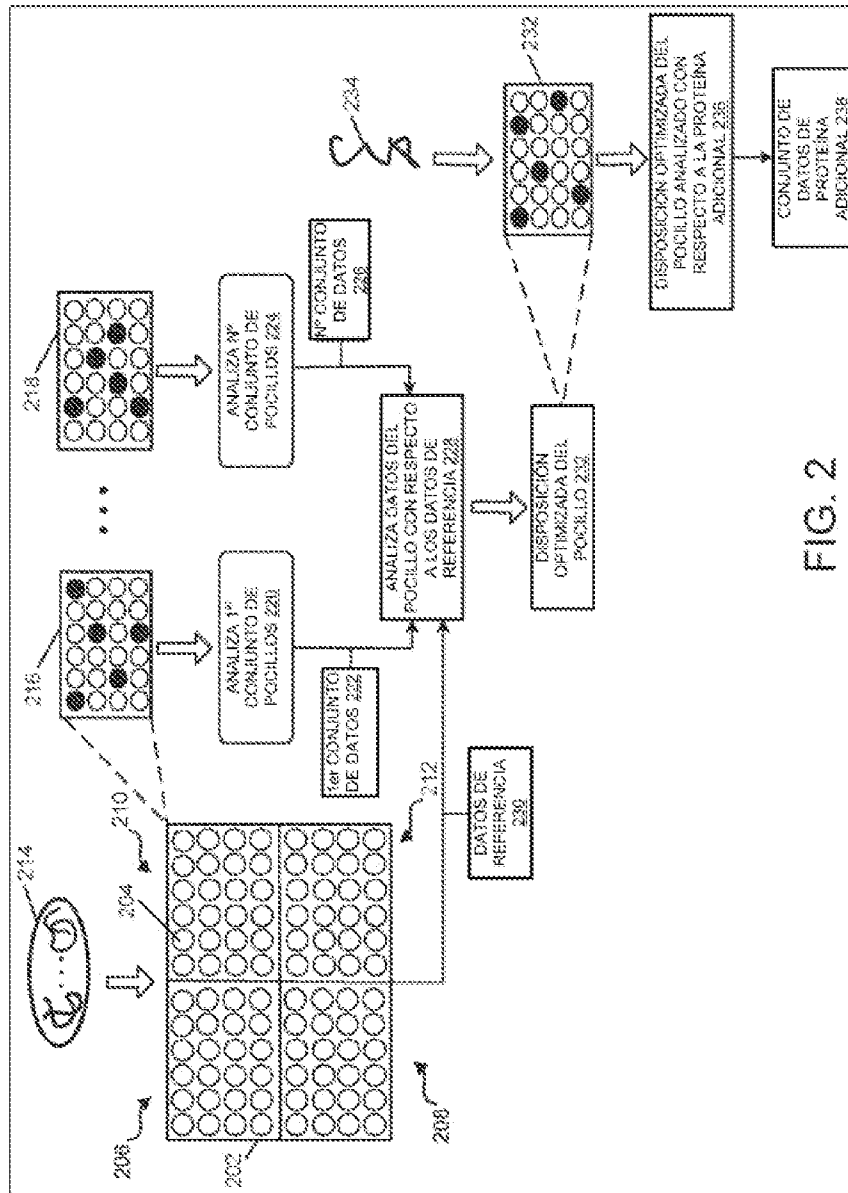


FIG. 2

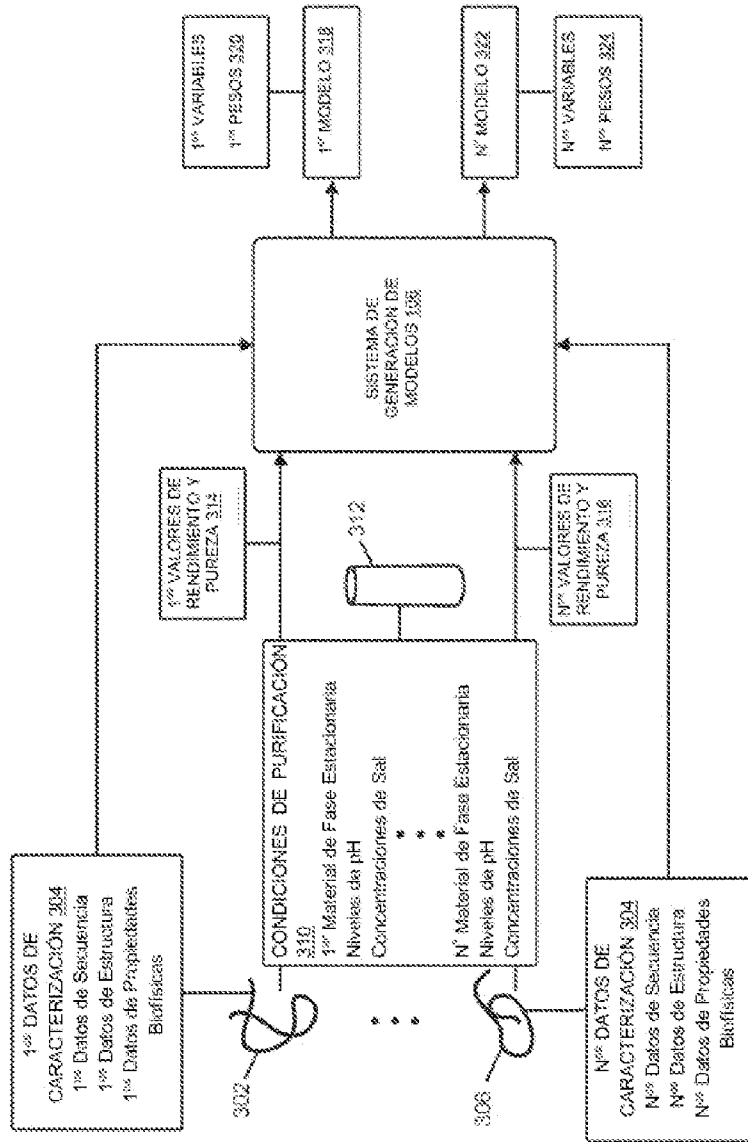


FIG. 3

