

DOMINIOS DE UNIÓN DEL RECEPTOR DE LA GLUCOPROTEÍNA SPIKE DEL CORONAVIRUS Y USOS DE ESTOS

REFERENCIA CRUZADA CON SOLICITUDES RELACIONADAS

[0001] La presente solicitud reivindica el beneficio de la solicitud provisional de patente de EE. UU. núm. 63/465.978 presentada el 12 de mayo de 2023 y la solicitud provisional de patente de EE. UU. núm. 63/598.691 presentada el 14 de noviembre de 2023, cada una de las cuales se incorpora a la presente mediante esta referencia en su totalidad.

LISTADO DE SECUENCIAS

[0002] Esta solicitud contiene un listado de secuencias que se presentó por vía EFS-WEB en formato .XML y se incorpora en su totalidad mediante esta referencia. Dicha copia .XML, creada el 8 de mayo de 2023, se denomina 061291-509001WO_SeqList_ST26.xml y tiene un tamaño de 335 kilobytes.

ANTECEDENTES

[0003] La aparición del SARS-Cov-2, el virus que causa la COVID-19, ha aumentado el interés en las vacunas contra los coronavirus, tanto para el SARS-CoV-2 como para otros coronavirus. Existe una necesidad insatisfecha de polipéptidos de coronavirus genomodificados que tengan propiedades mejoradas y usos mejorados.

COMPENDIO

[0004] En un aspecto, la descripción proporciona un polipéptido recombinante, que comprende un fragmento antigénico de una glucoproteína Spike de coronavirus que comprende un dominio de unión al receptor (RBD) y, opcionalmente, un subdominio 1 de coronavirus (SD1). En algunas modalidades, el fragmento antigénico de los polipéptidos comprende o consiste esencialmente en RDB-SD1.

[0005] En algunas modalidades, el SD1 comprende al menos una sustitución de un residuo de aminoácido no polar expuesto en la superficie con un residuo de aminoácido polar. En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, o tres aminoácidos en las posiciones 560, 562 o 569 con un residuo polar, en donde las sustituciones son relativas a una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende una sustitución de una o más, dos o más, o tres de las sustituciones de aminoácidos L560Q, F562T, F562Y o I569S en relación con una secuencia de referencia según la SEQ ID NO: 1.

[0006] En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, o tres o más aminoácidos en las posiciones 544, 546, 560, 562, 564, 569 o 582 en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende una sustitución de una o más, dos o más, tres o más, o cuatro o más aminoácidos

5 N544L, N544M, N544Q, L546V, L560Q, F562T, F562Y, Q564C, Q564L, Q564N, Q564W, I569S o L582S en relación con una secuencia de referencia según la SEQ ID NO: 1.

[0007] En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones 338, 358, 363, 365, 392 o 395 en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el

10 polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones F338L, I358F, A363L, Y365F, Y365M, Y365W, F392W o V395I en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos que comprenden F338L/Y365W, F392W, Y365F/V395I, Y365F/F392W/V395I, F338L/A363L/Y365M, y/o I358F/Y365F/V395I.

15 **[0008]** En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones 329, 348, 350, 367, 375, 402, 407, 410, 418, 429, 433, 435, 452, 464, 510, 512, 514, 517, 518, 519, 520, 522, 527 o 528 en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más

20 aminoácidos en las posiciones F329K, F329R, F329Y, A348P, V350L, V367F, F375Y, I402V, V407L, I410F, I418V, F429W, V433I, A435I, A435V, L452R, F464Y, P512Q, V510I, V512F, V512I, S514T, L517A, L517D, L517S, L517T, L518N, L518Q, L518V, H519D, H519G, H519R, H519S, H519T, A520C, A520D, A520G, A520H, P521A, P521D, P521N, P521Q, P521S, A522G, A522I, P527N, K528I, K528Q, K528T, L518G o L518S en relación con una secuencia

25 de referencia según la SEQ ID NO: 1.

[0009] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 348, 402, 464, 514, 520, 526 y 527, las sustituciones de aminoácidos en relación con una secuencia de referencia según SEQ ID NO: 1. En algunas

30 modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden A348P/Y365F/V395I/L560Q/F562Y/I569S, Y365F/V395I/I402V/L560Q/F562Y/I569S,

Y365F/V395I/S514T/L560Q/F562Y/I569S, Y365F/V395I/I402V/F464Y/L560Q/F562Y/I569S y/o Y365F/V395I/G526S/P527N/L560Q/F562Y/I569S.

[0010] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 329, 517, 519, 520 y 544, las sustituciones de aminoácidos en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden Y365F/V395I/L560Q/F562Y/I569S, Y365F/V395I/F329Y/L560Q/F562Y/I569S,

Y365F/V395I/L517T/H519S/L560Q/F562Y/I569S,

Y365F/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S,

Y365F/V395I/L517T/H519S/N544L/L560Q/F562Y/I569S, y/o

Y365F/V395I/F329Y/L517T/H519S/L560Q/F562Y/I569S.

[0011] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 392, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 329, 517, 519, 520 y 544, las sustituciones de aminoácidos en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden Y365F/F392W/V395I/L560Q/F562Y/I569S, Y365F/F392W/V395I/F329Y/L560Q/F562Y/I569S, Y365F/F392W/V395I/L517T/H519S/L560Q/F562Y/I569S,

Y365F/F392W/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S,

Y365F/F392W/V395I/L517T/H519S/N544L/L560Q/F562Y/I569S, y/o

Y365F/F392W/V395I/F329Y/L517T/H519S/L560Q/F562Y/I569S.

[0012] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519, 520, 560, 562, 564 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 521, 544 y 546, las sustituciones de aminoácidos en relación con una secuencia de referencia de acuerdo con SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden Y365F/V395I/L517T/H519G/A520C/Q564C/L560Q/F562Y/I569S, y/o

Y365F/V395I/L517S/H519R/A520H/P521Q/N544L/L546V/Q564W/F562T/L560Q/F562Y/I569S

[0013] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519 y 520; y una o más, dos o más, o tres o más sustituciones de

aminoácidos en las posiciones 518 y 392, las sustituciones de aminoácidos en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden Y365F/V395I/L517T/H519S/A520G, Y365F/V395I/L517T/L518G/H519D/A520G, Y365F/V395I/L517T/L518S/H519S/A520G, Y365F/F392W/V395I/L517T/H519S/A520G, Y365F/F392W/V395I/L517T/L518G/H519D/A520G, y/o Y365F/F392W/V395I/L517T/L518S/H519S/A520G.

[0014] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 329, 365, 395, 517, 519, 520, 560, 562, y 56, y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 392, 544 y 564, las sustituciones de aminoácidos en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden Y365F/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S/F329Y, Y365F/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S/F329Y, Y365F/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S/F329Y, Y365F/F392W/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S/F329Y, Y365F/F392W/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S/F329Y y/o Y365F/F392W/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S/F329Y. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden Y365F/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S, Y365F/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S, Y365F/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S, Y365F/F392W/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S, Y365F/F392W/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S, y/o Y365F/F392W/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S.

[0015] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519, 520, 521, 504, 560, 562, 569 y 329; y una o más sustituciones de aminoácidos en las posiciones 392, las sustituciones de aminoácidos en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden Y365F/V395I/L517T/H519G/P521A/A520C/Q564C/L560Q/F562Y/I569S/F329Y,

Y365F/V395I/L517T/H519G/P521N/A520C/Q564C/L560Q/F562Y/I569S/F329Y,
 Y365F/F392W/V395I/L517T/H519G/P521A/A520C/Q564C/L560Q/F562Y/I569S/F329Y, y/o
 Y365F/F392W/V395I/L517T/H519G/P521N/A520C/Q564C/L560Q/F562Y/I569S/F329Y.

[0016] En algunas modalidades, el polipéptido comprende una pluralidad de sustituciones de aminoácidos en las posiciones 329, 365, 395, 526, 527, 528, 560, 562 y 569. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que comprenden F329Y/Y365F/V395I/G526S/P527N/K528T/L560Q/F562Y/I569S, y/o F329Y/Y365F/V395I/G526S/P527N/K528I/L560Q/F562Y/I569S.

[0017] En algunas modalidades, el polipéptido comprende una secuencia polipeptídica al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a una o más de las SEQ ID NO: 11-19. En algunas modalidades, el polipéptido comprende una secuencia polipeptídica idéntica a una cualquiera de las SEQ ID NO: 11-19.

[0018] En algunas modalidades, el polipéptido se expresa a un nivel más alto que un polipéptido de referencia sin las sustituciones de aminoácidos, opcionalmente un polipéptido de referencia según las SEQ ID NO: 2 o 10. En algunas modalidades, el polipéptido expresado es más estable térmicamente o más antigénico que las SEQ ID NO: 2 o 10.

[0019] En algunas modalidades, el polipéptido comprende una secuencia polipeptídica al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 2-10.

[0020] En otro aspecto, la descripción proporciona una nanoestructura de proteína autoensamblable que comprende un primer componente que comprende un polipéptido de coronavirus recombinante descrito en la presente y, opcionalmente, un segundo componente que comprende una segunda proteína.

[0021] En algunas modalidades, la nanoestructura de proteína del polipéptido de coronavirus recombinante comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a una o más de las SEQ ID NO: 11-19.

al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

[0022] En algunas modalidades, la nanoestructura proteica del polipéptido de coronavirus recombinante comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en relación con una secuencia de referencia según la SEQ ID NO: 1.

[0023] En algunas modalidades, el componente polipeptídico de nanoestructura de proteína y un polipéptido de coronavirus recombinante están acoplados de forma no covalente. En algunas modalidades, el componente polipeptídico de nanoestructura de proteína y el polipéptido de coronavirus recombinante están acoplados de forma covalente.

[0024] En algunas modalidades, la nanoestructura de proteína comprende una proteína de fusión que incluye el polipéptido de coronavirus recombinante, un enlazador y el polipéptido componente de la nanoestructura de proteína. En algunas modalidades, la nanoestructura de proteína comprende una proteína de fusión que incluye, en el orden de extremo N a C, el polipéptido de coronavirus recombinante, un enlazador y el polipéptido componente de la nanoestructura de proteína.

[0025] En algunas modalidades, la nanoestructura de proteína del polipéptido de coronavirus recombinante comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

[0026] En algunas modalidades, la nanoestructura de proteína de la proteína de fusión comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

[0027] En algunas modalidades, el primer componente de la nanoestructura de proteína comprende I53-50A.

[0028] En algunas modalidades, la nanoestructura de proteína es una nanoestructura de

proteína I3-01/MI3. En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína I53-50AB.

[0029] En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de ferritina. En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de encapsulina. En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de cápside de bacteriófago CP3. En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de cápside de bacteriófago Q β . En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de cápside de bacteriófago AP205.

[0030] En algunas modalidades, el polipéptido recombinante de la nanoestructura de proteína comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

[0031] En algunas modalidades, el componente polipeptídico de la nanoestructura de proteína comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 20-21.

[0032] En algunas modalidades, la nanoestructura de proteína comprende un segundo componente polipeptídico.

[0033] En algunas modalidades, el segundo polipéptido componente es I53-50B.

[0034] En algunas modalidades, el segundo componente polipeptídico comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 22.

[0035] En algunas modalidades, el enlazador de la nanoestructura de proteína comprende entre

8 y 24 residuos de aminoácidos. En algunas modalidades, el enlazador de la nanoestructura de proteína se selecciona del grupo que consiste en GGSGGSGSGGSGGSGS, SGGGSGGSGSGGSGGSGS, EPEGGSGGSGSGGSGGSGS, YGGSGGSGGSGSGGSGGSGS, y GSGGSGGSGSGGSGGSGSGGSGGSGS.

5 **[0036]** En otro aspecto, la descripción proporciona un polinucleótido que codifica el polipéptido o la nanoestructura de proteína descrita en la presente. En algunas modalidades, el polinucleótido es ARNm.

[0037] En otro aspecto, la descripción proporciona una composición farmacéutica que comprende los polipéptidos, las nanoestructuras de proteína, o los polinucleótidos descritos
10 anteriormente.

[0038] En otro aspecto, la descripción proporciona vacunas que comprenden los polipéptidos, las nanoestructuras de proteína, o los polinucleótidos descritos anteriormente. En algunas modalidades, la vacuna comprende además un adyuvante. En algunas modalidades, el adyuvante comprende escualeno. En algunas modalidades, la vacuna es una vacuna bivalente.
15 En algunas modalidades, la vacuna bivalente comprende el polipéptido o la nanoestructura proteica descrita en la presente.

[0039] En otro aspecto, la descripción proporciona un método para prevenir o tratar una enfermedad relacionada con el coronavirus en un sujeto que lo necesita, que comprende administrar al sujeto los polipéptidos, las nanoestructuras de proteína, los polinucleótidos, el
20 componente farmacéutico o la vacuna descritos anteriormente. En otro aspecto, la descripción proporciona un método para inmunizar a un sujeto que lo necesita contra una infección por coronavirus, administrándole al sujeto la vacuna descrita anteriormente. En otro aspecto, la descripción proporciona kits que pueden usarse para preparar y administrar polipéptidos, nanoestructuras de proteínas, polinucleótidos, composiciones farmacéuticas o vacunas como
25 se describió anteriormente.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0040] El expediente de patente o solicitud contiene al menos un dibujo realizado en color. La Oficina proporcionará copias de esta patente o publicación de solicitud de patente con dibujos en color, previa solicitud y pago de la tarifa necesaria.

30 **[0041]** Las Figuras 1A-1C muestran la estructura de los fragmentos de antígeno RBD y RBD-SD1 (PDB 7UB5). La Figura 1A muestra la estructura de RBD solo, desde los residuos 328-

519. El motivo de unión al receptor (RBM) que se une a ACE2 se muestra en un recuadro. **La Figura 1B** muestra la estructura de RBD-SD1 de los residuos 319-591. Los residuos de SD1 están coloreados en negro. **La Figura 1C** muestra la misma estructura que **la Figura 1B**, con las regiones objetivo para las mutaciones diseñadas de acuerdo con la presente descripción
 5 sombreadas, con las regiones que comprenden mutaciones de Rpk, mutaciones de la interfaz RBD-SD1 y mutaciones proximales de RBM sombreadas y con mapas de densidad. Las mutaciones SD1resurf se muestran como esferas sombreadas y rodeadas por un cuadro.

[0042] Las Figuras 2A-2C muestran análisis de sobrenadantes de diseños iniciales probados con el fragmento ancestral RBD-SD1 fusionado genéticamente a CompA. **La Figura 2A**
 10 muestra transferencia Western de sobrenadantes. **La Figura 2B** muestra un primer conjunto de datos de interferometría de biocapa (BLI) recopilados para la unión de ACE2 contra muestras de sobrenadante seleccionadas. Las diferencias de enlace (en las construcciones RDBb1-5) se compararon usando antígenos WT RBD-SD1 (borde superior izquierdo) y las diferencias de mutación estabilizadora se compararon dentro de las construcciones con enlazadores de 16
 15 residuos. **La Figura 2C** muestra un segundo conjunto de datos de BLI recopilados para la unión de ACE2 contra muestras de sobrenadante seleccionadas. Los datos se centraron en diseños que contienen mutaciones Rpk8 (RBDb016 a 020), Rpk9 (RBDb021 a 025) o SD1resurf (RBDb036 a 040). RBD01 se usó como control para todos los estudios. Los valores de desplazamiento máximo para **las Figuras 2B-2C** se muestran respectivamente en **la Tabla 16**
 20 **y la Tabla 17**.

[0043] Las Figuras 3A-3F muestran trazas de BLI para CompAs purificadas de antígenos ancestrales. Se usó RBD01 como control, que carece de mutaciones estabilizadoras y del dominio SD1. **La Figura 3A** muestra la unión de ACE2, lo que representa la integridad de los epítomos clave que respaldan los anticuerpos neutralizantes. Los niveles de unión fueron
 25 similares para todas las construcciones. **La Figura 3B** muestra la unión de ACE2, lo que representa la integridad de los epítomos clave que respaldan los anticuerpos neutralizantes. Los niveles de unión fueron similares para todas las construcciones. **La Figura 3C y la Figura 3D** muestran la unión de CR3022. CR3022 es un anticuerpo no neutralizante que se dirige a un epítipo distal al sitio de unión de ACE2 que normalmente no está expuesto en la glucoproteína Spike. **La Figura 3E y la Figura 3F** muestran la unión de S309. S309 es un anticuerpo
 30 ampliamente neutralizante dentro de los sarbecovirus que se une fuera del sitio de unión de

ACE2. Los valores de desplazamiento máximo para **las Figuras 3A-F** se muestran en **la Tabla 18**.

[0044] **Las Figuras 4A-4K** muestran curvas de desnaturalización térmica para CompA ancestrales mediante fluorescencia intrínseca de proteínas (ITF) o tinción con gel de proteína naranja SYPRO™. **La Figura 4A y Figura 4B**, muestran, respectivamente, mediciones de ITF para el control RBD01 y RBDb001 en relación con RBDb002 para medir el impacto de la adición del dominio SD1 o la modificación de los enlazadores. Todos mostraron temperaturas de fusión similares. **La Figura 4C a la Figura 4F** muestran mediciones de ITF para CompAs RBD-SD1 que contienen diferentes mutaciones Rpk en comparación con WT RBD-SD1 (RBDb002). La adición de mutaciones Rpk9 condujo a las temperaturas de fusión más altas. **La Figura 4G** comparación ITF de RBDb002 y RBDb037 (mutaciones SD1resurf), que muestra que la adición de mutaciones SD1resurf no afecta la temperatura de fusión. Todas las gráficas en **la Figura 4A a la Figura 4G** muestran tres mediciones replicadas (cada una etiquetada como “rep1”, “rep2” o “rep3”), y las mediciones de ITF de RBDb002 que se muestran en cada gráfica para **la Figura 4A a la Figura 4G** son idénticas para permitir la comparación entre gráficas. Las temperaturas de fusión promedio para las mediciones de ITF de RBDb002 y muestras comparadoras se muestran respectivamente en líneas de puntos grises o líneas discontinuas negras, y también se muestran en **la Tabla 19**. **La Figura 4H a la Figura 4K** muestran comparaciones de SYPRO de CompAs RBD-SD1 que contienen diferentes mutaciones de Rpk en comparación con WT RBD-SD1 (RBDb002). Las mutaciones Rpk9 (RBDb022) mostraron una combinación de una línea base SYPRO más baja (lo que sugiere un plegamiento mejorado) y una temperatura de fusión más alta. Todas las gráficas en **la Figura 4H a la Figura 4K** muestran tres mediciones replicadas (cada una etiquetada como “rep1”, “rep2” o “rep3”), y las mediciones de ITF de RBDb002 que se muestran en cada gráfica para **la Figura 4H a la Figura 4K** son idénticas para permitir la comparación entre gráficas. Las temperaturas de fusión promedio para las mediciones de SYPRO de RBDb002 y muestras comparadoras se muestran respectivamente en líneas de puntos grises o líneas discontinuas negras, y también se muestran en **la Tabla 20**.