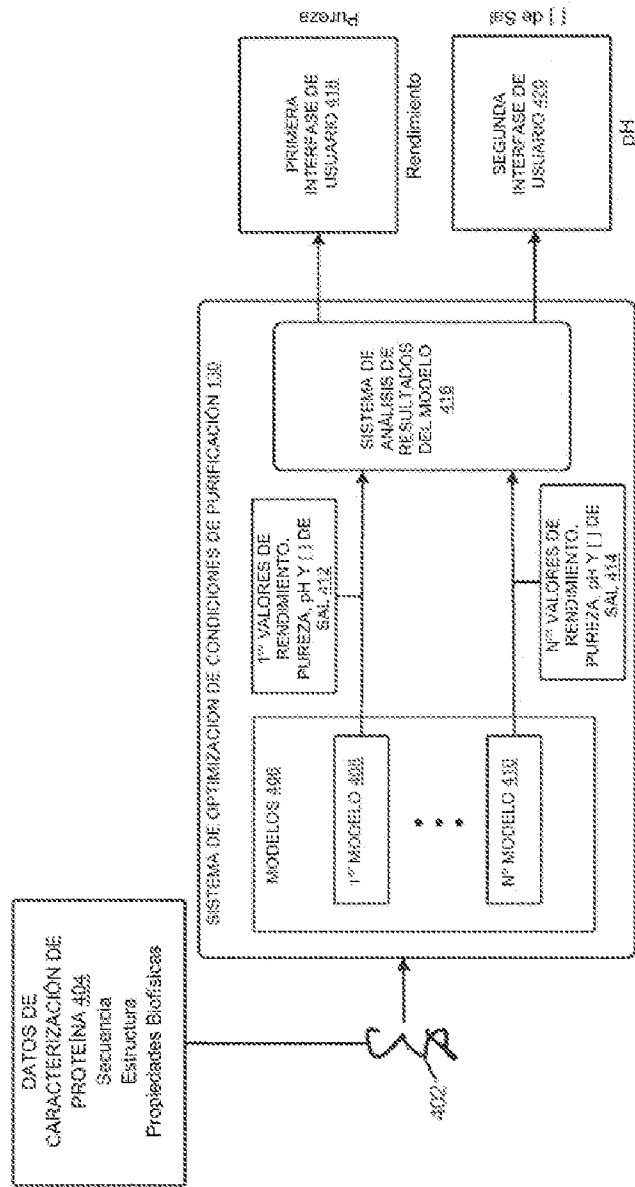


FIG. 4

500 →

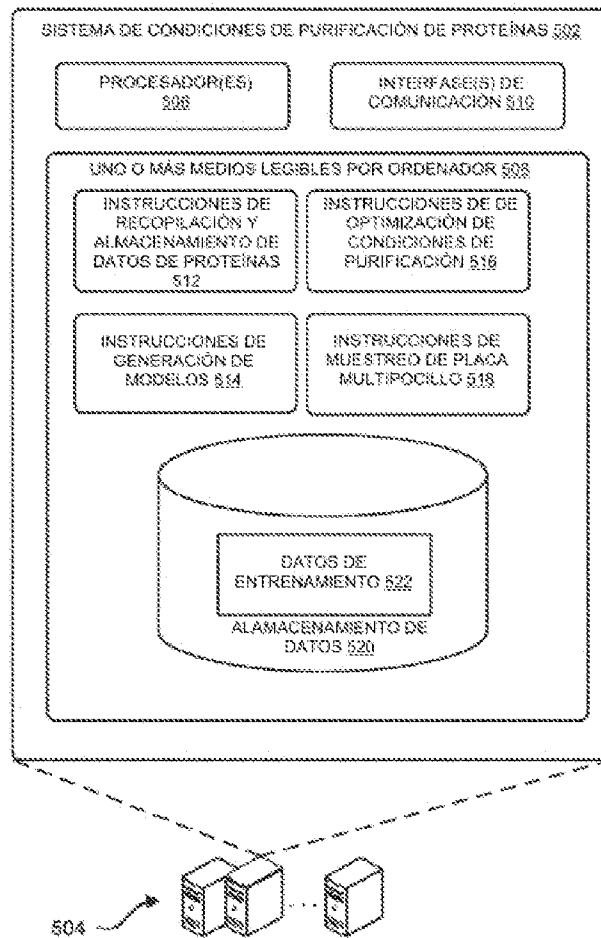


FIG. 5