[0045] **La Figura 5A y la Figura 5B** muestran el ensamblaje de CompAs ancestrales con la molécula CompB complementaria, medido mediante dispersión de luz dinámica (DLS) (todas se muestran como distribuciones de intensidad). **La Figura 5A** muestra mediciones de DLS

recolectadas después del ensamblaje en presencia de un amortiguador complementado con 50 mM de arginina y 4 % de sacarosa, que muestra un ensamblaje altamente eficiente de RBD_b022 y RBD_b037 basado en el rango de tamaño previsto de nanoestructuras de proteínas ensambladas, mientras que todos los demás grupos mostraron agregación. **La Figura 5B**

muestra mediciones de DLS recolectadas después del ensamblaje en presencia de un amortiguador complementado con 50 mM de arginina, que muestra un ensamblaje altamente eficiente de RBD_b022 y RBD_b037 basado en el rango de tamaño previsto de nanoestructuras de proteínas ensambladas, mientras que todos los demás grupos mostraron agregación. Las etiquetas para “MOPS” y “Tris/glicerol” indican fondos de amortiguador más allá de los excipientes agregados, que se refieren respectivamente a 50 mM de MOPS, 0,75 % de CHAPS, 150 mM de NaCl, pH 7,4, o 20 mM de Tris, 5 % de glicerol, 250 mM de NaCl, pH 8. Los resultados de DLS de **la Figura 5A** y **la Figura 5B** también se resumen en **la Tabla 21**, junto con condiciones de amortiguación más detalladas para el ensamblaje y análisis de la nanoestructura.

[0046] La Figura 6 muestra un análisis de ensamblaje para RBD_b092 (ancestral) y RBD_b093 (BA.5), que contienen la combinación de mutaciones Rpk9 y SD1resurf. Se usó DLS para el análisis (mostrado como una distribución de intensidad), y el ensamblaje se realizó en presencia de un amortiguador complementado con 100 mM de arginina y 5 % de glicerol. RBD_b092 se ensambló con éxito con alta eficiencia, mientras que RBD_b093 mostró agregación. Los resultados de DLS de **la Figura 6** también se resumen en **la Tabla 22**.

[0047] La Figura 7 muestra una alineación de secuencias RBD-SD1 (equivalentes a los residuos 319-591 del SARS-CoV-2) de diversos sarbecovirus representativos. Los dominios SD1 están marcados con líneas horizontales y los residuos SD1 están en cursiva. Las regiones que comprenden mutaciones Rpk8/Rpk9, las mutaciones de la interfaz RBD-SD1 o las mutaciones SD1resurf están marcadas con símbolos de acuerdo con la leyenda. Los espacios con respecto a otras secuencias están marcados con guiones. Esta alineación muestra que las regiones que comprenden las mutaciones Rpk8/Rpk9 y las mutaciones de la interfaz RBD-SD1 están en gran medida conservadas, lo que verifica que estas mutaciones se extienden a través de diversas cepas virales. Además, esta alineación muestra que las mutaciones L560Q e I569S de SD1resurf están presentes en otros sarbecovirus y, por lo tanto, son aplicables a diversas cepas virales. I569V también aparece con frecuencia en la alineación.

[0048] La **Figura 8** muestra un análisis de transferencia Western de diseños BA.5 RBD-SD1 CompA en sobrenadantes después de la expresión a pequeña escala de células HEK293. Todos los números de construcción de dos dígitos hacen referencia a los dos últimos dígitos de la nomenclatura “RBDb0XX”, en el que los dos dígitos reemplazan a “XX”. Con la excepción de RBDb046, todos los diseños contienen mutaciones Rpk8 y SD1resurf, y las construcciones RBDb051-091 también contienen otras mutaciones. Las construcciones RBDb044-050 prueban la combinación de mutaciones Rpk8 y SD1resurf en múltiples cepas de ómicron, con RBDb047 usando la cepa BA.5.

[0049] La **Figura 9A** a **Figura 9C** muestran la caracterización *in vitro* de diseños purificados seleccionados para BA.5 RBD-SD1 CompAs. RBDb046 usa un antígeno WT RBD-SD1, mientras que todas las demás construcciones incluyen además mutaciones estabilizadoras. La **Figura 9A** muestra las temperaturas de fusión (T_m) y de inicio (T_{onset}) medidas usando ITF. La **Figura 9B** muestra mediciones de la tasa de unión de BLI contra ACE2. Las barras de error muestran la desviación estándar entre replicados. La **Figura 9C** muestra los desplazamientos máximos de BLI frente a ACE2. Las barras de error muestran la desviación estándar entre replicados.

[0050] La **Figura 10** muestra un ensamblaje de polipéptidos purificados seleccionados para BA.5 RBD-SD1 CompAs medido por DLS (mostrado como distribución de intensidad). El ensamblaje se realizó mediante amortiguador complementado con 100 mM de arginina. Varios polipéptidos se ensamblaron bien sin agregación, las excepciones fueron RBDb045, RBDb078 y RBDb081.

[0051] La **Figura 11A** y la **Figura 11B** muestran la inmunogenicidad de CompAs seleccionados y nanoestructuras de proteína de antígenos ancestrales y BA.5 en ratones. La **Figura 11A** muestra los títulos de unión de ELISA contra antígenos de glucoproteína ancestral o Spike BA.5 coincidentes con la cepa. La **Figura 11B** muestra la neutralización de pseudovirus frente a pseudovirus ancestrales o BA.5 de cepa coincidente. Estos datos muestran diferencias en la inmunogenicidad entre muchos CompA, particularmente cuando se compara RBDb046 con otros CompA BA.5 que contienen mutaciones estabilizadoras.

[0052] La **Figura 12A** muestra la relación de unión de ACE2/anti-CompA de diseños silvestres y estabilizados seleccionados en diversas cepas. La **Figura 12B** muestra la producción total de proteína de diseños silvestres y estabilizados seleccionados en diversas cepas. La **Figura 12C**

muestra la temperatura de fusión de diseños silvestres y estabilizados seleccionados en diversas cepas.

DESCRIPCIÓN DETALLADA

Definiciones

5 **[0053]** En la presente se hace referencia a características particulares (que incluyen etapas del método). Se debe entender que la descripción en esta memoria descriptiva incluye todas las publicaciones posibles de dichas características particulares. Por ejemplo, cuando se describe una característica particular en el contexto de un aspecto o modalidad particular, o una reivindicación particular, la característica también se puede usar, en la medida de lo posible, en combinación y/o en el contexto de otros aspectos y modalidades particulares

10 **[0054]** Donde se hace referencia en la presente a un método que comprende dos o más etapas definidas, las etapas definidas se pueden llevar a cabo en cualquier orden o de forma simultánea (excepto cuando el contexto excluye esa posibilidad) y el método puede incluir una o más etapas llevadas a cabo antes de cualquiera de las etapas definidas, entre dos de las etapas definidas o después de todas las etapas definidas (excepto cuando el contexto excluye esa posibilidad).

15 **[0055]** La práctica de la tecnología empleará, a menos que se indique específicamente lo contrario, métodos químicos, bioquímicos, de química orgánica, biología molecular, microbiología, genética, inmunobiología, biología celular y técnicas de ADN recombinante comprendidas en la técnica, muchas de las cuales se describen más adelante con fines
20 ilustrativos. Dichas técnicas se explican en su totalidad en la bibliografía. Se debe entender que esta descripción no se limita a la metodología, protocolos y reactivos particulares descritos, ya que pueden variar de acuerdo con el contexto en el que los expertos en la técnica los usan.

[0056] Todas las publicaciones y patentes que se mencionan en la presente se incorporan en su totalidad mediante esta referencia como si cada publicación o patente se indicara específica e individualmente como incorporada mediante esta referencia. En caso de conflicto, rige la presente solicitud, lo que incluye cualesquiera definiciones en ella. Sin embargo, la mención de cualquier referencia, artículo, publicación, patente, publicación de patente y solicitud de patente que se cite en la presente no es un reconocimiento de que constituyan técnica previa o formen parte del conocimiento general en cualquier país, y no se debe tomar como un reconocimiento
25 o sugerencia.

30 **[0057]** Los títulos de las secciones que se usan en la presente tienen únicamente fines de

organización y no se ha de interpretar que limiten la materia descrita. Específicamente, las características descritas en una sección se pueden combinar con características en cualquier otra sección de la descripción.

[0058] Aunque se describen y se ilustran ejemplos de modalidades, se entenderá que es posible hacer varios cambios a estos ejemplos de modalidades sin apartarse del espíritu y el alcance de la invención.

[0059] Salvo que se indique lo contrario en la presente, todos los términos técnicos y científicos que se utilizan en este documento tienen el mismo significado comúnmente entendido por un experto en la técnica a la que pertenece la presente invención. Diversos diccionarios científicos que incluyen los términos incluidos en la presente son bien conocidos y están a disposición de los expertos en la materia. Aunque cualquier método y material similar o equivalente a los descritos en la presente puede utilizarse en la práctica o en las pruebas de la descripción, se describen algunos métodos y materiales preferidos. Por consiguiente, los términos que se definen inmediatamente a continuación se describen de manera más completa en relación con la memoria descriptiva en su totalidad.

[0060] Se pretende que las formas singulares «un», «una», «el» y «la» incluyan también las formas plurales, a menos que el contexto indique claramente lo contrario.

[0061] Tal como se usa en la presente, el término «alrededor de» significa un intervalo de valores que incluyen el valor indicado, que un experto en la técnica podría considerar razonablemente similares al valor indicado. En algunas modalidades, aproximadamente significa dentro de una desviación estándar usando mediciones generalmente aceptables en la técnica. En algunas modalidades, aproximadamente significa un rango que se extiende hasta $\pm 10\%$, $\pm 5\%$, $\pm 3\%$ o $\pm 1\%$ del valor especificado.

[0062] El término «al menos» seguido de un número se utiliza en la presente para indicar el inicio de un intervalo que comienza con ese número (que puede ser un intervalo que tiene un límite superior o ningún límite superior, dependiendo de la variable que se defina). Por ejemplo, "al menos 1" se refiere a 1 o más de 1.

[0063] El término "como máximo", seguido de un número, se utiliza en la presente para indicar el final de un intervalo que termina con ese número (que puede ser un intervalo que tiene 1 o 0 como su límite inferior, o un intervalo que no tiene límite inferior, dependiendo de la variable definida). Por ejemplo, "como máximo 4" significa 4 o menos de 4, y "como máximo 40 %"

significa 40 % o menos de 40 %. Cuando, en esta memoria descriptiva, un intervalo se indica como "de (un primer número) a (un segundo número)" o "(un primer número)-(un segundo número)", se trata de un intervalo cuyo límite inferior es el primer número y cuyo límite superior es el segundo número. Por ejemplo, 25 a 100 mm significa un intervalo cuyo límite inferior es 25 mm y cuyo límite superior es 100 mm.

[0064] Tal como se utiliza en este documento, "sustancialmente" o "sustancial" se refiere al alcance o grado completo o casi completo de una acción, característica, propiedad, estado, estructura, elemento o resultado. Por ejemplo, un objeto que está "sustancialmente" encerrado significaría que el objeto está completamente encerrado o casi completamente encerrado. El grado exacto de desviación permisible respecto de la integridad absoluta puede depender en algunos casos del contexto específico. Sin embargo, en términos generales, la proximidad de la finalización será tal que se obtendrá el mismo resultado general que si se obtuviera la finalización absoluta y total. El uso de "sustancialmente" es igualmente aplicable cuando se utiliza en una connotación negativa para referirse a la falta total o casi total de acción, característica, propiedad, estado, estructura, elemento o resultado. Por ejemplo, una composición que está "sustancialmente libre de" otros agentes activos carecería completamente de otros agentes activos o carecería casi por completo de otros agentes activos de tal manera que el efecto sería el mismo que si careciera completamente de otros agentes activos. En otras palabras, una composición que está «sustancialmente libre de» un ingrediente, elemento u otro agente activo todavía puede contener dicho elemento siempre que no tenga un efecto mensurable.

[0065] En esta memoria descriptiva, a menos que el contexto requiera lo contrario, se comprenderá que «comprender», «comprende» y «que comprende» implican la inclusión de una etapa, un elemento o un grupo de etapas o elementos establecidos, pero no la exclusión de cualquier otra etapa o elemento o grupo de etapas o elementos. Por ejemplo, una composición «que comprende» los componentes A, B y C puede consistir en (es decir, contener solamente) los componentes A, B y C, o pueden contener no solo los componentes A, B y C, sino también uno o más de otros componentes.

[0066] Tal como se usa en la presente, la expresión «que consiste en» hace referencia a incluir, de modo taxativo, lo posterior a la frase «que consiste en». Por lo tanto, la frase «que consiste en» indica que los elementos enumerados son necesarios u obligatorios, y que no pueden estar

presentes otros elementos. Se pretende que «que consiste esencialmente en» incluya cualesquiera elementos enumerados después de la frase y se limite a otros elementos que no interfieran o contribuyan a la actividad o la acción especificada en la descripción para los elementos mencionados. Por consiguiente, la frase «que consiste esencialmente en» indica que los elementos mencionados son necesarios u obligatorios, pero ningún otro elemento es opcional y pueden o no estar presentes de acuerdo con que afecten o no la actividad o la acción de los elementos mencionados.

[0067] La referencia a lo largo de esta memoria descriptiva, por ejemplo, "una modalidad", "otra modalidad", "una modalidad particular", "una modalidad relacionada", "en ciertas modalidades", "en algunas modalidades", "una modalidad adicional" o "una modalidad adicional" o combinaciones de estas significa que una característica, estructura o aspecto particular descritos en relación con la modalidad está incluida en al menos una modalidad de la presente descripción. Por lo tanto, las apariciones de las frases anteriores en varios lugares a lo largo de esta memoria descriptiva no necesariamente se refieren todas a la misma modalidad. Además, las características, estructuras o aspectos particulares pueden combinarse de cualquier manera adecuada en una o más modalidades.

[0068] La siguiente descripción incluye información que puede ser útil para comprender la presente invención. No constituye una admisión de que alguna de la información aquí proporcionada sea técnica anterior o relevante para las invenciones actualmente reivindicadas, o que alguna publicación a la que se hace referencia específica o implícita sea técnica anterior.

[0069] El término "idéntico/a" o "porcentaje de identidad", en el contexto de dos o más secuencias de ácidos nucleicos o polipeptídicas, se refieren a dos o más secuencias o subsecuencias que son iguales o tiene un porcentaje especificado de nucleótidos o residuos aminoácidos que son iguales, cuando se comparan y alinean para una correspondencia máxima. Los métodos de alineamiento de secuencias para comparación son bien conocidos en la técnica. Una vez que estén alineadas, la cantidad de coincidencias se determina contando la cantidad de posiciones donde un residuo de aminoácido o nucleótido idéntico se presenta en ambas secuencias. El porcentaje de identidad de secuencia se determina dividiendo la cantidad de coincidencias en la alineación por la longitud de la secuencia de referencia, seguido de la multiplicación del valor resultante por 100. Por ejemplo, una secuencia peptídica que tiene 1166 coincidencias cuando se alinea con una secuencia de prueba que tiene 1554 aminoácidos es

un 75,0 por ciento idéntica a la secuencia de prueba ($1166 \div 1554 * 100 = 75,0$). Tal como se utilizan los términos en la presente, los huecos en la alineación no disminuyen el porcentaje de identidad de secuencia.

[0070] A menos que se especifique lo contrario, la alineación óptima de las secuencias para la comparación se realiza mediante el algoritmo de alineación global de Needleman y Wunsch, Mol. Biol. 48:443 (1970) implementado por EMBOSS Needle (disponible en línea en ebi.ac.uk/Tools/psa/emboss_needle/) (Madeira et al. *Nucleic Acids Res.* 50(W1):W276-W279 (2022)). En modalidades, pueden utilizarse otros métodos de alineación, incluidos, entre otros, los descritos en Devereux, et al, *Nucleic Acids Res.* 12:387-95 (1984); Atschul et al. *J. Mo. Biol.* 215:403-10 (1990) (BLAST); Carrillo y Lipman *Siam J. Appl. Math.* 48(5) (1988); Computational Molecular Biology (Lesk, AM, ed., 1989); Biocomputing Informatics and Genome Projects, (Smith, DW, ed., 1993); Computer Analysis of Sequence Data, Parte I, (Griffin y Griffin, eds., 1994); Sequence Analysis in Molecular Biology (von Heinje, 2012); Sequence Analysis Primer (Gribskov y Devereux, J., eds. 1993). En modalidades, la identidad de secuencia se calcula usando la implementación del algoritmo Needleman-Wunsch proporcionado por la Biblioteca Nacional de Medicina (disponible en línea en blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=GlobalAln).

[0071] Por ejemplo, la identidad de secuencia se puede determinar mediante métodos estándar que se utilizan comúnmente para comparar la similitud de dos secuencias de polipéptidos o de polinucleótidos. Al usar un programa informático tal como EMBOSS Needle o BLAST, dos polipéptidos o dos secuencias de polinucleótidos se alinean para lograr una coincidencia óptima de sus residuos respectivos (ya sea a través de la longitud completa de una o ambas secuencias o a través de una porción predeterminada de una o ambas secuencias). Los programas proporcionan una penalización de apertura predeterminada y una penalización de espacio predeterminada, y una matriz de puntuación como PAM 250 (una matriz de puntuación estándar; véase Dayhoff et al., en *Atlas of Protein Sequence and Structure*, vol. 5, sup. 3 (1978)) que se puede utilizar junto con el programa informático.

[0072] Una «posición» de aminoácido o base nucleotídica se denota mediante un número que identifica secuencialmente cada aminoácido (o base nucleotídica) en la secuencia de referencia basándose en su posición relativa al extremo N (o extremo 5'). Debido a las delecciones, inserciones, truncamientos, fusiones y similares que se deben tener en cuenta al determinar un

alineamiento óptimo, en general el número de residuo de aminoácido en una secuencia de prueba determinada simplemente contando desde el extremo N no será necesariamente el mismo que el número de su posición correspondiente en la secuencia de referencia. Por ejemplo, en un caso de que una variante tenga una delección relativa a una secuencia de referencia alineada, no habrá ningún aminoácido en la variante que corresponda a una posición en la secuencia de referencia en el sitio de la delección. En el caso de que haya una inserción en una secuencia de referencia alineada, dicha inserción no corresponderá a una posición de aminoácido numerada en la secuencia de referencia. En el caso de truncamientos o fusiones puede haber tramos de aminoácidos en la secuencia de referencia o alineada que no correspondan a ningún aminoácido en la secuencia correspondiente.

[0073] Los términos "en relación con" o "numerado con referencia a" o "correspondiente a", cuando se usan en el contexto de la numeración de una secuencia dada, se refieren a la numeración de los residuos de la secuencia dada cuando se alinea a una secuencia de referencia especificada, tal como cada posición en la secuencia numerada asocia un número correspondiente al número de la posición contraparte en la secuencia de referencia.

[0074] El término "coronavirus" se refiere a los virus de la subfamilia Coronavirinae de la familia Coronaviridae, del orden Nidovirales. Los coronavirus son virus con envoltura y tienen un genoma de ARN monocatenario de sentido positivo con una nucleocápside de simetría helicoidal. El término "coronavirus" abarca todas las cepas, genotipos, protectotipos y serotipos del virus de la bronquitis infecciosa. Ejemplos de coronavirus aviares son el virus de la bronquitis infecciosa (IBV), el coronavirus de la gallina de Guinea (GfCo) y el coronavirus del pavo (virus de la enteritis del pavo TCO y virus de la enfermedad de Bluecomb).

[0075] El término "SARS-CoV-2", también conocido como Coronavirus de Wuhan o nuevo Coronavirus de 2019, se refiere a un virus de ARN monocatenario de sentido positivo del género betaCoronavirus que ha surgido como una causa altamente mortal de infección respiratoria aguda grave. El término "SARS-CoV-2" incluye variantes del mismo, tales como, entre otras, alfa (linajes B.1.1.7 y Q); beta (B.1.351 y linajes descendientes); delta (linajes B.1.617.2 y AY); gamma (P.1 y linajes descendientes); épsilon (B.1.427 y B.1.429); eta (B.1.525); iota (B.1.526); kappa (B.1.617.1); 1.617.3; mu (B.1.621, B.1.621.1) y zeta (P.2)).

[0076] B.1.1.529, también conocida como variante ómicron, es una variante del SARS-CoV-2 original. Esta variante tiene un total de 60 mutaciones en comparación con la cepa original del

SARS-CoV-2, concretamente 50 mutaciones no sinónimas, 8 mutaciones sinónimas, 35 y 2 mutaciones no codificantes. Treinta y dos mutaciones afectan a la glucoproteína Spike (A67V, A69-70, T95I, G142D, A143-145, A211, L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K y L981F con referencia a SEQ ID NO: 1), o de las cuales aproximadamente la mitad se encuentran en el dominio de unión al receptor (319-530).

[0077] El término “glucoproteína Spike” se refiere a una glucoproteína de fusión de clase I sintetizada inicialmente como una proteína precursora de aproximadamente 1273 aminoácidos para el SARS-CoV-2. La subunidad S1 está distal a la membrana del virus y contiene el dominio de unión al receptor (RBD) que se cree que media la unión del virus a su receptor huésped. En la presente se presenta una secuencia ejemplar de glucoproteína Spike del SARS-CoV-2 tal como SEQ ID NO: 1 (secuencia de referencia de NCBI: YP_009724390.1 cuya secuencia se incorpora por referencia en la presente en su totalidad). Como se usa en la presente, el fragmento antigénico son los aminoácidos 319-591 en la secuencia de referencia NCBI: YP_009724390.1.

[0078] Como se usa en la presente, el término “dominio de unión al receptor (RBD)” se refiere a una parte de la glucoproteína Spike. Por ejemplo, el RBD corresponde a los aminoácidos 319-541 en la secuencia de referencia del NCBI: YP_009724390.1. Como se usa en la presente, el RBD corresponde a los aminoácidos 328-531 en la secuencia de referencia del NCBI: YP_009724390.1. El RBD incluye el motivo de unión al receptor (RBM), que facilita la entrada celular al unirse al receptor ECA2.

[0079] Como se usa en la presente, el término “subdominio-1 (SD1)” se refiere a una parte de la glucoproteína Spike. Por ejemplo, el SD1 son los aminoácidos 532-591 en la secuencia de referencia NCBI: YP_009724390.1.

[0080] El término “ancestral” se refiere a antígenos o virus basados directamente en la secuencia del aislado original Wuhan-Hu-1 del SARS-CoV-2 (identificador de Genbank NC_045512.2).

[0081] El término “recombinante” indica que el material (por ejemplo, un ácido nucleico o una proteína) ha sido alterado artificial o sintéticamente (no naturalmente) mediante intervención humana. La alteración puede realizarse sobre el material dentro o fuera de su entorno o estado

natural. En concreto, por ejemplo, un coronavirus es recombinante cuando se produce mediante la expresión de un ácido nucleico recombinante. Por ejemplo, un “ácido nucleico recombinante” es uno que se produce mediante la recombinación de ácidos nucleicos, por ejemplo, durante la clonación, la mezcla de ADN u otros procedimientos, o mediante mutagénesis química u otra mutagénesis; un “polipéptido recombinante” o una “proteína recombinante” es un polipéptido o proteína que se produce mediante la expresión de un ácido nucleico recombinante; y un “virus recombinante”, por ejemplo, un coronavirus recombinante, se produce mediante la expresión de un ácido nucleico recombinante. La “proteína de fusión” o los “tetrámeros” descritos en la presente son recombinantes y no existen en la naturaleza.

10 **[0082]** Los términos "polipéptido", "péptido" y "proteína" se usan indistintamente en la presente para referirse a un polímero de residuos de aminoácidos, en donde el polímero puede estar conjugado con un resto que no consiste en aminoácidos, tal como un glicano u otra modificación post-traducciona, o un resto químico no natural. Los términos se aplican a polímeros de aminoácidos en los que uno o más residuos de aminoácidos son miméticos químicos artificiales de un aminoácido natural correspondiente, así como a polímeros de aminoácidos naturales y polímeros de aminoácidos no naturales. Una "proteína de fusión" se refiere a una proteína quimérica que codifica dos o más secuencias de proteínas separadas que se expresan de forma recombinante como una única molécula de polímero.

15 **[0083]** El término “polipéptido recombinante” se refiere a un polipéptido que no se produce de forma natural o que tiene una secuencia formada por una combinación artificial de dos segmentos de secuencia que de otro modo estarían separados. En varias modalidades, un polipéptido recombinante está codificado por un ácido nucleico heterólogo (por ejemplo, recombinante) que se ha introducido en una célula huésped, tal como una célula bacteriana o eucariota.

20 **[0084]** Los términos «antígeno» o «inmunógeno» hacen referencia a su significado común y corriente de un compuesto o composición que induce una respuesta inmunitaria, celular o humoral, por ejemplo, una respuesta de linfocitos T citotóxicos (CTL), una respuesta de linfocitos B (por ejemplo, la producción de anticuerpos que se unen específicamente al epítopo), una respuesta de linfocitos NK o cualesquiera combinaciones de estos, cuando se administran a un sujeto inmunocompetente o se expresan en este. Los antígenos pueden incluir polipéptidos (lo que incluye glucoproteína). En algunos aspectos, un antígeno es un polipéptido o complejo

polipeptídico que incluye al menos un componente diseñado para provocar una respuesta inmunitaria. Por ejemplo, un antígeno puede incluir uno o más epítopos inmunogénicos asociados con un patógeno bacteriano. Tal como se usa en la presente, el término antígeno no se limita a la parte del polipéptido o complejo polipeptídico que contiene epítopos antigénicos.

- 5 Un «epítopo» o «determinante antigénico» hace referencia a su significado común y corriente como parte de un antígeno reconocido por el sistema inmunitario, específicamente, por anticuerpos, linfocitos B o linfocitos T.

[0085] El término "variante" hace referencia a un polipéptido que tiene una o más inserciones, deleciones o sustituciones de aminoácidos con respecto a un polipéptido de referencia, pero
10 que retiene una o más propiedades de la proteína de referencia.

[0086] El término "variante antigénica" se refiere a una variante que tiene uno o más epítopos en común con un polipéptido de referencia y/o genera la misma respuesta inmune o una similar cuando se administra a un sujeto como polipéptido de referencia.

[0087] El término "fragmento antigénico" se refiere a cualquier fragmento de una proteína que
15 genera una respuesta inmune (respuesta humoral o de células T) a la proteína in vivo. El fragmento antigénico puede ser un epítipo lineal, un epítipo discontinuo o un epítipo de conformación (por ejemplo, un dominio plegado). El fragmento antigénico puede preservar la estructura secundaria, terciaria y/o cuaternaria de la proteína de longitud completa.

[0088] La frase «variante funcional» hace referencia a una variante que exhibe al menos parte
20 de la actividad como un polipéptido de referencia. Por ejemplo, una variante funcional de un dominio de ensamblaje puede promover la multimerización y el ensamblaje en la misma medida, o en una medida similar, como un dominio de ensamblaje de referencia y/o se puede multimerizar y ensamblar con los mismos dominios de ensamblaje afines como un dominio de ensamblaje de referencia.

[0089] La expresión "sustitución de aminoácidos" hace referencia al reemplazo de un único
25 aminoácido en una secuencia con otro residuo de aminoácido. Se utilizan las formas estándar de abreviaturas para la sustitución de aminoácidos. Por ejemplo, V94R se refiere a la sustitución de valina (V) en una secuencia de referencia con arginina (R). La abreviatura Arg94 se refiere a cualquier secuencia en la que el residuo 94, con respecto a una secuencia de referencia, es
30 arginina (Arg). El término "combinación de sustituciones" se refiere a una pluralidad de sustituciones de aminoácidos que están incluidas en un solo polipéptido. Como se usa en la

presente, un carácter de barra “/” identifica una combinación de sustituciones en cada una de las posiciones de aminoácidos separadas por la barra. A modo de ejemplo, “Y365F/V395I” se refiere a una sustitución Y-->F en la posición de aminoácido 365 en combinación con una sustitución V-->I en la posición de aminoácido 395; “Y365F/ F392W /V395I” se refiere a una sustitución Y-->F en la posición de aminoácido 365 en combinación con una sustitución F-->W en la posición de aminoácido 392 y en combinación con una sustitución V-->I en la posición de aminoácido 395.

[0090] Cada uno de los siguientes ocho grupos contienen aminoácidos que son sustituciones conservadoras entre sí:

- 1) Alanina (A), Glicina (G);
- 2) Ácido aspártico (D), Ácido glutámico (E);
- 3) Asparagina (N), Glutamina (Q);
- 4) Arginina (R), Lisina (K);
- 5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V);
- 6) Fenilalanina (F), Tirosina (Y), Triptófano (W);
- 7) Serina (S), Treonina (T); y
- 8) Cisteína (C), Metionina (M)

(véase, por ejemplo, Creighton, Proteins (1984)).

[0091] Como se usa en la presente, “un residuo de aminoácido expuesto en la superficie” se refiere a un aminoácido que se presenta al exterior del polipéptido del que forma parte. En algunas modalidades, un residuo de aminoácido expuesto en la superficie está disponible para interactuar con otros polipéptidos o con solvente. En otras palabras, el residuo de aminoácido no está enterrado dentro de la cadena polipeptídica.

[0092] Como se usa en la presente, “residuo de aminoácido no polar” se refiere a un aminoácido o residuo hidrófobo que tiene una cadena lateral que no está cargada a pH fisiológico y que tiene enlaces en los que el par de electrones compartido en común por dos átomos generalmente está retenido de manera igual por cada uno de los dos átomos (es decir, la cadena lateral no es polar). Los aminoácidos no polares codificados genéticamente incluyen L-Gly (G), L-Leu (L), L-Val (V), L-Ile (I), L-Met (M) y L-Ala (A).

[0093] Como se usa en la presente, “residuo de aminoácido polar” se refiere a un aminoácido o residuo hidrófilo que tiene una cadena lateral que no está cargada a pH fisiológico, pero que

tiene al menos un enlace en el que el par de electrones compartido en común por dos átomos está retenido más estrechamente por uno de los dos átomos. Los aminoácidos polares codificados genéticamente incluyen L-Asn (N), L-Gln (Q), L-Ser (S) y L-Thr (T).

5 [0094] Una «posición» de aminoácido o base nucleotídica se denota mediante un número que identifica secuencialmente cada aminoácido (o base nucleotídica) en la secuencia de referencia basándose en su posición relativa al extremo N (o extremo 5'). Debido a las deleciones, inserciones, truncamientos, fusiones y similares que se deben tener en cuenta al determinar un alineamiento óptimo, en general el número de residuo de aminoácido en una secuencia de prueba determinada simplemente contando desde el extremo N no será necesariamente el mismo que el número de su posición correspondiente en la secuencia de referencia. Por ejemplo, en un caso de que una variante tenga una deleción relativa a una secuencia de referencia alineada, no habrá ningún aminoácido en la variante que corresponda a una posición en la secuencia de referencia en el sitio de la deleción. En el caso de que haya una inserción en una secuencia de referencia alineada, dicha inserción no corresponderá a una posición de aminoácido numerada en la secuencia de referencia. En el caso de truncamientos o fusiones puede haber tramos de aminoácidos en la secuencia de referencia o alineada que no correspondan a ningún aminoácido en la secuencia correspondiente.

20 [0095] El término “expresión” incluye cualquier etapa involucrada en la producción del polipéptido que incluye, de modo no taxativo, la transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción. En algunos aspectos, la expresión se puede detectar usando técnicas convencionales para detectar proteínas (*por ejemplo*, ELISA, transferencia Western, citometría de flujo, inmunofluorescencia, inmunohistoquímica, *etc.*).

25 [0096] Una “célula”, como se usa en la presente, se refiere a una unidad biológica unida a una membrana capaz de llevar a cabo una función metabólica o de otro tipo suficiente para preservar o replicar su ADN genómico. Una célula puede identificarse mediante métodos bien conocidos en la técnica, que incluyen, por ejemplo, la presencia de una membrana intacta, la tinción con un tinte particular, la capacidad de producir progenie o, en el caso de un gameto, la capacidad de combinarse con un segundo gameto para producir una descendencia viable. Las células pueden incluir células procariotas y eucariotas. Las células procariotas incluyen, entre otras, las bacterias. Las células eucariotas incluyen, entre otras, células de levadura y células derivadas de plantas y animales, por ejemplo, células de mamíferos, insectos (*por ejemplo*, *spodoptera*) y

30

células humanas. Las células pueden ser útiles cuando son naturalmente no adherentes o han sido tratadas para no adherirse a las superficies, por ejemplo mediante tripsinización.

[0097] La expresión «célula hospedadora» hace referencia a cualquier célula capaz de ser usada en la expresión de un polipéptido o polinucleótido recombinante. En algunas modalidades, la célula hospedadora es una célula bacteriana. En algunas modalidades, la célula bacteriana es una célula de *E. coli*.

[0098] Los términos "cultivo" y "medio de cultivo" se refieren al cultivo celular estándar y a las técnicas de expresión de proteínas recombinantes.

[0099] El término "producción" hace referencia a la producción de una nanoestructura proteica o polipeptídica recombinante en cualquier escala, lo que incluye al menos 25 mL, 50 mL, 1 L, 1.000 L, 50.000 L o a una escala mayor.

[0100] El término "purificar" se refiere a separar una molécula de otras sustancias presentes en una composición. Los polipéptidos se pueden purificar por afinidad (por ejemplo, a un anticuerpo o a una etiqueta, por ejemplo, utilizando una resina de captura con etiqueta His), por carga (por ejemplo, cromatografía de intercambio iónico), por tamaño (por ejemplo, ultracentrifugación preparativa, cromatografía de exclusión por tamaño) o de otro modo.

[0101] El término "estabilidad" se refiere a la estabilidad física y química de las formulaciones de polipéptidos. La inestabilidad física de una formulación de proteína puede ser causada por la agregación de las moléculas de proteína para formar polímeros de orden superior o incluso precipitados. Una formulación "estable" es aquella en la que el grado de agregación de proteínas que contiene está aceptablemente controlado y no aumenta de forma inaceptable con el tiempo. La estabilidad física se puede evaluar mediante métodos bien conocidos en la técnica, incluida la medición de la atenuación aparente de la luz de una muestra (absorbancia o densidad óptica). Esta medición de la atenuación de la luz se relaciona con la turbidez de una formulación. La turbidez se produce por agregación o precipitación de proteínas o complejos en la formulación.

[0102] El término "antigenicidad" se refiere a la capacidad de un antígeno de unirse específicamente a un receptor de linfocitos T o anticuerpo e incluye la reactividad de un antígeno hacia anticuerpos preexistentes en un sujeto. "Inmunogenicidad" se refiere a la capacidad de cualquier antígeno para inducir una respuesta inmune e incluye la capacidad intrínseca de un antígeno para generar anticuerpos en un sujeto.

[0103] El término «neutralizante» (por ejemplo, «respuesta de anticuerpo neutralizante») hace

referencia a los anticuerpos que previenen la infección y/o reducen el nivel de infección con un patógeno. Es posible medir una respuesta de anticuerpo neutralizante ya sea en ensayos *in vitro* (por ejemplo, destrucción de bacterias diana en presencia del anticuerpo) o en un ensayo *in vivo* (por ejemplo, mediante la determinación de una dosis de protección de un anticuerpo a través de la administración del anticuerpo a un sujeto antes de la exposición a una dosis de infección de un patógeno).

[0104] Un anticuerpo «se une a», «es específico para» o «se une específicamente» (usados de forma indistinta en la presente) a una diana (por ejemplo, la proteína bacteriana) son términos conocidos en la técnica y también se conocen en la técnica métodos para determinar dicha unión específica o preferencial. Se dice que una molécula exhibe «unión específica» o «unión preferencial» si reacciona o se asocia más frecuentemente, más rápidamente, con más duración y/o con mayor afinidad con una célula o sustancia en particular que con células o sustancias alternativas.

[0105] Tal como se usa en la presente, la expresión «nanoestructura proteica» hace referencia a arreglos proteicos simétricos en los que las subunidades se ensamblan en solución acuosa, sin que sean necesarios lípidos o macromoléculas diferentes de la nanoestructura proteica para el ensamblaje. Se describen ejemplos de nanoestructuras proteicas en Hsia et ál. *Nature* 35:136-9 (2016) y Bale et ál. *Science* 353:389-394 (2016). En algunas modalidades, la nanoestructura proteica es una nanoestructura proteica de un componente, en la que un único tipo de polipéptido proporciona los componentes básicos que se ensamblan para formar la nanoestructura proteica. En algunas modalidades, la nanoestructura proteica es una nanoestructura proteica de dos componentes, en la que dos tipos de polipéptidos proporcionan los componentes básicos que se ensamblan para formar la nanoestructura proteica. En algunas modalidades, los tipos de polipéptidos incluyen un dominio de ensamblaje que hace que el polipéptido forme componentes diméricos, triméricos, tetraméricos, hexaméricos u otros componentes multiméricos. En una nanoestructura de dos componentes, los dos componentes difieren en la selección de un dominio de ensamblaje. En algunas modalidades, el dominio de ensamblaje del primer tipo de polipéptido hace que el polipéptido forme un trímero y el dominio de ensamblaje del segundo tipo de polipéptido hace que el polipéptido forme un pentámero.