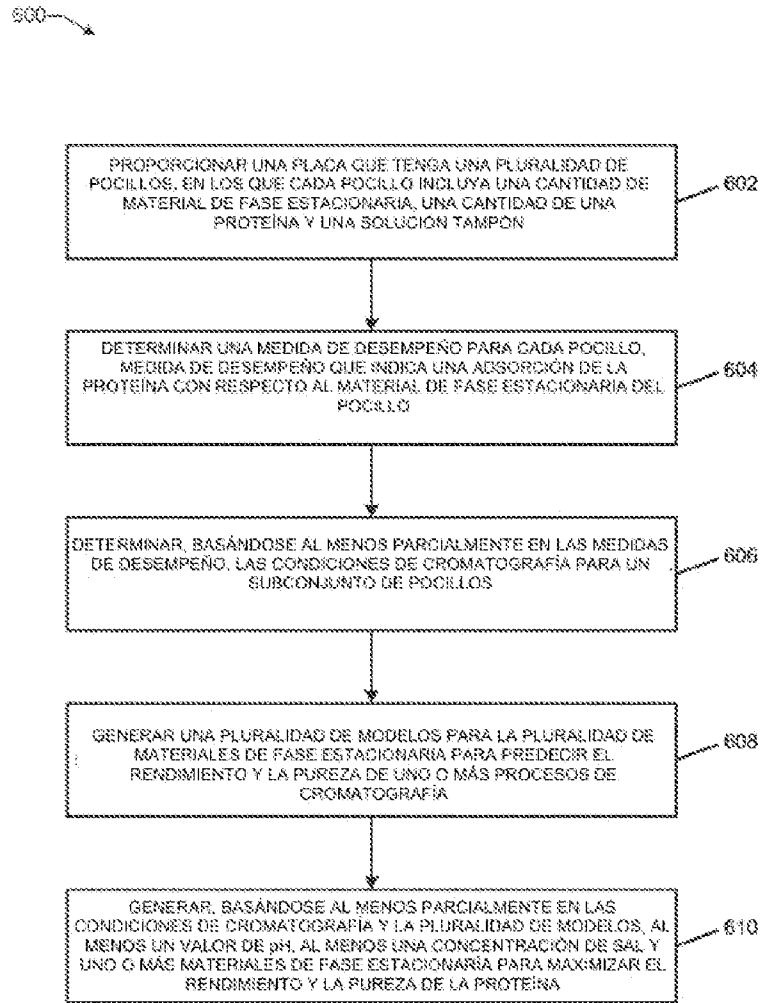


FIG. 6

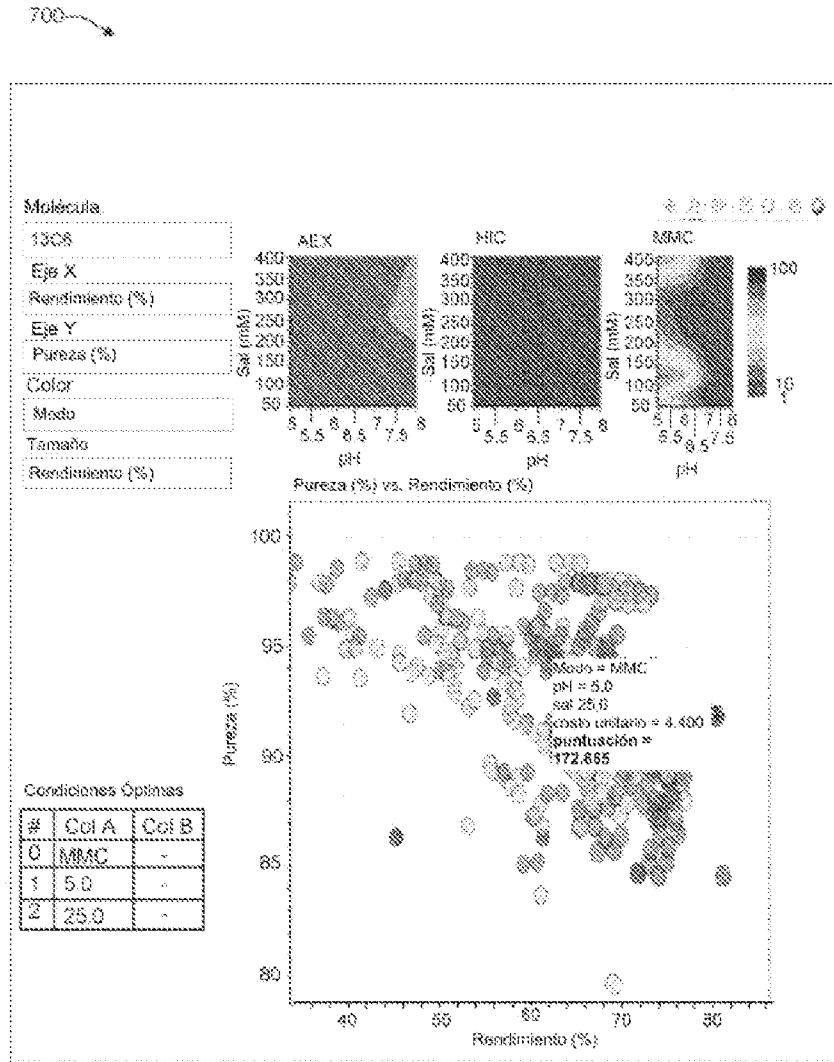


FIG. 7

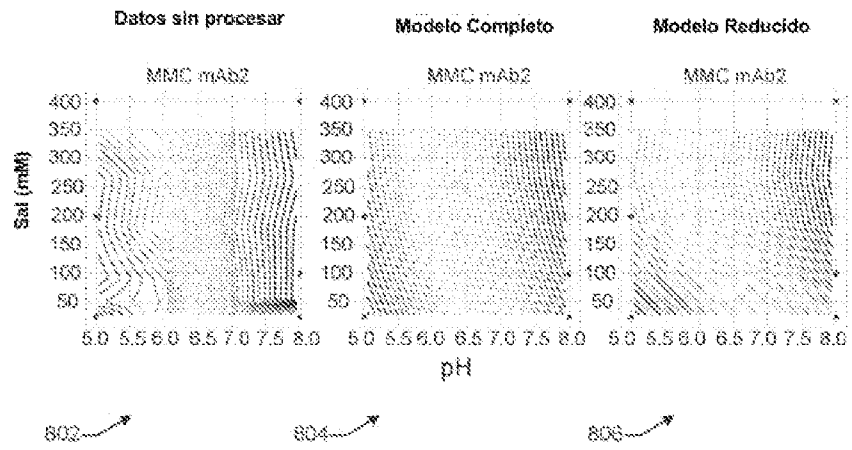