[0106] En algunas modalidades de nanoestructuras de un componente, o relacionadas con estas, dos o más copias del componente se ensamblan adicionalmente de forma simétrica para

formar la nanoestructura. En algunas modalidades de nanoestructuras de dos componentes, o relacionadas con estas, dos o más de cada uno de los dos componentes diferentes se ensamblan adicionalmente de forma simétrica para formar la nanoestructura.

[0107] Tal como se usa en la presente, «dominio de ensamblaje» hace referencia a la parte de una subunidad de un componente involucrada en la formación de la nanoestructura proteica a través de interacciones intracomponente e interacciones con otras copias del mismo componente (en una nanoestructura de un componente) o con otros componentes (por ejemplo, en una nanoestructura de dos componentes).

[0108] La expresión «partícula icosaédrica» hace referencia a nanoestructuras oriteucas con un núcleo con simetría icosaédrica. I53 hace referencia a una partícula icosaédrica construida a partir de pentámeros y trímeros. I52 se refiere a una partícula icosaédrica construida a partir de pentámeros y dímeros. T33 se refiere a una partícula tetraédrica construida a partir de dos conjuntos de trímeros. T32 se refiere a una partícula tetraédrica construida a partir de trímeros y dímeros. Como se usan en la presente, las abreviaturas I53-50A e I53-50B se refieren a los dos componentes o cadenas polipeptídicas (A y B) de una partícula de tipo I53 particular (es decir, una partícula I53-50).

[0109] Los polipéptidos potencialmente antigénicos pueden estar unidos al núcleo de la nanoestructura proteica ya sea de forma no covalente o covalente, lo que incluye una proteína de fusión o de otra forma descrita en la presente. Opcionalmente, se pueden presentar polipéptidos a lo largo de un eje de simetría de la nanoestructura proteica. También se proporcionan proteínas y moléculas de ácido nucleico que codifican dichas proteínas, formulaciones y métodos de uso.

[0110] Tal como se usa en la presente, el término «unido» con referencia a dos restos significa que los dos restos están unidos, en donde el enlace o los enlaces que conectan los dos restos pueden ser covalentes o no covalentes. En algunas modalidades, los dos restos están enlazados de forma covalente entre sí (por ejemplo, directamente o a través de un intermedio enlazado de forma covalente (un "enlazador")). En algunas modalidades, los dos restos se enlazan de forma no covalente (por ejemplo, a través de enlaces iónicos, interacciones o enlaces de Van der Waal, enlaces de hidrógeno, enlaces polares o combinaciones o mezclas de estos). En un polipéptido de fusión unido al enlazador en un extremo N y un extremo C es un enlace peptídico.

[0111] Tal como se usan en la presente, los términos «bioconjugado» y «enlazador bioconjugado» hacen referencia a la asociación resultante entre átomos o moléculas de «grupos reactivos bioconjugados» o «restos reactivos bioconjugados». La asociación puede ser directa o indirecta. Por ejemplo, un conjugado entre el primer grupo reactivo bioconjugado (por ejemplo, $-\text{NH}_2$, $-\text{C}(\text{O})\text{OH}$, $-\text{N}$ -hidrosuccinimida o maleimida) y un segundo grupo reactivo bioconjugado (por ejemplo, sulfhidrilo, un aminoácido que contiene azufre, una amina, un aminoácido que contiene una cadena lateral de amina o un carboxilato) proporcionado en la presente puede ser directo, por ejemplo, mediante un enlace covalente o enlazador (por ejemplo, un primer enlazador o un segundo enlazador), o indirecto, por ejemplo, mediante un enlace no covalente (por ejemplo, interacciones electrostáticas (por ejemplo, enlace iónico, puente de hidrógeno, enlace de halógeno), interacciones de van der Waals (por ejemplo, dipolo-dipolo, dipolo-dipolo inducido, fuerza de dispersión de London), apilamiento de anillos (efecto pi), interacciones hidrófobas y similares). En algunas modalidades, los bioconjugados o enlazadores bioconjugados se forman con química de bioconjugación (es decir, la asociación de dos grupos reactivos bioconjugados) que incluye, de modo no taxativo, sustituciones nucleofílicas (por ejemplo, reacciones de aminas y alcoholes con haluros de acilo, ésteres activos), sustituciones electrofílicas (por ejemplo, reacciones de enamina) y adiciones a enlaces múltiples carbono-carbono y carbono-heteroátomo (por ejemplo, reacción de Michael, adición de Diels-Alder). Estas y otras reacciones útiles se mencionan, por ejemplo, en March, ADVANCED ORGANIC CHEMISTRY, 3.a ed., John Wiley & Sons, Nueva York, 1985; Hermanson, BIOCONJUGATE TECHNIQUES, Academic Press, San Diego, 1996; y Feeney et ál., MODIFICATION OF PROTEINS; Advances in Chemistry Series, Vol. 198, American Chemical Society, Washington, D.C., 1982. En algunas modalidades, el primer grupo reactivo bioconjugado (por ejemplo, resto de maleimida) se une de forma covalente al segundo grupo reactivo bioconjugado (por ejemplo, un sulfhidrilo). En algunas modalidades, el primer grupo reactivo bioconjugado (por ejemplo, resto de haloacetilo) se une de forma covalente al segundo grupo reactivo bioconjugado (por ejemplo, un sulfhidrilo). En algunas modalidades, el primer grupo reactivo bioconjugado (por ejemplo, resto de piridilo) se une de forma covalente al segundo grupo reactivo bioconjugado (por ejemplo, un sulfhidrilo). En algunas modalidades, el primer grupo reactivo bioconjugado (por ejemplo, resto de $-\text{N}$ -hidroxisuccinimida) se une de forma covalente al segundo grupo reactivo bioconjugado (por ejemplo, una amina). En algunas modalidades, el primer grupo

reactivo bioconjugado (por ejemplo, resto de maleimida) se une de forma covalente al segundo grupo reactivo bioconjugado (por ejemplo, un sulfhidrilo). En algunas modalidades, el primer grupo reactivo bioconjugado (por ejemplo, resto de –sulfo–N-hidroxisuccinimida) se une de forma covalente al segundo grupo reactivo bioconjugado (por ejemplo, una amina).

5 **[0112]** Por ejemplo, los restos reactivos bioconjugados usados para reacciones químicas de bioconjugación en la presente incluyen:

10 (a) grupos carboxilos y varios derivados de estos que incluyen, de modo no taxativo, ésteres de N-hidroxisuccinimida, ésteres de N-hidroxibenzotriazol, haluros ácidos, imidazoles de acilo, tioésteres, ésteres de p-nitrofenilo, ésteres de alquilo, alquenilo, alquinilo y aromáticos;

(b) grupos hidroxilo que se pueden convertir en ésteres, éteres, aldehídos, etc.

15 (c) grupos haloalquilo en donde el haluro puede ser desplazado por un grupo nucleofílico, tal como, por ejemplo, una amina, un anión carboxilato, un anión tiol, un carbanión o un ion alcóxido, lo que resulta en la unión covalente de un nuevo grupo en el sitio del átomo de halógeno;

(d) grupos dienófilos capaces de participar en reacciones de Diels-Alder tales como, por ejemplo, grupos maleimido o maleimida;

20 (e) grupos aldehído o cetona tales que sea posible la derivación posterior mediante la formación de derivados de carbonilo como, por ejemplo, iminas, hidrazonas, semicarbazonas u oximas, o mediante mecanismos como adición de Grignard o adición de alquil-litio;

(f) grupos de haluro de sulfonilo para reacciones posteriores con aminas, por ejemplo, para formar sulfonamidas;

25 (g) grupos tiol, que se pueden convertir en disulfuros, reaccionar con haluros de acilo o enlazarse con metales como oro, o reaccionar con maleimidas;

(h) grupos amina o sulfhidrilo (por ejemplo, presente en cisteína), que, por ejemplo, pueden estar acetilados, alquilados u oxidados;

(i) alquenos, que se pueden someter, por ejemplo, a cicloadiciones, acilación, adición de Michael, etc;

30 (j) epóxidos, que pueden reaccionar, por ejemplo, con aminas y compuestos de hidroxilo;

(k) fosforamiditas y otros grupos funcionales estándar útiles en la síntesis de ácidos nucleicos;

(l) enlace de óxido de silicio metálico; y

(m) enlace metálico a grupos de fósforo reactivos (por ejemplo, fosfinas) para formar, por ejemplo, enlaces de diéster de fosfato.

(n) azidas acopladas a alquinos que usan química de clic de cicloadición catalizada por cobre.

(o) conjugado de biotina que puede reaccionar con avidina o estreptavidina para formar un complejo de avidina y biotina o un complejo de estreptavidina y biotina.

10 **[0113]** Es posible seleccionar los grupos reactivos bioconjugados de forma que no participen en la estabilidad química del conjugado descrito en la presente ni interfieran con ella. De manera alternativa, es posible proteger a un grupo funcional reactivo de la participación en la reacción de formación de enlaces mediante la presencia de un grupo protector. En algunas modalidades, el bioconjugado comprende una entidad molecular derivada de la reacción de un enlace no saturado, tal como una maleimida, y un grupo sulfhidrilo.

15 **[0114]** El sistema SpyTag/SpyCatcher es otro método de enlace covalente de dos polipéptidos. El péptido SpyTag (13 aminoácidos) reacciona de forma espontánea con la proteína SpyCatcher (12,3 kDa) para formar un enlace isopeptídico intermolecular entre el par. La secuencia polinucleotídica que codifica SpyTag o SpyCatcher se puede introducir de forma recombinante en la secuencia polinucleotídica que codifica polipéptidos de interés para formar una proteína de fusión. Tres proteínas de fusión se pueden enlazar de forma covalente cuando se mezclan en una reacción a través del sistema SpyTag/SpyCatcher.

20 **[0115]** Es posible lograr con el par de etiqueta y receptor (Tag/Catcher) la bioconjugación entre dos proteínas recombinantes que, de otra manera, sería limitada o imposible con la fusión genética directa tradicional entre las dos proteínas. Por ejemplo, los problemas relacionados con el plegamiento de proteínas, la expresión subóptima del huésped y las modificaciones postraduccionales especializadas pueden mitigarse separando la producción de las proteínas con la modularidad del sistema Tag/Catcher. Las nanoestructuras proteicas pueden emplear diversas técnicas de acoplamiento para unir un antígeno, incluyendo, entre otras, el sistema
 25 SpyCatcher descrito en, *por ejemplo*, Escolano et al. Nature 570:468-473 (2019), He et al. Sci Adv. 7(12):eabf1591 (2021) y Tan et al. Nat. Commun. 12(1):542 (2021).

[0116] El término “dominio” se refiere a cualquier porción de un polipéptido que adopta una estructura terciaria.

[0117] Los términos «dominio de ensamblaje» y «multimerizar» hacen referencia a la capacidad de un polipéptido o dominio de un polipéptido para formar una estructura terciaria con otro polipéptido o dominio de un polipéptido. En algunas modalidades, los dominios de ensamblaje pueden formar dímeros, trímeros, tetrameros, pentámeros o hexámeros y/o pueden formar heterómeros con otros dominios de ensamblaje. En algunas modalidades, el dominio de ensamblaje forma una nanoestructura proteica.

[0118] La expresión «dominio de trimerización» hace referencia a un dominio de ensamblaje que forma trímeros.

[0119] El término «fragmento» hace referencia a un polipéptido que tiene uno o más truncamientos en el extremo N o el extremo C en comparación con un polipéptido de referencia.

[0120] La expresión «fragmento funcional» hace referencia a un fragmento que retiene al menos una función de su polipéptido de referencia.

[0121] Como se usa en la presente, el término «conjugado» con referencia a dos restos significa que los dos restos están unidos, en donde el enlace o los enlaces que conectan los dos restos pueden ser covalentes o no covalentes. En algunas modalidades, los dos restos están enlazados de forma covalente entre sí (*por ejemplo*, directamente o a través de un intermedio enlazado de forma covalente). En algunas modalidades, los dos restos están unidos de forma no covalente (*por ejemplo*, interacciones electrostáticas (*por ejemplo*, enlace iónico, enlace de hidrógeno, enlace de halógeno), enlaces/interacciones de Van der Waals (*por ejemplo*, dipolo-dipolo, dipolo-dipolo inducido, dispersión de London), apilamiento de anillos (efectos pi), interacciones hidrófobas y similares), enlaces polares o combinaciones o mezclas de estos).

[0122] El término “proteína de fusión” se refiere a una proteína quimérica que codifica dos o más secuencias de proteínas separadas que se expresan de forma recombinante como un polipéptido, logrando que se acoplen de forma covalente. En un polipéptido de fusión unido al enlazador en un extremo N y un extremo C puede ser un enlace peptídico o péptido de cualquier longitud.

[0123] El término “enlazador” se refiere al enlace de un péptido o polipéptido que une químicamente el extremo C de un primer segmento polipeptídico al extremo N de un segundo segmento polipeptídico, formando así una proteína de fusión que contiene los dos segmentos

polipeptídicos.

[0124] El término "secuencia señal" se refiere a una secuencia polipeptídica, típicamente en el extremo N de un polipéptido expresado en una célula huésped que dirige el polipéptido a un compartimento celular particular. Una secuencia señal puede ser una señal de secreción para hacer que la célula huésped secrete el polipéptido en el medio en el que se cultiva la célula huésped. La secuencia señal puede ser una secuencia señal "natural", una secuencia señal que ocurre en la naturaleza como parte del polipéptido. La secuencia señal puede ser una secuencia que no ocurre de forma natural con el polipéptido tal como se encuentra en la naturaleza. Se conocen varias secuencias señal y está dentro de la capacidad de un experto en la técnica para seleccionar una secuencia señal apropiada.

[0125] El término «secretar» hace referencia a la capacidad de las células hospedadoras para liberar polipéptidos expresados en el medio en el que se cultivan.

[0126] Como se pueden usar en la presente, los términos "ácido nucleico", "molécula de ácido nucleico", "oligómero de ácido nucleico", "oligonucleótido", "secuencia de ácido nucleico", "fragmento de ácido nucleico" y "polinucleótido" se usan indistintamente y pretenden incluir, entre otros, una forma polimérica de nucleótidos unidos covalentemente entre sí que pueden tener varias longitudes, ya sean desoxirribonucleótidos o ribonucleótidos, o análogos, derivados o modificaciones de estos. Diferentes polinucleótidos pueden tener diferentes estructuras tridimensionales y pueden realizar diversas funciones, conocidas o desconocidas. Los ejemplos no limitantes de polinucleótidos incluyen un gen, un fragmento de gen, un exón, un intrón, ADN intergénico (incluido, entre otros, ADN heterocromático), ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, una ribozima, ADNc, un polinucleótido recombinante, un polinucleótido ramificado, un plásmido, un vector, ADN aislado de una secuencia, ARN aislado de una secuencia, una sonda de ácido nucleico y un cebador. Los polinucleótidos útiles en los métodos de la descripción pueden comprender secuencias de ácidos nucleicos naturales y variantes de estas, secuencias de ácidos nucleicos artificiales o una combinación de dichas secuencias.

[0127] El término "ARN mensajero" o "ARNm" se refiere a un polinucleótido que codifica al menos un polipéptido. El término ARNm, como se usa en la presente, abarca tanto el ARN modificado como el no modificado. El ARNm puede contener uno o más marcos o regiones de lectura.

[0128] Las expresiones «excipientes farmacéuticamente aceptables» y «vehículo farmacéuticamente aceptable» hacen referencia a una sustancia que favorece la administración de un agente activo a un sujeto y su absorción, y se puede incluir en las composiciones de la presente descripción sin provocar un efecto toxicológico adverso en el paciente y pueden significar excipientes aprobados por una agencia reguladora del gobierno federal o estatal o mencionados por la Farmacopea de los Estados Unidos u otra farmacopea generalmente reconocida para su uso en animales y, más particularmente, en humanos. Los ejemplos no limitantes de excipientes farmacéuticamente aceptables incluyen agua, NaCl, soluciones salinas normales, Ringer lactado, sacarosa normal, glucosa normal, aglutinantes, rellenos, desintegrantes, lubricantes, recubrimientos, edulcorantes, saborizantes, soluciones salinas (como solución de Ringer), alcoholes, aceites, gelatinas, carbohidratos como lactosa, amilosa o almidón, ésteres de ácidos grasos, hidroximetilcelulosa, polivinilpirrolidina y colorantes, y similares. Dichas preparaciones se pueden esterilizar y, si se desea, se pueden mezclar con agentes auxiliares tales como lubricantes, conservantes, estabilizantes, agentes humectantes, emulsionantes, sales para influir en la presión osmótica, amortiguadores, colorantes y/o sustancias aromáticas y similares que no reaccionan de forma nociva con los compuestos de la descripción. Un experto en la técnica reconocerá que son útiles otros excipientes farmacéuticos en la presente descripción.

[0129] El término "vacuna" se refiere a una composición que puede proporcionar inmunidad adquirida activa a una enfermedad particular (*por ejemplo*, COVID) o un patógeno (*por ejemplo*, coronavirus). Usualmente, una vacuna comprende uno o más agentes que pueden inducir una respuesta inmunitaria en un sujeto contra un patógeno o enfermedad, *es decir*, una enfermedad o patógeno objetivo. El agente inmunogénico estimula la respuesta inmunitaria del cuerpo para que reconozca al agente como una amenaza o indicación de la presencia del patógeno o enfermedad diana, lo que induce la memoria inmunológica, de manera que el sistema inmunitario pueda reconocer y destruir más fácilmente a cualquiera de los patógenos en una exposición posterior. Las vacunas pueden ser profilácticas (*por ejemplo*, prevenir o mejorar los efectos de una infección futura con cualquier organismo natural o patógeno, o de una aparición anticipada del cáncer en un sujeto predispuesto) o terapéuticas (*por ejemplo*, tratar una enfermedad, tal como, entre otros, tratar la meningitis en un sujeto que lo necesita). Se hace referencia a la administración de vacunas como vacunación. En algunos ejemplos, una

composición de vacuna puede proporcionar moléculas antigénica o un ácido nucleico, *por ejemplo*, ARNm, que codifica las moléculas antigénicas (*por ejemplo*, péptidos) a un sujeto. La molécula antigénica o el ácido nucleico que se administra a través de la composición de vacuna en el sujeto se puede expresar en moléculas antigénicas y permite al sujeto adquirir inmunidad

5 contra las moléculas antigénicas. En el contexto de la vacunación contra una enfermedad infecciosa, la composición de vacuna puede proporcionar moléculas antigénicas o ARNm que codifica moléculas antigénicas asociadas con un patógeno determinado, *por ejemplo*, uno o más péptidos que se sabe que se expresan en el patógeno (*por ejemplo*, virus o bacteria patógena).