FIG. 8

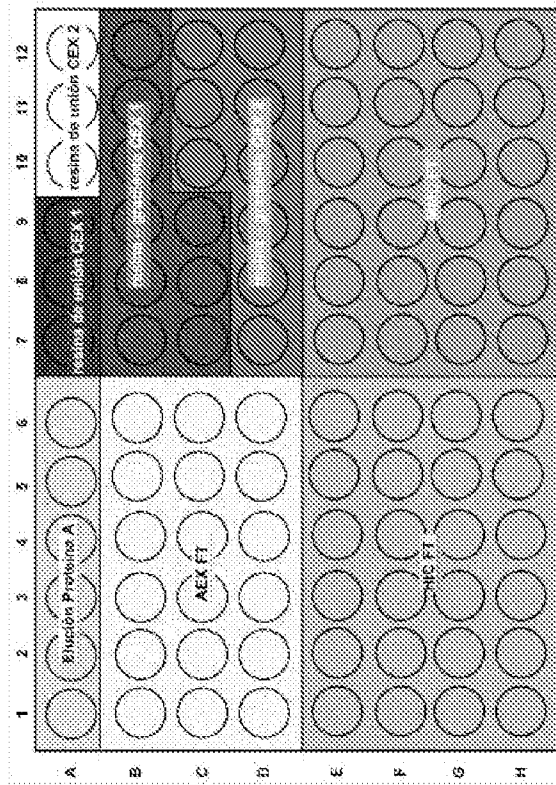


FIG. 9

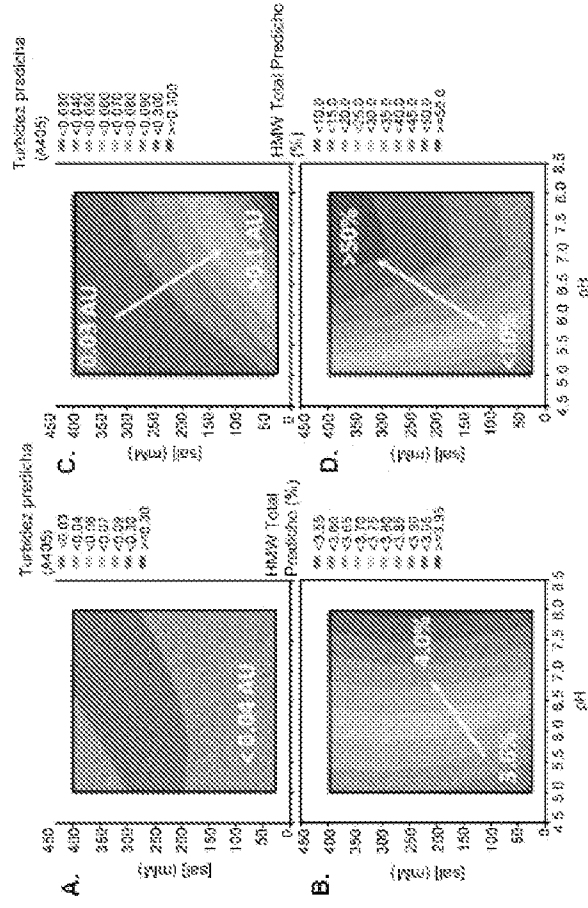