10 **[0130]** El término "vacuna bivalente" se refiere a una vacuna preparada a partir de dos cepas diferentes. Dichas vacunas pueden combinarse con otros componentes inmunogénicos o antigénicos para formar una composición de vacuna (p. ej., combinadas con uno o más antígenos proteicos recombinantes).

15 **[0131]** El término "adyuvantes" se refiere a una sustancia farmacéuticamente aceptable que mejora la respuesta inmune a un antígeno cuando se coadministra con el antígeno o se administra antes, durante o después de la administración del antígeno a un sujeto.

[0132] El término "sujeto" se refiere a un animal humano o no humano al que se le puede administrar una composición para vacunación, tratamiento u otro propósito. En algunas modalidades, el animal no humano es un primate no humano, conejo, hámster, gerbil, cerdo,

20 vaca, oveja, cabra, conejillo de Indias, rata, ratón, ardilla, lobo, zorro, caballo, jirafa, hiena elefante, gato, perro, llama o hurón.

[0133] La expresión "cantidad eficaz" hace referencia a la cantidad de una composición que, cuando se administra a un paciente para tratar un estado, trastorno o afección es suficiente para lograr dicho tratamiento o cuando se administra a un paciente para generar una respuesta

25 inmunitaria es suficiente para generar dicha respuesta inmunitaria. Las cantidades exactas dependerán del principio activo, el estado, trastorno o afección a tratar y su gravedad, y la edad, el peso, el estado físico y la capacidad de respuesta del sujeto a tratar, y las podrá determinar un experto en la técnica con técnicas conocidas (vea, por ejemplo, Lieberman, Pharmaceutical Dosage Forms (tomos. 1-3, 1992); Lloyd, The Art, Science and Technology of Pharmaceutical

30 Compounding (1999); Pickar, Dosage Calculations (1999); y Remington: The Science and Practice of Pharmacy, 20.ª edición, 2003, Gennaro, Ed., Lippincott, Williams & Wilkins).

[0134] Como se usa en la presente, el término “administrar” significa proporcionar una entidad a un destino. Por ejemplo, administrar un agente terapéutico y/o profiláctico a un sujeto puede implicar administrar una composición de nanopartículas que incluye el agente terapéutico y/o profiláctico al sujeto (por ejemplo, por vía intravenosa, intramuscular, intradérmica, intraperitoneal, intratumoral o subcutánea). La administración de una composición de nanopartículas a un mamífero o a una célula de mamífero puede implicar el contacto de una o más células con la composición de nanopartículas.

[0135] Como se usa en la presente, el término “prevenir” se refiere a una disminución en la aparición de síntomas de una enfermedad en un paciente. La prevención puede ser completa (sin síntomas detectables) o parcial, de modo que se observen menos síntomas de los que probablemente ocurrirían sin tratamiento.

[0136] “Tratar” o “tratamiento” como se usa en la presente (y como se entiende bien en la técnica) se refiere ampliamente a cualquier enfoque para obtener resultados beneficiosos o deseados en la condición de un sujeto, incluidos los resultados clínicos. Los resultados clínicos beneficiosos o deseados pueden incluir, entre otros, alivio o mejora de uno o más síntomas o condiciones, disminución de la extensión de una enfermedad, estabilización (*es decir*, no empeoramiento) del estado de la enfermedad, prevención de la transmisión o propagación de una enfermedad, retraso o desaceleración de la progresión de la enfermedad, mejora o paliación del estado de la enfermedad, disminución de la reaparición de la enfermedad y remisión, ya sea parcial o total y ya sea detectable o indetectable. En otras palabras, “tratamiento” como se usa en la presente incluye cualquier cura, mejora o prevención de una enfermedad. El tratamiento puede prevenir la aparición de la enfermedad, inhibir su propagación, aliviar los síntomas de la enfermedad, eliminar total o parcialmente la causa subyacente de la enfermedad, acortar la duración de la enfermedad o realizar una combinación de estas cosas.

[0137] El término “administrar” se refiere a proporcionar una composición a un sujeto de una manera que permita que la composición tenga el efecto previsto. La administración para vacunación o profilaxis posterior a la exposición puede realizarse mediante inyección intramuscular, inyección intravenosa, inyección intraperitoneal o cualquier otra vía adecuada.

[0138] La “coadministración” significa que una composición descrita en la presente se administra en el mismo momento, inmediatamente antes o inmediatamente después de la

administración de uno o más tratamientos adicionales. Las composiciones proporcionadas en la presente se pueden administrar solos o en coadministración al sujeto. La coadministración pretende incluir la administración simultánea o secuencia de los compuestos de forma individual o en combinación (más de un compuesto). Por lo tanto, las preparaciones también se pueden
5 combinar, cuando se desea, con otras sustancias activas (por ejemplo, para reducir la degradación metabólica).

[0139] Los términos «inmunización» e «inmunizar» hacen referencia a la administración de una composición a un sujeto en una cantidad suficiente para provocar, después de una o más etapas de administración, una respuesta inmunitaria deseada. La inmunización puede comprender
10 entre una y diez, o más, administraciones (por ejemplo, inyecciones) de la composición, tal como 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más administraciones. Es posible que la primera administración no produzca una respuesta inmunitaria detectable, ya que generalmente cada administración posterior reforzará la respuesta inmunitaria generada por administraciones anteriores. Tal como se usa en la presente, el término «inmunización» incluye la profilaxis posterior a la exposición.

[0140] El término "respuesta inmunitaria protectora" se refiere a una respuesta inmunitaria que previene y/o reduce la gravedad de la infección con un patógeno cuando el sujeto se expone
15 posteriormente al patógeno, o a una respuesta inmunitaria que genera un nivel de respuesta inmunitaria que se correlaciona con la protección. Por ejemplo, la vacunación puede generar una respuesta inmune protectora si da como resultado la producción, en el plasma o suero, del
20 sujeto (por ejemplo, un ser humano, una mascota o un animal agrícola), de anticuerpos neutralizantes que protegen al sujeto contra una infección posterior y/o están presentes en una cantidad que se observa que confiere protección a los sujetos de prueba (por ejemplo, conejos blancos de Nueva Zelanda (NZW)).

[0141] El término "anticuerpo" se refiere a un polipéptido codificado por un gen de
25 inmunoglobulina o fragmentos funcionales del mismo que se une específicamente y reconoce un antígeno. Los genes de inmunoglobulina reconocidos incluyen los genes de la región constante kappa, lambda, alfa, gamma, delta, épsilon y mu, así como una gran cantidad de genes de la región variable de inmunoglobulina. Las cadenas ligeras se clasifican como kappa o lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta o épsilon, que a su
30 vez definen las clases de inmunoglobulinas, IgG, IgM, IgA, IgD e IgE, respectivamente.

[0142] El término "infección" o "enfermedad infecciosa" se refiere a una enfermedad o condición

que puede ser causada por organismos tales como bacterias, virus, hongos o cualquier otro agente microbiano patógeno. En algunas modalidades, la enfermedad infecciosa es causada por una bacteria patógena. En algunas modalidades, la enfermedad infecciosa es una infección por *Campylobacter jejuni*, *Enterococcus faecalis*, *Haemophilus influenzae*, *Helicobacter pylori*,
 5 *Klebsiella pneumoniae*, *Legionella pneumophila*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Staphylococcus aureus*, *Streptococcus pneumoniae* o *Vibrio cholera*.

[0143] El término “infección viral” o “enfermedad viral” se refiere a una enfermedad o condición causada por un virus. Los ejemplos no limitantes de infecciones virales incluyen enfermedades virales hepáticas (p. ej., hepatitis A, B, C, D, E), infección por el virus del herpes (p. ej., HSV-1, HSV-2, herpes zóster), infección por flavivirus, infección por el virus del Zika, infección por citomegalovirus, una infección viral respiratoria (p. ej., infección por adenovirus, influenza, síndrome respiratorio agudo severo, infección por coronavirus (p. ej., SARS-CoV-1, SARS-CoV-2, MERS-CoV, COVID-19, MERS)), una infección viral gastrointestinal (p. ej., infección por norovirus, infección por rotavirus, infección por astrovirus), una infección viral exantemática (p.
 10 ej., sarampión, herpes zóster, viruela, rubéola), enfermedad hemorrágica viral (p. ej., Ébola, fiebre de Lassa, dengue, fiebre amarilla), una infección viral neurológica (p. ej., infección por el virus del Nilo Occidental, polio, meningitis viral, encefalitis viral, encefalitis japonesa, rabia) e infección por el virus del papiloma humano.

Polipéptido recombinante del coronavirus

[0144] En un aspecto, la descripción proporciona un polipéptido recombinante, que comprende un fragmento antigénico de una glucoproteína Spike de coronavirus que comprende un dominio de unión al receptor (RBD) y, opcionalmente, un subdominio 1 de coronavirus (SD1). El fragmento antigénico comprende RBD-SD1. En algunas modalidades, el fragmento antigénico comprende esencialmente RBD-SD1. En algunas modalidades, el fragmento antigénico
 25 comprende la porción RBD-SD1 aislada del remanente de la glucoproteína Spike. Es posible que se use alguna secuencia de flaqueo alrededor del dominio. En algunas modalidades, el fragmento antigénico comprende RBD. Las sustituciones de aminoácidos pueden introducirse en las glucoproteína Spike del coronavirus, o en fragmentos de las mismas, de diversas cepas de coronavirus (incluidas, entre otras, MERS, SARS-CoV-1 y SARS-CoV-2).

[0145] Se establece en la presente una secuencia ilustrativa de glucoproteína Spike de SARS-CoV-2 como la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende una

secuencia polipeptídica al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID

5 NO: 1.

```

1 MFVFLVLLPL VSSQCVNLTT RTQLPPAYTN SFTRGVYYPD KVERSSVLHS
51 TQDLFLPFFS NVTWFHAIHV SGTNGTKRFD NPVLPFNDGV YFASTEKSNI
101 IRGWIFGTTL DSKTQSLIV NNATNVVIKV CEFQFCNDPF LGVYYHKNNK
151 SWMESEFRVY SSANNCTFEY VSQPFLMDLE GKQGNFKNLR EFVFKNIDGY
201 FKIYSKHTPI NLVRDLPQGF SALEPLVDLP IGINITRFQT LLALHRSYLT
251 PGDSSSGWTA GAAAYYVGYL QPRTFLLKYN ENGTITDAVD CALDPLSETK
301 CTLKSFTVEK GIYQTSNFRV QPTESIVREF NITNLCPFGE VFNATREFASV
351 YAWNRKRISN CVADYSVLYN SASFSTFKCY GVSPTKLNDL CFTNVYADSF
401 VIRGDEVROI APGQTGKIAD YNYKLPDDET GCVIAWNSNN LDSKVGGNYN
451 YLYRLFRKSN LKPFERDIST EIQAGSTPC NGVEGFNCYF PLQSYGFQPT
501 NGVGYQPYRV VVLSFELLHA PATVCGPKKS TNLVKNKCVN FNFNGLTGTG
551 VLTESNKKFL PFQQFGRDIA DTTDAVRDPQ TLEILDITPC SFGGVSVITP
601 GTNTSNQVAV LYQDVNCTEV PVAIHADQLT PTWRVYSTGS NVFQTRAGCL
651 IGAEHVNNYSY ECDIPIGAGI CASYQTQTN SRRARSVASQ SIIAYTMSLG
701 AENSVAYSNN SIAIPTNFTI SVTTEILPVS MTKTSVDCTM YICGDSTEC S
751 NLLLQYGSFC TQLNRALTGI AVEQDKNTQE VFAQVKQIYK TPPIKDFGGF
801 NFSQILPDPS KPSKRSFIED LLFNKVTLAD AGFIKQYGDC LGDIAARDLI
851 CAQKFENGLTV LPPLLTDEMI AQYTSALLAG TITSGWTFGA GAALQIPFAM
901 QMAYRFNGIG VTQNVLYENQ KLIANQFNSA IGKIQDSLSS TASALGKLQD
951 VVNQNAQALN TLVKQLSSNF GAISSVLNDI LSRLDKVEAE VQIDRLITGR
1001 LQSLQTYVTQ QLIRAAEIRA SANLAATKMS ECVLGQSKRV DFCGKGYHLM
1051 SFPQSAPHGV VFLHVTYVPA QEKNETTAPA ICHDGKAHFP REGVFSNGT
1101 HWFVTQRNFY EPQIITDNT FVSGNCDVVI GIVNNTVYDP LQPELDSFKE
1151 ELDKYFKNHT SPDVDLGDIS GINASVVNIQ KEIDRLNEVA KNLNESLIDL
1201 QELGKYEQYI KWPWYIWLGF IAGLIAIVMV TIMLCCMTSC CSCCLKGCCSC
1251 GSCCKFDEDD SEPVLKGVKL HYT

```

(SEQ ID No: 1)

[0146] En la **Tabla 1** se proporcionan secuencias de fragmentos antigénicos ilustrativos de diversas cepas de coronavirus. En algunas modalidades, la secuencia de fragmento antigénico es al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos

85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 99 %, o 100 % idéntica a cualquier secuencia polipeptídica enumerada en la

Tabla 1.

5 **Tabla 1.**

SEQ ID NO	Secuencia	Cepa
2	RVQPTESIVRFPNITNLPCPFGEVFNATRFASVYAWNRR KRISNCVADYSVLVNSASFSTFKCYGVSP TKLNDLCFTN VYADSFVIRGDEV RQIAPGQTGKIADYNYKLPDDFTGC VIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDIST EIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQP YRVVVL SFELLHAPATVCGPKKSTNLVKNKCVNFNFNG LTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEI LDITPCS	Ancestral
3	RVQPTESIVRFPNITNLCPFDEVFNATRFASVYAWNRRK RISNCVADYSVLVNFAPFFAFKCYGVSP TKLNDLCFTNV YADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVI AWNSNKLDSKVGGNYNYRYRLFRKSNLKPFERDISTEI YQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGHQPHY RVVVL SFELLHAPATVCGPKKSTNLVKNKCVNFNFNGL TGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEIL DITPCS	BA.5
4	RVQPTESIVRFPNITNLCPFDEVFNATRFASVYAWNRRK RISNCVADYSVLVNLAPFFTFKCYGVSP TKLNDLCFTNV YADSFVIRGDEV RQIAPGQTGNIADYNYKLPDDFTGCVI AWNSNKLDSKVSGNYNYLYRLFRKSNLKPFERDISTEIY QAGNKPCNGVAGFNCYFPLRSYSFRPTYGVGHQPHYRV VVL SFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDIT PCS	BA.1
5	RVQPTESIVRFPNITNLCPFHEVFNATTFASVYAWNRRKR	XBB

	ISNCVADYSVIYNFAPFFAFKCYGVSP TKLNDLCFTNVY ADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVIA WNSNKLD SKPSGNYNYLYRLFRKSKLKPFERDISTEIYQ AGNKPCNGVAGSNCYSPLQSYGFRPTYGVGHQPYRV VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDIT PCS	
6	RVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERK KISNCVADYSVLYNSTFFSTFKCYGV SATKLNDLCFSNV YADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCV LAWNTRNIDATSTGNYNYKYRYLRH GKLRPFERDISNV PFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRV VVLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGT GVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDIS PCS	SARS-CoV-1
7	RVQPTESIVRFPNITNLCPFDEVFNATRFASVYAWN RK RISNCVADYSVLYNFAPFFAFKCYGVSP TKLNDLCFTNV YADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVI AWNSNKLD SKVGGNYNYLYRLFRKSNLKPFERDISTEI YQAGNKPCNGVAGFNCYFPLRSYGFRPTYGVGHQPY RVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGL TGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEIL DITPCS	BA.2
8	RVQPTESIVRFPNITNLCPFDEVFNATRFASVYAWN RK RISNCVADYSVLYNFAPFFAFKCYGVSP TKLNDLCFTNV YADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVI AWNSNKLD SKVGGNYNYLYRLFRKSNLKPFERDISTEI YQAGNKPCNGVAGFNCYFPLRSYGFRPTYGVGHQPY RVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGL TGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEIL DITPCS	BA.2.12

9	RVQPTESIVRFPNITNLCPFDEVFNATRFASVYAWNRRK RISNCVADYSVLYNFAPFFAFKCYGVSP TKLNDLCFTNV YADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVI AWNSNKLD SKVGGNYNYQYRLFRKSNLKPFERDISTEI YQAGNKPCNGVAGFNCYFPLRSYGFRPTYGVGHQPY RVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGL TGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEIL DITPCS	BA.2.75
10	RVQPTESIVRFPNITNLCPFHEVFNATTFASVYAWNRRK ISNCVADYSVIYNFAPFFAFKCYGVSP TKLNDLCFTNVY ADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVIA WNSNKLD SKPSGNYNYLYRLFRKSKLKPFERDISTEIYQ AGNKPCNGVAGPNCYSPLQSYGFRPTYGVGHQPYRV VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDIT PCS	XBB.1.5
245	RVQPTESIVRFPNITNLCPFHEVFNATTFASVYAWNRRK ISNCVADYSVIYNFAPFFAFKCYGVSP TKLNDLCFTNVY ADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVIA WNSNKLD SKPSGNYNYLYRLLRKS KLKPFERDISTEIYQ AGNKPCNGVAGPNCYSPLQSYGFRPTYGVGHQPYRV VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDIT PCS	EG.5.1
246	RVQPTESIVRFPNVTNLCPFHEVFNATRFASVYAWNRT RISNCVADYSVLYNFAPFFAFKCYGVSP TKLNDLCFTNV YADSFVIKNEVSQIAPGQTGNIADYNYKLPDDFTGCVI AWNSNKLD SKHSGNYDYWYRLFRKSKLKPFERDISTEI YQAGNKPCCKGKGPNCYFPLQSYGFRPTYGVGHQPYR VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLT GTGVLTKSNKKFLPFQQFGRDIVDTTDAVRDPQTLEILD	BA.2.86

	ITPCS	
247	RVQPTESIVRFPNVTNLCPFHEVFNATRFASVYAWNRT RISNCVADYSVLYNFAPFFAFKCYGVSP TKLNDLCFTNV YADSFVIKGNEVSQIAPGQTGNIADYNYKLPDDFTGCVI AWNSNKLD SKHSGNYDYWYRSFRKSKLPFERDISTEI YQAGNKPC KGKGPNCYFPLQSYGFRPTYGVGHQP YR VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLT GTGVLTKSNKKFLPFQQFGRDIVDTTDAVRDPQTLEILD ITPCS	JN.1
248	RVQPTDSIVRFPNITNLCPFGEVFNATTFASVYAWN RK RISNCVADYSVLYNSTSFSTFKCYGVSP TKLNDLCFTNV YADSFVITGDEV RQIAPGQTGKIADYNYKLPDDFTGCVI AWNSKHIDAKEGGNFNYLYRLFRKANLKPFERDISTEIY QAGSKPCNGQTGLNCYYPLYRYGFYPTDGVGHQP YR VVLSFELLNAPATVCGPKKSTNLVKNKCVNFNFNGLT GTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILD ITPCS	RaTG13
249	RVQPTISIVRFPNITNLCPFGEVFNASKFASVYAWN RKR ISNCVADYSVLYNSTSFSTFKCYGVSP TKLNDLCFTNVY ADSFVVKGDEV RQIAPGQTGVIADYNYKLPDDFTGCVIA WNSVKQDAL TGGNYGYLYRLFRKSKLPFERDISTEIY QAGSTPCNGQVGLNCYYPLERYGFHPTTGVNYQPFRV VVLSFELLNGPATVCGPKLSTTLVKDKCVNFNFNGLTG TGVLTTSKKQFLPFQQFGRDISDTTDAVRDPQTLEILDIT PCS	Pangolin17
272	RVAPSKEVVRFPNITNLCPFGEVFNATTFPSVYAWERK RISNCVADYSVLYNSTSFSTFKCYGV SATKLNDLCFSNV YADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFTGCV LAWNTRNIDATQTGN YNYKYRSLRHGKL RPFERDISNV PFSPDGKPCTPPAFNCYWPLNDYGFYITNGIGYQP YRV VVLSFELLNAPATVCGPKLSTD LIKNQCVNFNFNGLTGT	WIV1

	GVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDIS PCS	
273	RVAPSKEVVRFPNITNLCPFGEVFNATTFPSVYAWERK RISNCVADYSVLYNSTSFSTFKCYGVSATKLNDLCFSNV YADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFLGCV LAWNTNSKDSSTSGNYYLYRWVRRSKLNPYERDLSN DIYSPGGQSCSAVGPNPCYNPLRPYGFFTTAGVGHQPY RVVVLSELLNAPATVCGPKLSTDLIKNQCVNFNFNGLT GTGVLTPSSKRFQPFQQFGRDVSDFTDSVRDPKTSEIL DISPCS	SHC014
274	RILPSTEVRFPNITNFCPFDKVFNATRFPNVYAWQRTK ISDCIADYTVLYNSTSFSTFKCYGVSPSKLIDLCFTSVYA DTFLIRFSEVRQIAPGETGVIADYNYKLPDDFTGCVLAW NTAQQDIGSYFYRSHRAVKLKPFERDLSSDENGVRTLS TYDFNPVPLDYQATR VVVLSELLNAPATVCGPKLST QLVKNRCVNFNFNGLRGTGVLTDSDKRFQSFQQFGRD SADFTDSVRDPQTLQILDISPCS	RmYN02
275	RVSPSTEVRFPNITNLCPFQGVFNASNFPVYAWERL RISDCVADYAVLYNSSSSSFSTFKCYGVSPTKLNDLCFSS VYADYFVVKGDDVRQIAPAQTGVIADYNYKLPDDFTGC VLAWNTNSVDSKSGNNFYRLFRHGKIKPYERDISNVL YNSAGGTCSSISQLGCYEPLKSYGFTPTVGVGYQPYRV VVLSFELLNAPATVCGPKKSTELVKNKCVNFNFNGLTG TGVLTSSSTKKFQPFQQFGRDVSDFTDSVRDPKTFEILDI SPCS	BTKY72
238	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRGVYY PDKVFRSSVLHSTQDLFLPFFSNVTWFHAISGTNGTKR FDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSL LIVNNATNVVIKVCEFQFCNDPFLGVYHKNNKSWMESE FRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFV FKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGIN	SARS-CoV-2 Alfa (B.1.1.7)

	ITRFQTLALHRSYLTPGDSSSGWTAGAAAYYVGYLQP RTFLLKYNENGITITDAVDCALDPLSETKCTLKSFTVEKGI YQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVY AWNRRKRISNCVADYSVLYNSASFSTFKCYGVSP TKLND LCFTNVYADSFVIRGDEV RQIAPGQTGKIADYNYKLPDD FTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFER DSTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTYGVG YQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFN FNGLTGTGVLTESNKKFLPFQQFGRDIDDTTDAVRDPQ TLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTE VPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVN NSYECDIPIGAGICASYQTQTNSHRRARSVASQSIIAYT MSLGAENSVAYSNNNSIAIPINFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQD KNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSK RSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQK FNGLTVLPPLLTDemiaQYTSALLAGTITSGWTFGAGAA LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN SAIG KIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNF GAISSVLNDILARLDKVEAEVQIDRLITGRLQSLQTYVTQ QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH LMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICH DGK AHFPREGVFVSNGTHWFVTQRNFYEPQIITHTNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQEL GKYEQYIKWPWYIWLGFIAGLIAVMVTIMLCCMTSCCS CLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
239	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRGVYY PDKVFRSSVLHSTQDLFLPFFSNVTW FHAIHVSGTNGT KRFANPVL PFNDGVYFASTEKSNIIRGWIFGTTLDSKTQ SL LIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWM	SARS-CoV-2 Beta (B.1.351)

	ESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLR EFVFKNIDGYFKIYSKHTPINLVRGLPQGFSALEPLVDLP IGINITRFQTLHRSYLTPGDSSSGWTAGAAAYYVGYLQP RTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGI YQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVY AWNRRKRISNCVADYSVLYNSASFSTFKCYGVSP TKLND LCFTNVYADSFVIRGDEV RQIAPGQTGNIADYNYKLPDD FTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFR DISTEIYQAGSTPCNGVKGFNCYFPLQSYGFQPTYGVG YQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFN FNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQ TLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTE VPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVN NSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYT MSLGVENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSV DCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQ DKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPS KRSFIEDLLFNKVT LADAGFIKQYGDCLGDIAARDLICAQ KFNGLTVLPPLLTDemiaQYTSALLAGTITSGWTFGAGA ALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN SAI GKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSN FGAISSVLNDILSR LDKVEAEVQIDRLITGRLQSLQTYVT QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGY HLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICH DG KAHFPREGVFVSNGTHWFVTQRNFYEPQIITDNTFVS GNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSP DVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE LGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCC SCLKGCCSCGSCCKFDEDDSEPV LKGVKLHYT	
240	MFVFLVLLPLVSSQCVNFTNRTQLPSAYTNSFTRGVYY PDKVFRSSVLHSTQDLFLPFFSNVTW FHAIHVSGTNGT	SARS-CoV-2 Gamma (P.1)

KRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQ
 SLLIVNNATNVVIKVCEFQFCNYPFLGVYYHKNNKSWM
 ESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLS
 EFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP
 IGINITRFQTLALHRSYLT PGDSSSGWTAGAAAYYVG
 LQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTV
 EKGIIQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRF
 ASVYAWNRRKRISNCVADYSVLYNSASFSTFKCYGVSP
 T KLNDLCFTNVYADSFVIRGDEV RQIAPGQTGTIADYNYK
 LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK
 PFERDISTEIIYQAGSTPCNGVKGFNCYFPLQSYGFQPT
 YGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKC
 VNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV
 RDPQTL EILDITPCSFGGVSVITPGTNTSNQVAVLYQGV
 NCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA
 EYVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSI
 IAYTMSLGAENSVAYSNNNSIAIPTNFTISVTTEILPVSM
 TK TSV DCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAV
 EQDKNTQE VFAQVKQIYKTPPIKDFGGFNFSQILPDPSK
 PSKRSFIEDLLFNKVT LADAGFIKQYGDCLGDIAARDLIC
 AQKFNGLT VLPPLL TDEMIAQYTSALLAGTITSGWTFGA
 GAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN
 SAIGKIQDSL SSTASALGKLQDVVNQNAQALNTLVKQLS
 SNFGAISSVLNDILSR LDKVEAEVQIDRLITGRLQSLQTY
 VTQQLIRAAEIRASANLAAIKMSECVLGQSKRVDFCGKG
 YHLM SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICH
 D GKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFV
 SGNC DVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTS
 PDVDLGD ISGINASFVNIQKEIDRLNEVAKNLNESLIDLQ
 ELGKYE QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSC
 CSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

241	MFVFLVLLPLVSIQCVNLTTRTQLPPAYTNSFTRGVYYP DKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTK RFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQS LLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSCME SEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLRE FVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPI GINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYVGYL QPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVE KGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFA SVYAWNRRKRISNCVADYSVLVNSASFSTFKCYGVSPTK LNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKL PDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPT NGVGYPYRVVLSFELLHAPATVCGPKKSTNLVKNKC VNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAV RDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGV NCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGA EHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSI IAYTMSLGAENSVAYSNNNSIAIPTNFTISVTTEILPVSMTK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAV EQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSK PSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLIC AQKFNGLTVLPPLLDEMIAQYTSALLAGTITSGWTFGA GAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN SAIGKIQDSLSSSTASALGKLQDVVNQNAQALNTLVKQLS SNFGAISSVLNDILSRDKVEAEVQIDRLITGRLQSLQTY VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGK GYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICH DGKAHFPREGVVFVSNGTHWFVTQRNFYEPQIITDNTF VSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHT SPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL	SARS-CoV-2 Épsilon (B.1.427)
-----	---	------------------------------------

	QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTS CCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
242	MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYY PDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGT KRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQ SLLIVNNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWM ESGVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREF VFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGI NITRFQTLALHRSYLTPGDSSSGWTAGAAAYYVGYLQ PRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEK GIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFAS VYAWNRRKRISNCVADYSVLVNSASFSTFKCYGVSPTKL NDLCFTNVYADSFVIRGDEVQRQIAPGQGTGKIADYNYKLP DDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLKP FERDISTEIQAGSKPCNGVEGFNCYFPLQSYGFQPTN GVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCV NFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVR DPQTLLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVN CTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAE HVNNSECDIPIGAGICASYQTQTNSRRRRARSVASQSII AYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMK TSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAV EQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSK PSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLIC AQKFNGLTVLPPLLTDemiaQYTSALLAGTITSGWTFGA GAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN SAIGKIQDSLSTASALGKLQNVVNQNAQALNTLVKQLS SNFGAISSVLNDILSRDKVEAEVQIDRLITGRLQSLQTY VTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGK GYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICH DGKAHFPREGVFSNGTHWFVTQRNFYEPQIITDNTF	SARS-CoV-2 Delta (B.1.617.2) pico completo

	VSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHT SPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDL QELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTS CCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
243	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRGVYY PDKVFRSSVLHSTQDLFLPFFSNVTWFHVISGTNGTKR FDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLDSKTQSLL IVNNATNVVIKVCEFQFCNDPFLDHKNNKSWMESEFRV YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNI DGYFKIYSKHTPIIVREPEDLPQGFSALEPLVDLPIGINIT RFQTLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPR TFLLKYNENGTITDAVDCALDPLSETKCTLSFTVEKGIY QTSNFRVQPTESIVRFPNITNLCPFDEVFNATRFASVYA WNRKRISNCVADYSVLNLAFFFTFKCYGVSPTKLNDL CFTNVYADSFVIRGDEVQRQIAPGQTGNIADYNYKLPDDF TGCVIAWNSNKLDSKVSGNLYLYRLFRKSNLKPFERD ISTEIYQAGNKPCNGVAGFNCYFPLRSYSFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNF NGLKGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEV PVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAIEYVNN SYECDIPIGAGICASYQTQTKSHRRARSVASQSIIAYTM SLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLKRALTGIAVEQD KNTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPSK RSFIEDLLFNKVTADAGFIKQYGDCLGDIAARDLICAQK FKGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAA LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIG KIQDSLSTASALGKLQDVVNHNAQALNTLVKQLSSKF GAISSVLNDIFSRLDKVEAEVQIDRLITGRLQSLQTYVTQ QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH	SARS-CoV-2 ómicron (BA1, anteriorment e B.1.1.529) pico completo

	LMSFPQSAPHGVVFLHVTYVPAQEKNF T TAPAICH DGK AHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG NCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQEL GKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCS CLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
244	MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKV FRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFD NPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV NNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESEF RVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVF KNIDGYFKIYSKHTPINLGRDLPQGFSALEPLVDLPIGINI TRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQP RTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGI YQTSNFRVQPTESIVRFPNITNLCPFDEVFNATRFASVY AWNRRKRISNCVADYSVLNFAFFAFKCYGVSP TKLND LCFTNVYADSFVIRGNEVSQIAPGQTGNIADYNYKLPDD FTGCVIAWNSNKLDSKVGGNYNYLYRLFRKSNLKPFER DISTEIYQAGNKPCNGVAGFNCYFPLRSYGFRPTYGVG HQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFN FNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQ TLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTE VPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEYVN NSYECDIPIGAGICASYQTQTKSHRRARSVASQSIIAYT MSLGAENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSV DCTMYICGDSTECSNLLLQYGSFCTQLKRALTGIAVEQ DKNTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPS KRSFIEDLLFNKVTADAGFIKQYGDCLGDIAARDLICAQ KFNGLTVLPPLLTDemiaQYTSALLAGTITSGWTFGAGA ALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAI GKIQDSLSTASALGKLQDVVNHNAAQALNTLVKQLSSK	SARS-CoV-2 pico completo

	FGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVT QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGY HLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDG KAHFPREGVFVSNNGTHWFVTQRNFYEPQIITTDNTFVS GNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSP DVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE LGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCC SCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
254	MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKV FRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFD NPALPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV NNATNVVIKVCEFQFCNDPFLDVYQKNNKSWMESEFR VYSSANNCTFEYVSQPFLMDLEGKEGNFKNLREFVFKN IDGYFKIYSKHTPINLERDLPQGFSALEPLVDLPIGINITR FQTLALHRSYLTPVDSSSGWTAGAAAYYVGYLQPRTF LLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQ TSNFRVQPTESIVRFPNITNLCPFHEVFNATTFASVYAW NRKRISNCVADYSVIYNFAPFFAFKCYGVSP TKLNDLCF TNVYADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTG CVAWNSNKLDSKPSGNYNYLYRLFRKSKLKPFERDIST EIYQAGNKPCNGVAGPNCYSPLQSYGFRPTYGVGHQP YRVVVL SFELLHAPATVCGPKKSTNLVKNKCVNFNFNG LTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEI LDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPV AIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEYVNNSY ECDIPIGAGICASYQTQTKSHRRARSVASQSIIAYTMSL GAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCT MYICGDSTECSNLLLQYGSFCTQLKRALTGIAVEQDKN TQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPSKRS FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFN GLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQI	XBB.1.5 pico completo

	PFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQ DSLSSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAI SSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLI RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGDKAH FPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNC DVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVD LGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGK YEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCL KGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
255	MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKV FRSSVLHSTHDLFLPFFSNVTWFHAIHVSGTNGTKRFD NPALPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIV NNATNVVIKVCEFQFCNDPFLDVYQKNNKSWMESEFR VYSSANNCTFEYVSQPFLMDLEGKEGNFKNLREFVFKN IDGYFKIYSKHTPINLERDLPQGFSALEPLVDLPIGINITR FQTLALHRSYLTVPDSSSGWTAGAAAYYVGYLQPRTF LLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQ TSNFRVQPTESIVRFPNITNLCPFHEVFNATTFASVYAW NRKRISNCVADYSVIYNFAPFFAFKCYGVSP TKLNDLCF TNVYADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTG CVIAWNSNKLDSKPSGNYNYLYRLLRKS KLKPFERDIST EIYQAGNKPCNGVAGPNCYSPLQSYGFRPTYGVGHQP YRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNG LTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEI LDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPV AIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEYVNNSY ECDIPIGAGICASYQTQTKSHRRARSVASQSIIAYTMSL GAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCT MYICGDSTECNLLLQYGSFCTQLKRALTGIAVEQDKN TQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPSKRS	EG.5.1 pico completo

	FIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFN GLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQI PFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQ DSLSSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAI SSVLNDILSRDKVEAEVQIDRLITGRLQSLQTYVTQQLI RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHGDKAH FPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNC DVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVD LGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGK YEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCL KGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
256	MFVFLVLLPLVSSQCVMPFLNLITTTQSYTNSFTRGVYY PDKVFRSSVLHLTQDLFLPFFSNVTWFHAISGTNGTKR FDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSL LIVNNATNVFIKVCEFQFCNDPFLDVYHKNNKSWMESE SGVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFV FKNIDGYFKIYSKHTPIIGRDFPQGFSALEPLVDLPIGINIT RFQTLALNRSYLT PGDSSSGWTAGAADYYVGYLQPR TFLLKYNENGTITDAVDCALDPLSETKCTLKSTVEKGIY QTSNFRVQPTESIVRFPNVTNLCPFHEVFNATRFASVY AWNRTNISNCVADYSVLYNFAPFFAFKCYGVSP TKLND LCFTNVYADSFVIKGNEVSQIAPGQTGNIADYNYKLPDD FTGCVIAWNSNKLDSKHSGNYDYWYRLFRKSKLKPFE RDISTEIQAGNKPCKGKGPNCYFPLQSYGFRPTYGVG HQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFN FNGLTGTGVLTKSNKKFLPFQQFGRDIVDTTDAVRDPQ TLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTE VSVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEYVN NSYECDIPIGAGICASYQTQTKSRRRARSVASQSIIAYT MSLGAENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSV	BA.2.86 pico completo

	DCTMYICGDSTECSNLLLQYGSFCTQLKRALTGIAVEQ DKNTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPS KRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQ KFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGA ALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN SAI GKIQDSLSTASALGKLQDVVNHNAQALNTLVKQLSSK FGAISSVLNDILSRDKVEAEVQIDRLITGRLQSLQTYVT QQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGY HLMSFPQSAPHGVVFLHVTVPAQEKNFTTAPAICH D KAHFPREGVFVSNGTHWFVTQRNFYEPQIITDNTFVS GNCDVVIGIVNNTVYDPLQLELDSFKEELDKYFKNHTSP DVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQE LGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCC SCLKGCCSCGSCCKFDEDDSEPVLKGVLHYT	
257	MFVFLVLLPLVSSQCVNLITTTQSYTNSFTRGVYYPDKV FRSSVLHLTQDLFLPFFSNVTWFHAISGTNGTKRFDNP VLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNN ATNVFIKVCEFQFCNDPFLDVYHKNNKSWMESESGVY SSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNID GYFKIYSKHTPIIGRDFPQGFSALEPLVDLPIGINITRFQT LLALNRSYLTPGDSSSGWTAGAADYYVGYLQPRTFLLK YNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSN FRVQPTESIVRFPNVTNLCPFHEVFNATRFASVYAWNR TRISNCVADYSVLNFAFFAFKCYGVSPTKLNDLCFTN VYADSFVIKGNEVSQIAPGQTGNIADYNYKLPDDFTGCV IAWNSNKLDSKHSGNYDYWYRSFRKSKLKPFERDISTE IYQAGNKPCCKGKGPNCYFPLQSYGFRPTYGVGHQP PYR VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLT GTGVLTKSNKKFLPFQQFGRDIVDTTDAVRDPQTLEILD ITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVSVAIH ADQLTPTWRVYSTGSNVFQTRAGCLIGAAYVNNSEYEC	JN.1 pico completo