FIG. 10

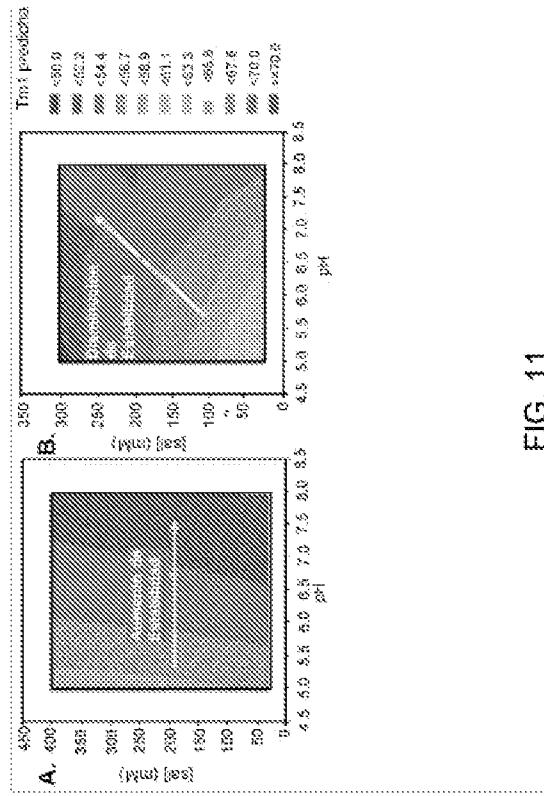


FIG. 11

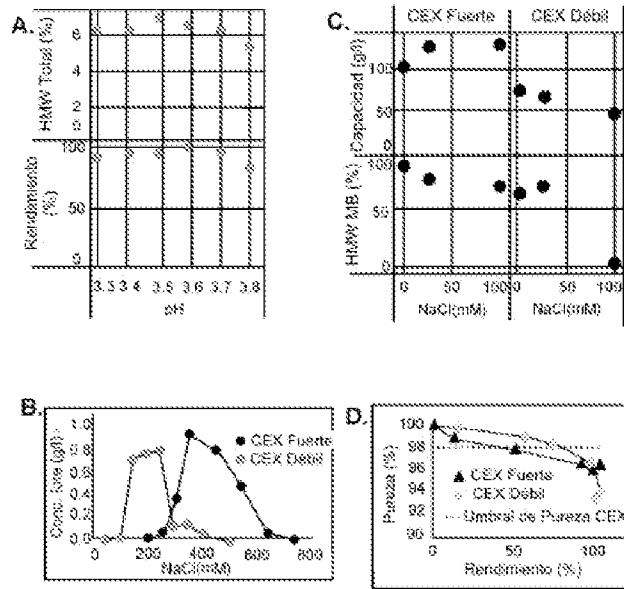


FIG. 12

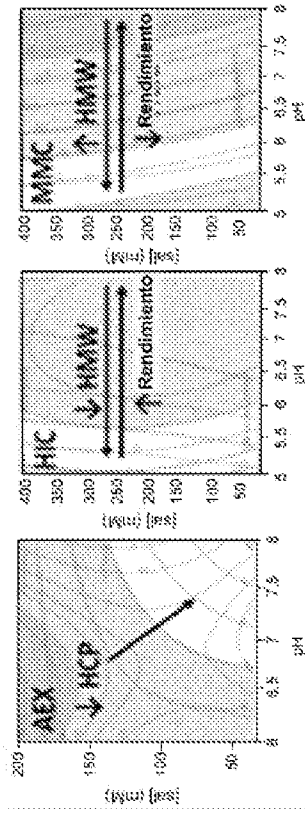


FIG. 13

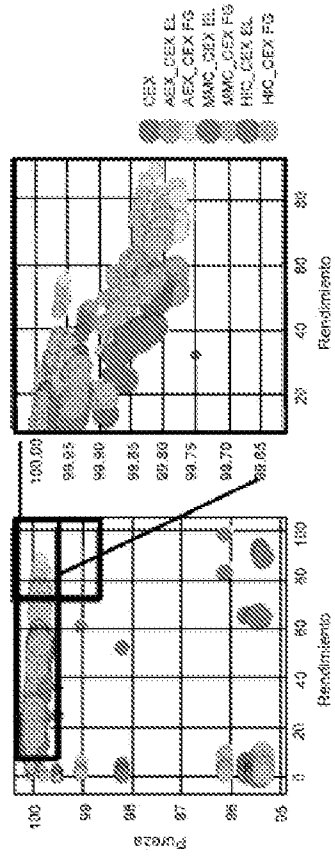


FIG. 14