	DIPIGAGICASYQTQTKSRRRARSVASQSIIAYTMSLGAE NSVAYSNNNSIAIPTNFTISVTTEILPVSMKTSVDCTMYIC GDSTECNLLLQYGSFCTQLKRALTGIAVEQDKNTQEV FAQVKQIYKTPPIKYFGGFNFSQILPDPSKPSKRSFIEDL LFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTV LPPLLTDemiaQYTSALLAGTITSGWTFGAGAALQIPFA MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSL FSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISSVL NDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAE IRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQ SAPHGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPRE GVFSVNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVI GIVNNTVYDPLQLELDSFKEELDKYFKNHTSPDVDLGI SGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEYQ IKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCCLKGCC SCGSCCKFDEDDSEPVKGVKLHYT	
258	MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMR GVYYPDEIFRSDTLYLTLQDLFLPFYSNVTGFH TINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMN NKSQSVIIINNSTNVVIRACNFELCDNPFFAV SKPMGTQTHMIFDNAFNCTFEYISDAFSLDVSEKSGN FKHLREFVFKNKDGFLYVYKGYQPIDVVRDLP SGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSA YFVGYLKPTTFMLKYDENGITDAVDCSQ NPLAELKCSVKSFEIDKGIYQTSNFRVVPBGDVVRFPNI TNLCPFGEVFNATKFPSVYAWERKKISNCVA DYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFV VKGDDVRQIAPGQTGVIADYNYKLPDFMGCV LAWNTRNIDATSTGNVNYKYRYLRHGKLRPFERDISNV PFSPDGKPCTPPALNCYWPLNDYGFYTTTGIG YQPYRVVLSFELLNAPATVCGPKLSTDLIKNQCVNFNF	SARS-CoV-1 pico completo

	NGLTGTGVLTPSSKRFPFQQFGRDVSDFD SVRDPKTSEILDISPCSFGGVSVITPGTNASSEVAVLYQ DVNCTDVSTAIHADQLTPAWRIYSTGNNVFQ TQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLRSTS QKSIVAYTMSLGADSSIAYSNNTIAIPTNF SISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYG SFCTQLNRALSGIAAEQDRNTREVFAQVKQM YKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTL ADAGFMKQYGECLGDINARDLICAQKFNGL TVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIP FAMQMAYRFNGIGVTQNVLYENQKQIANQFN KAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLS SNFGAISSVLNDILSRDKVEAEVQIDRLIT GRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQ SKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYV PSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQ RNFFSPQIITDNTFVSGNCDVVIGIINNTVY DPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVV NIQKEIDRLNEVAKNLNESLIDLQELGKYEQ YIKWPWYVWLGFIAGLIAIVMTILLCCMTSCCCLKGA CSCGSCCKFDEDDSEPVLKGVKLHYT	
259	MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSSTRGVYY PDKVFRSSVLHLTQDLFLPFFSNVTWFHAIHVSGTNGIK RFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQS LLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWME SEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLRE FVFKNIDGYFKIYSKHTPINLVRDLPPGFSALEPLVDLPI GINITRFQTLALHRSYLT PGDSSSGWTAGAAAYYVGYL QPRTFLLKYNENGITDAVDCALDPLSETKCTLKSFTVE KGIYQTSNFRVQPTDSIVRFPNITNLCPFGEVFNATTFA SVYAWNRRKRISNCVADYSVLYNSTSFSTFKCYGVSPTK	RaTG13 pico completo

	LNDLCFTNVYADSFVITGDEVQRQIAPGQTGKIADYNYKL PDDFTGCVIAWNSKHIDAKEGGNFNYLYRLFRKANLKP FERDISTEIYQAGSKPCNGQTGLNCYYPYRYGFYPTD GVGHQPYRVVLSFELLNAPATVCGPKKSTNLVKNKCV NFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVR DPQTLILDITPCSFGGVSVITPGTNASNQVAVLYQDVN CTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAE HVNNSYECDIPIGAGICASYQTQTNSRSVASQSIIAYTM SLGAENSVAYSNNNSIAIPTNFTISVTTEILPVSMTKTSVD CTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQD KNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSK RSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQK FNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAA LQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIG KIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNF GAISSVLNDILSRDLKVEAEVQIDRLITGRLQSLQTYVTQ QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH LMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGK AHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSG SCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQEL GKYEQYIKWPWYIWLGFIAGLIAIMVTIMLCCMTSCCSC LKGCCSCGSCCKFDEDDSEPVLKGVKLHYT	
260	MFVFLFVLPLVSSQCVNLTTTRTGIPPGYTNSSTRGVYYP DKVFRSSILHLTQDLFLPFFSNVTWFNTINYQGGFKKFD NPVLPFNDGVYFASTEKSNIIRGWIFGTTL DARTQSLLIV NNATNVVIKVCEFQFCTDPFLGVYYHNNNKTWVENEFR VYSSANNCTFEYISQPFLMDLEGKQGNFKNLREFVFKN VDGYFKIYSKHTPIDLVRDLPRGFAALEPLVDLPIGINITR FQTLALHRSYLTPGKLESGWTTGAAAYYVGYLQQRTF LLSYNQNGTITDAVDCSLDPLSETKCTLKSLTVEKGIYQ	Pangolin17 pico completo

	TSNFRVQPTISIVRFPNITNLCPFGEVFNASKFASVYAW NRKRISNCVADYSVLYNSTSFSTFKCYGVSP TKLNDLC FTNVYADSFVVKGDEV RQIAPGQTGVIADYNYKLPDDF TGCVIAWNSVKQDALTGGNYGYLYRLFRKSKLKPFERD ISTEIYQAGSTPCNGQVGLNCYYPLERYGFHPTTGVNY QPFRVVVLSFELLNGPATVCGPKLSTTLVKDKCVNFN NGLTGTGVLTTSKKQFLPFQQFGRDISDTTDAVRDPQT LEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEV PMAIHAEQLTPAWRVYSAGANVFQTRAGCLVGAEHVN NSYECDIPVGAGICASYHSMSSFRSVNQRSIIAYTMSLG AENSVAYSNN SIAIPTNFTISVTTEILPVSMTKTSVDCTM YICGDSIECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQ EVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIE DLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGL TVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPF AMQMAYRFNGIGVTQNVLYENQKLIANQFN SAIGKIQD SLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAIS SVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIR AAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM SF PQSAPHGVVFLHVTYVPAQEKNFTTAPAICHEGKAHFP REGVFVSNGTHWFITQRNFYEPQIITDNTFVSGSCDV VIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE QYIKWPWYIWLGFIAGLIAIIMVTIMLCCMTSCC SCLKGC CSCGSCCKFDEDDSEPVLKGVKLHYT	
261	MKLLVLVFATLVSSYTIEKCLDFDDRTPPANTQFLSSHR GVYYPDDIFRSNVLHLVQDHFLPFDSNVTRFITFGLNFD NPIIPFKDGIYFAATEKSNVIRGWVFGSTMNNKSQS VIIM NNSTNLVIRACNFELCDNPFFVLKSNNTQIPSYIFNNAF NCTFEYVSKDFNLDLGEKPGNFKDLREFVFRNKDGFLH VYSGYQPISAASGLPTGFNALKPIFKLPLGINITNFR TLLT	WIV1 pico completo

	AFPPRPDYWGTSAAAAYFVGYLKPTTFMLKYDENGITD AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVAPSKEV VRFPNITNLCPFGEVFNATTFPSVYAWERKRISNCVAD YSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVV KGDDVRQIAPGQTGVIADYNYKLPDDFTGCVLAWNTR NIDATQTGNYNKYRSLRHGKLRPFERDISNVPFSPDG KPCTPPAFNCYWPLNDYGFYITNGIGYQPYRVVVLSE LLNAPATVCGPKLSTDLIKNCVNFNFNGLTGTGVLTPS SKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCSF GVSVITPGTNTSSEVAVLYQDVNCTDVPVAIHADQLTPS WRVHSTGNNVFQTQAGCLIGAEHVDTSECDIPGAGI CASYHTVSSLRSTSQKSIVAYTMSLGADSSIAYSNNTIAI PTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLL LQYGSFCTQLNRALSGIAVEQDRNTREVFAQVKQMYK TPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVT LADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPL LTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQ MAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQES LTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGA ISSVLNDILSRDKVEAEVQIDRLITGRLQSLQTYV TQQLIRAAEIRASANLAATKMSECVLGQSKRVDF CGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFT TAPAICHEGKAYFPREGVVFVNGTSWFITQRNFF SPQIITTDNTFVSGSCDVVIGIINNTVYDPLQPE LDSFKEELDKYFKNHTSPDVDLGDISGINASVVN IQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKW PWYVWLGFIAGLIAIVMTILLCCMTSCCSCLKGAC SCGSCCKFDEDDSEPVLKGVKLHYT	
262	MKLLVLVFATLVSSYTIEKCLDFDDRTPPANTQFL SSHRGVYYPDIFRSNVLHLVQDHFLPFDSNVTRF ITFGLNFDNPIIPFRDGIYFAATEKSNVIRGWV FVGSTMNNKSQSIVIIMNNSTNLVIRACNFELC DNPFFVVLKSNNTQIPSYIFNNAF	SHC014 pico completo

	NCTFEYVSKDFNLDLGEKPGNFKDLREFVFRNKDGFLH VYSGYQPISAASGLPTGFNALKPIFKLPLGINITNFRTLLT AFPPRPDYWGTSAAAYFVGYLKPTTFMLKYDENGITD AVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVAPSKEV VRFPNITNLCPFGEVFNATTFPSVYAWERKRISNCVAD YSVLYNSTSFSTFKCYGVSATKLNDLCFSNVYADSFVV KGDDVRQIAPGQTGVIADYNYKLPDDFLGCVLAWNTNS KDSSTSGNYYLYRWVRRSKLNPYERDLSNDIYSPGG QSCSAVGPNCYNPLRPYGFFTTAGVGHQPYRVVLSF ELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTP SSKRFQPFQQFGRDVSDFDTSVRDPKTSEILDSPCSF GGVSVITPGTNTSSEVAVLYQDVNCTDVPVAIHADQLT PSWRVYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGA GICASYHTVSSLRSTSQKSIVAYTMSLGADSSIAYSNNTI AIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECAN LLLQYGSFCTQLNRALSGIAVEQDRNTREVFAQVKQMY KTPTLKDFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLA DAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLTD DMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYR FNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTAL GKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRL DKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANL AATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVV FLHVTYVPSQERNFTTAPAICHEGKAYFPREGVVFVNG TSWFITQRNFFSPQIITTDNTFVSGSCDVVIGIINNTVYD PLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVN IQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVW LGFIAGLIAIVMTILLCCMTSCCSCLKGACSCGSCCKF DEDDSEPVKGVKLHYT	
263	MFILLLIGYTAATTCVTGPTIENKQNVSSLMRGVYYPDDI YRSNVNVLFTVPFLRFNSTLTWYNFWNQAYTSRVMEF	RmYN02 pico completo

	GDGIYFSTVDKSNAVRGWIFGTTLDNTTQSALLINNGSL VTIQVCYFQFCANPAFLVAGGQQTSAAVYISSHNCTYS EVLSHISLDFTEDTGSFKRLREFVFKNNDGFLHIYGAYQ PHVIDIGATPSLPSTFKPLIPLWKLPLGLNITNFKVVLTFR SNSQPLQANFAVGSLKLTTIMLSFDVNGTIDKAVDCSSD PLSELKCTLKSFNVSKGIYPTSNFRILPSTEVVRFPNITN FCPFDKVFNATRFPNVYAWQRTKISDCIADYTVLYNSTS FSTFKCYGVSPSKLIDLCFTSVYADTFLIRFSEVRQIAPG ETGVIADYNYKLPPDDFTGCVLAWNTAQQDIGSYFYRSH RAVKLKPFERDLSSDENGVRTLSTYDFNPNVPLDYQAT RVVVLSEFELNAPATVCGPKLSTQLVKNRCVNFNFNGL RGTGVLTDSDKRFQSFQQFGRDSADFTDSVRDPQTLQ ILDISPCSFGGVSVITPGTQTSSKVAVLYQDVNCTDVPT ALGLDQISAAWRVYAIGNDFQTQAGCLVGAEHTNISY ECDIPIGAGVCASYNSPAARVGTNSIIAYAMSIGAESSIA YSNNIAIPTNFSIQVTTEVLPVSVSKTSVDCTMYICGDS QECNKLLLQYGSFCAQLNRALSGVAVEQDKNTEAVFA QVKQIYKAPVIKDLGGFNFSQLLPDPTKPSQRSFIEDLLF DKVTLSDAGFIKQYGDCLGDIAARDLICAQKFNGLTVLP PLLTDEMIAAYTSALVSGTATAGWTFGIGAALQVPFAM QMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSL STASALGKLQDVVNQNAQALNTLVKQLSSNFGAIVSVL NDILSRDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAE IRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQ SAPHGVVFLHVTYVPAQEKNFTTAPAICYDGKAHFPRE GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGSCDVVI GIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDI SGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEYQ IKWPWYIWLGFIAGLIAIMVTIMLCCMTSCCSCCLKGCCS CGSCCKFDEDDSEPVLKGVKLHYT	
264	MKFFILLSLLSFTTAQEGCGILSNKSNPALTQYFSSRRG	BTKY72 pico

FYYFDDTFRSSVRVLTGGYFLPFNSNLTYGSSRNSVTG RLIQFDNPNIPIFKDGLYFAATERSNVIRGWIFGSTLDNTT QSAVLFNNGTHIVINVCNFYFCQDPMLAVANGSHFKSW VFLNATNCTYNRVHGFEDPSPNTGSFIHLREHVFRNVD GFLYVYHNYERVDVYDNFPQGFSVLKPIFKLPFGLNITQ FKVIMTLFSPTTSSFNADASVYFVGHLKPLTMLAEFDEN GTITDAVDCSQDPLSELKCTTKSLTVEKGIYQTSNFRVS PSTEVVRFPNITNLCPFGQVFNASNFPSVYAWERLRIS DCVADYAVLYNSSSSSFSTFKCYGVSPTKLNDLCFSSVY ADYFVVKGDDVRQIAPAQTGVIADYNYKLPDDFTGCVL AWNTNSVDSKSGNMFYRLFRHGKIKPYERDISNVLYN SAGGTCSSISQLGCYEPLKSYGFTPTVGVGYQPYRVVV LSFELLNAPATVCGPKKSTELVKNKCVNFNFNGLTGTG VLTSSSTKKFQPFQQFGRDVSDFTDSVRDPKTFEILDISP CSYGGVSVITPGTNTSKAVAVLYQDVNCTDVPTMIHVE QVSSDWRVYAFNSYGNMFQTQAGCLVGAIYENTTYEC DIPIGAGICAKFGSDKIRMGQESIVAYTMSIGEDQSIAYS NNIIAIPTNFSISVTTEVLVPSMTKTSVDCNMYICGDSTE CSNLLLQYGSFCTQLNRALSGIAVEQDRNTRDVFAQTK SIYKTPNIKDFGGFNFSQILPDPKKLSYRSFIEDLLYNKV TLSDPGFMKQYGDCLGGINARDLICAQKFNGLTVLPPLL TDDMIAAYTAALISGTATAGYTFGAGAALQIPFAMQMAY RFNGIGVTQNVLYENQKQIANQFNNAISKIQDSLTTTSA ALGKLQDVINQNAVALNTLVKQLSSNFGAISSVLNDILSR LDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASAN LAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGV VFLHVITYVPSQQQNFTTAPAICHNGKAYFPREGVFMN GTHWFITQRNFYSPQVITTDNTFESGSCDVVIGIINNTVY DPLQPELESFKQELDKYFKNHTSPDVDFGDISGINASVV DIKKEIAHLNEIAKNLNESSLIDLQELGKYEQYVKWPWYV WLG FVAGIVAILMSVIMLCCMTNCCSCFKGMCSCGSCC	completo
---	----------

	KFDEDHSEPVSGVKLHYT	
--	--------------------	--

[0147] En algunas modalidades, el SD1 comprende al menos una sustitución de un residuo de aminoácido no polar expuesto en la superficie con un residuo de aminoácido polar. En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más o tres aminoácidos en las posiciones 560, 562 o 569 con un residuo polar, en donde las sustituciones son relativas a una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende una sustitución de una o más, dos o más o tres de las sustituciones de aminoácidos L560Q, F562T, F562Y o I569S con respecto a una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más o tres o más aminoácidos en las posiciones 544, 546, 560, 562, 564, 569 o 582 con respecto a una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones N544L, N544M, N544Q, L546V, L560Q, F562T, F562Y, Q564C, Q564L, Q564N, Q564W, I569S o L582S en relación con una secuencia de referencia según la SEQ ID NO: 1

[0148] En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones 338, 358, 363, 365, 392 o 395 en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones F338L, I358F, A363L, Y365F, Y365M, Y365W, F392W o V395I en relación con una secuencia de referencia según la SEQ ID NO: 1. En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos que comprenden que se muestran en la Tabla 2.

Tabla 2

F338L /Y365W
F392W
Y365F/V395I
Y365F/F392W/V395I
F338L/A363L/Y365M
I358F/Y365F/V395I

[0149] En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones 329, 348, 350, 367, 375, 402, 407, 410, 418, 429, 433, 435, 452, 464, 510, 512, 514, 517, 518, 519, 520, 522, 527 o 528 con respecto a una secuencia de referencia según la SEQ ID NO: 1.

[0150] En algunas modalidades, el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones F329K, F329R, F329Y, A348P, V350L, V367F, F375Y, I402V, V407L, I410F, I418V, F429W, V433I, A435I, A435V, L452R, F464Y, P512Q, V510I, V512F, V512I, S514T, L517A, L517D, L517S, L517T, L518N, L518Q, L518V, H519D, H519G, H519R, H519S, H519T, A520C, A520D, A520G, A520H, P521A, P521D, P521N, P521Q, P521S, A522G, A522I, P527N, K528I, K528Q, K528T, L518G o L518S en relación con una secuencia de referencia según la SEQ ID NO: 1.

[0151] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 348, 402, 464, 514, 520, 526 y 527, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

[0152] En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que se muestran en **la Tabla 3**.

Tabla 3.

A348P/Y365F/V395I/L560Q/F562Y/I569S
Y365F/V395I/I402V/L560Q/F562Y/I569S
Y365F/V395I/S514T/L560Q/F562Y/I569S
Y365F/V395I/I402V/F464Y/L560Q/F562Y/I569S
Y365F/V395I/G526S/P527N/L560Q/F562Y/I569S

[0153] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 348, 402, 464, 514, 520, 526 y 527, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

[0154] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de

aminoácidos en las posiciones 329, 517, 519, 520, y 544, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

[0155] En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que se muestran en la **Tabla 4**.

5 **Tabla 4**

Y365F/V395I/L560Q/F562Y/I569S
Y365F/V395I/F329Y/L560Q/F562Y/I569S
Y365F/V395I/L517T/H519S/L560Q/F562Y/I569S
Y365F/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S
Y365F/V395I/L517T/H519S/N544L/L560Q/F562Y/I569S
Y365F/V395I/F329Y/L517T/H519S/L560Q/F562Y/I569S

[0156] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 392, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 329, 517, 519, 520, y 544, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

10

[0157] En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que se muestran en la **Tabla 5**.

Tabla 5

Y365F/VF392W/395I/L560Q/F562Y/I569S
Y365F/VF392W/V395I/F329Y/L560Q/F562Y/I569S
Y365F/VF392W/V395I/L517T/H519S/L560Q/F562Y/I569S
Y365F/VF392W/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S
Y365F/VF392W/V395I/L517T/H519S/N544L/L560Q/F562Y/I569S
Y365F/VF392W/V395I/ F329Y/L517T/H519S/L560Q/F562Y/I569S

15 **[0158]** En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519, 520, 560, 562, 564, y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 521, 544, y 546, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

[0159] En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que

se muestran en la **Tabla 6**.

Tabla 6.

Y365F/V395I/L517T/H519G/A520C/Q564C/L560Q/F562Y/I569S
Y365F/V395I/L517S/H519R/A520H/P521Q/N544L/L546V/Q564W/F562T/L560Q/F562Y/I569S

[0160] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519, y 520; y una o más, o dos o más sustituciones de aminoácidos en las posiciones 518 y 392, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

[0161] En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que se muestran en la **Tabla 7**.

Tabla 7.

Y365F/V395I/L517T/H519S/A520G
Y365F/V395I/L517T/L518G/H519D/A520G
Y365F/V395I/L517T/L518S/H519S/A520G
Y365F/F392W/V395I/L517T/H519S/A520G
Y365F/F392W/V395I/L517T/L518G/H519D/A520G
Y365F/F392W/V395I/L517T/L518S/H519S/A520G

[0162] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 329, 365, 395, 517, 519, 520, 560, 562, y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 392, 544, y 564, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

[0163] En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que se muestran en la **Tabla 8**.

Tabla 8.

Y365F/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S/F329Y
Y365F/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S/F329Y
Y365F/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S/F329Y

Y365F/F392W/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S/F329Y
Y365F/F392W/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S/F329Y
Y365F/F392W/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S/F329Y
Y365F/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S
Y365F/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S
Y365F/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S
Y365F/F392W/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S
Y365F/F392W/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S
Y365F/F392W/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S

[0164] En algunas modalidades, el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519, 520, 521, 504, 560, 562, 569, y 329; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 392, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

[0165] En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que se muestran en la **Tabla 9**.

Tabla 9.

Y365F/V395I/L517T/H519G/P521A/A520C/Q564C/L560Q/F562Y/I569S/F329Y
Y365F/V395I/L517T/H519G/P521N/A520C/Q564C/L560Q/F562Y/I569S/F329Y
Y365F/F392W/V395I/L517T/H519G/P521A/A520C/Q564C/L560Q/F562Y/I569S/F329Y
Y365F/F392W/V395I/L517T/H519G/P521N/A520C/Q564C/L560Q/F562Y/I569S/F329Y

[0166] En algunas modalidades, el polipéptido comprende una pluralidad de sustituciones de aminoácidos en las posiciones 329, 365, 395, 526, 527, 528, 560, 562 y 569.

[0167] En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos que se muestran en la **Tabla 10**.

Tabla 10.

F329Y/Y365F/V395I/G526S/P527N/K528T/L560Q/F562Y/I569S
F329Y/Y365F/V395I/G526S/P527N/K528I/L560Q/F562Y/I569S

[0168] En algunas modalidades, el polipéptido comprende una secuencia polipeptídica al menos

80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 2-10.

5 **[0169]** En algunas modalidades, el polipéptido comprende una secuencia polipeptídica al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a una o más de las SEQ ID NO: 11-19.

10 **[0170]** En algunas modalidades, el polipéptido comprende una secuencia polipeptídica idéntica a una cualquiera de las SEQ ID NO: 11-19.

[0171] En algunas modalidades, el polipéptido se expresa a un nivel más alto que un polipéptido de referencia sin las sustituciones de aminoácidos, opcionalmente un polipéptido de referencia según las SEQ ID NO: 2 o 10.

15 **[0172]** En algunas modalidades, la expresión se puede detectar usando técnicas convencionales para detectar proteínas (*por ejemplo*, ELISA, transferencia Western, citometría de flujo, inmunofluorescencia, inmunohistoquímica, *etc.*).

[0173] En algunas modalidades, el polipéptido expresado es más estable térmicamente o más antigénico que las SEQ ID NO: 2 o 10.

20 **[0174]** En algunas modalidades, el polipéptido comprende una pluralidad de sustituciones de aminoácidos en las posiciones 329, 367, 383, 547, 568, 570, y 619. En algunas modalidades, el polipéptido comprende las sustituciones de aminoácidos en las posiciones F329Y, F329S, V367F, S383C, T547C, D568N, A570C y E619Q.

25 **[0175]** En algunas modalidades, la antigenicidad se mide a través de interferometría de biocapa (BLI) usando anticuerpos ACE2-Fc, S309, CR3022 inmovilizados.

30 **[0176]** Como se usa en la presente, el término "estabilidad térmica" o "termoestabilidad" de una proteína se entiende en la presente como la estabilidad de la proteína en términos de desnaturalización debido a la alteración de la temperatura y puede determinarse por la cinética de la desnaturalización de la proteína durante el calentamiento. La estabilidad térmica se puede evaluar midiendo la temperatura de fusión (T_m) o la temperatura de inicio (T_{onset}) de una composición de la invención usando cualquier técnica adecuada. La temperatura de fusión es

la temperatura a la que la primera derivada de la curva de fusión alcanza un máximo local. La temperatura de inicio es la temperatura a la que una regresión lineal del inicio se interseca con una regresión lineal de la transición de fusión.

[0177] En algunas modalidades, se puede usar una fluorimetría de barrido diferencial nano (nanoDSF) para determinar la estabilidad térmica mediante el monitoreo de la fluorescencia intrínseca del triptófano. En algunas modalidades, se puede usar la dispersión de luz estática (SLS) para revelar una temperatura de agregación.

[0178] En algunas modalidades, la estabilidad térmica se puede evaluar mediante espectroscopia analítica. Un método de espectroscopia analítica ejemplar es la espectroscopia de dicroísmo circular (CD). La espectrometría de CD mide la actividad óptica de una composición en función del aumento de la temperatura. La espectroscopia de dicroísmo circular (CD) mide las diferencias en la absorción de la luz polarizada hacia la izquierda frente a la luz polarizada hacia la derecha que surgen debido a la asimetría estructural. Una estructura desordenada o desplegada da como resultado un espectro CD muy diferente al de una estructura ordenada o plegada. El espectro CD refleja la sensibilidad de las proteínas a los efectos desnaturalizantes del aumento de la temperatura y, por lo tanto, es indicativo de la estabilidad térmica de una proteína (véase, por ejemplo, van Mierlo y Steemsma, J. Biotechnol., 79(3):281-98, 2000).

[0179] Otro método de espectroscopia analítica ejemplar para medir la estabilidad térmica es la espectroscopia de emisión de fluorescencia. Los métodos basados en fluorescencia para evaluar la estabilidad térmica controlan los cambios en la fluorescencia de los fluoróforos intrínsecos (p. ej., aminoácidos triptófano y tirosina) o fluoróforos extrínsecos (p. ej., ANS o SYPRO Orange) tras el desdoblamiento térmico (véase, por ejemplo, Niesen F. PL, Berglund H. y Vedadi M.: Vedadi M. : The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols 2007, 2:2212-21.) Otro método de espectroscopia analítica ejemplar para medir la estabilidad térmica de los anticuerpos, de la cadena ligera o pesada, es la espectroscopia de resonancia magnética nuclear (RMN) (véase, por ejemplo, van Mierlo y Steemsma, J. Biotechnol., 79(3):281 -98, 2000).

[0180] En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 30 °C. En algunas modalidades, la temperatura de fusión (T_m) es de aproximadamente 35 °C . En algunas modalidades, la

temperatura de fusión (T_m) es de aproximadamente 40 °C . En algunas modalidades, la temperatura de fusión (T_m) es de aproximadamente 45 °C . En algunas modalidades, la temperatura de fusión (T_m) es de aproximadamente 50 °C . En algunas modalidades, la temperatura de fusión (T_m) es de aproximadamente 55 °C . En algunas modalidades, la temperatura de fusión (T_m) es de aproximadamente 60 °C . En algunas modalidades, la temperatura de fusión (T_m) es de aproximadamente 65 °C . En algunas modalidades, la temperatura de fusión (T_m) es de aproximadamente 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de aproximadamente 75 °C .

[0181] En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de al menos aproximadamente 30 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 35 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 40 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 45 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 50 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 55 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 60 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 65 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de al menos aproximadamente 75 °C .

[0182] En algunas modalidades, la temperatura de fusión (T_m) es de 30 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 35 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 40 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 45 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 50 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 55 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 60 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 65 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 70 °C a 75 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 30 °C a 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 35 °C a 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 40 °C a 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 45 °C a 70 °C . En algunas

- modalidades, la temperatura de fusión (T_m) es de 50 °C a 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 55 °C a 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 60 °C a 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 65 °C a 70 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 30 °C a 65 °C .
- 5 En algunas modalidades, la temperatura de fusión (T_m) es de 35 °C a 65 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 40 °C a 65 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 45 °C a 65 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 50 °C a 65 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 55 °C a 65 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 60 °C a 65 °C .
- 10 En algunas modalidades, la temperatura de fusión (T_m) es de 35 °C a 60 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 40 °C a 60 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 45 °C a 60 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 50 °C a 60 °C . En algunas modalidades, la temperatura de fusión (T_m) es de 55 °C a 60 °C .
- 15 **[0183]**En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,1 °C . En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,2 °C . En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,3 °C . En
- 20 algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,4 °C . En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,5 °C . En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,6 °C . En
- 25 algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,7 °C . En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,8 °C . En algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 0,9 °C . En
- 30 algunas modalidades, la temperatura de fusión (T_m) de un polipéptido expresado como se describe en la presente es de aproximadamente 1,0 °C .

Enlazadores

[0184] En algunas modalidades, el polipéptido recombinante y una nanoestructura proteica pueden fusionarse genéticamente de tal manera que ambos estén presentes en un único polipéptido, denominado "proteína de fusión". El enlace unión entre el polipéptido y la nanoestructura proteica permite que el polipéptido recombinante se muestre en el exterior de la nanoestructura proteica autoensamblable.

[0185] Puede utilizarse una amplia variedad de secuencias polipeptídicas para enlazar las proteínas, o fragmentos antigénicos de estas y la nanoestructura proteica. En algunos casos, el enlazador comprende una secuencia polipeptídica que puede incluirse en la secuencia polinucleotídica codificante. Se puede utilizar cualquier polipéptido enlazador adecuado. En algunas modalidades, el enlazador impone una orientación relativa rígida de la proteína antigénica (*por ejemplo*, polipéptido de coronavirus que comprende un dominio de unión al receptor (RBD) y un subdominio 1 adyacente a RBD (SD1) de la glucoproteína Spike) o el fragmento antigénico de esta a la nanoestructura proteica. En algunas modalidades, el enlazador se une de forma flexible a la proteína antigénica (*por ejemplo*, polipéptido de Coronavirus que comprende un dominio de unión al receptor (RBD) y un subdominio 1 adyacente a RBD (SD1) de la glucoproteína Spike) o el fragmento antigénico de esta a la nanoestructura proteica. En algunas modalidades, los polipéptidos codificados pueden incluir un enlazador entre regiones. En algunas modalidades, el polipéptido es una proteína de fusión que incluye el polipéptido de coronavirus recombinante, un enlazador y el polipéptido componente de la nanoestructura de proteína. En algunas modalidades, el polipéptido es una proteína de fusión, que incluye, en el orden de extremo N a C, el polipéptido de coronavirus recombinante, un enlazador y el polipéptido componente de la nanoestructura de proteína. El enlazador puede ser un polipéptido. Puede utilizarse una amplia variedad de secuencias polipeptídicas conocidas en la técnica. En algunas modalidades, el enlazador puede comprender un enlazador Gly-Ser (*es decir*, un enlazador formado por residuos de glicina y serina) de cualquier longitud adecuada. En algunas modalidades, el enlazador Gly-Ser puede tener una longitud de 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más residuos de aminoácidos. A continuación se presentan ejemplos no limitantes de enlazadores Gly-Ser:

[0186] En la **Tabla 11** se proporcionan enlazadores polipeptídicos ilustrativos. En algunas modalidades, el enlazador polipeptídico es cualquier secuencia en la **Tabla 11**.

[0187] En algunas modalidades, el enlazador comprende entre 3 y 30 residuos de aminoácidos.

En algunas modalidades, el enlazador comprende entre 4 y 24 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 8 y 24 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 10 y 24 residuos de aminoácidos. En

5 algunas modalidades, el enlazador comprende entre 12 y 24 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 16 y 24 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 18 y 24 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 20 y 24 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 4 y 20 residuos de aminoácidos. En

10 algunas modalidades, el enlazador comprende entre 8 y 20 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 10 y 20 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 12 y 20 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 16 y 20 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende entre 8 y 18 residuos de aminoácidos. En

15 algunas modalidades, el enlazador comprende entre 12 y 16 residuos de aminoácidos.

[0188] En algunas modalidades, el enlazador comprende 3 residuos de aminoácidos. En

algunas modalidades, el enlazador comprende 4 residuos de aminoácidos. En algunas

modalidades, el enlazador comprende 7 residuos de aminoácidos. En algunas modalidades, el

enlazador comprende 8 residuos de aminoácidos. En algunas modalidades, el enlazador

20 comprende 10 residuos de aminoácidos. En algunas modalidades, el enlazador comprende 12

residuos de aminoácidos. En algunas modalidades, el enlazador comprende 16 residuos de

aminoácidos. En algunas modalidades, el enlazador comprende 18 residuos de aminoácidos.

En algunas modalidades, el enlazador comprende 19 residuos de aminoácidos. En algunas

modalidades, el enlazador comprende 20 residuos de aminoácidos. En algunas modalidades,

25 el enlazador comprende 24 residuos de aminoácidos. En algunas modalidades, el enlazador

comprende 30 residuos de aminoácidos.

[0189] En algunas modalidades, los polipéptidos codificados pueden incluir un enlazador entre

regiones. En algunas modalidades, el polipéptido es una proteína de fusión que incluye el

polipéptido del SARS-CoV-2 recombinante, un enlazador, un enlazador de extensión de

30 extremo N y el polipéptido componente de la nanoestructura proteica. En algunas modalidades,

el polipéptido es una proteína de fusión, que incluye, en orden de extremo N a C, el polipéptido

del SARS-CoV-2 recombinante, un enlazador, un enlazador de extensión de extremo N y el polipéptido componente de la nanoestructura proteica. En algunas modalidades, el enlazador de extensión de extremo N es de extensión helicoidal I53-50A. En algunas modalidades, la secuencia polipeptídica del enlazador de extensión de extremo N es EKAAKAEEAARK (SEQ ID NO: 237).

[0190] En algunas modalidades, el enlazador se selecciona del grupo que se muestra en la **Tabla 11**.

Tabla 11.

Secuencia	Longitud	SEQ ID NO:
GSS	3	85
GSGS	4	86
GGSGEKP	7	87
GGSGQKP	7	88
GGSGGSGS	8	89
GGSGGSGEKP	10	90
GGSGGSGQKP	10	91
GGSGGSGGSGGS	12	92
GSGGSGSGSGGS	12	93
GGGGSGGGSGGGGS	15	94
GGGGSGGGSGGGGS	15	95
GGSGGSGSGGSGGSGS	16	96
GGGGSGGGSGGGSGG	17	97
SGGGSGGSGSGGSGGSGS	18	98
EPEGSGGSGSGGSGGSGS	19	99
YGGSGGSGGSGSGGSGGSGS	20	100
GGSGGSGSGGSGGSGSGGSGSGGS	24	101
GSGGSGGSGGSGGSGSGGSGGSGGS	24	102
KSDELLGSGGSGSGSGGSEKAAKAEEAARK	30	103

10 Nanoestructuras proteicas autoensamblables

[0191] En otro aspecto, la descripción proporciona una nanoestructura de proteína

autoensamblable que incluye un primer componente que comprende un polipéptido recombinante que incluye un fragmento antigénico de una glucoproteína Spike de coronavirus como se describió anteriormente. Opcionalmente, la nanoestructura de proteína autoensamblable puede incluir un segundo componente que comprende una segunda proteína.

5 El primer componente, como un ensamblaje trimérico, comprende una interfaz proteína-proteína que induce la autoasociación de una región del primer componente para formar componentes básicos triméricos. En nanoestructuras proteicas que tienen dos o más componentes, cada copia del primer dominio de ensamblaje comprende también una interfaz expuesta a la superficie que interactúa con una interfaz expuesta a la superficie complementaria en el
10 segundo dominio de ensamblaje. De manera similar, el segundo componente, como un ensamblaje pentamérico, está adaptado para multimerizarse con el primer componente. Tal como se describe en King et al. (Nature 510, 103-108, 2014), Bale et al. (Science 353, 389-394, 2016) y las publicaciones de patente WO2014124301 A1 y US20160122392 A1, la interfaz proteína-proteína complementaria entre un primer dominio de ensamblaje y un segundo dominio
15 de ensamblaje guía el ensamblaje de múltiples copias del dominio de ensamblaje trimérico y el segundo dominio de ensamblaje en una nanoestructura proteica diana. En algunas modalidades, cada uno de los dominios de ensamblaje triméricos de la nanoestructura proteica tiene una proteína antigénica, o un fragmento antigénico de esta, enlazado (*por ejemplo*, como una fusión genética), y estas nanoestructuras proteicas exhiben las proteínas en su valencia
20 completa. En otras modalidades, las nanoestructuras proteicas de la divulgación comprenden uno o más primeros dominios de ensamblaje con proteínas de antígenos, o fragmentos antigénicos de estas (por ejemplo, como fusiones genéticas), así como uno o más primeros dominios de ensamblaje que no tienen proteínas antigénicas, y dichas nanoestructuras proteicas exhiben las proteínas con valencia parcial. Los primeros dominios de ensamblaje
25 pueden ser cualquier secuencia polipeptídica que forma un trímero y que interactúa con un segundo dominio de ensamblaje para guiar el ensamblaje a una nanoestructura proteica diana. En algunas modalidades, la nanoestructura proteica comprende un primer polipéptido y un segundo polipéptido seleccionados entre los descritos en US 20130274441 A1, US 20150356240 A1, US 20160122392 A1, WO 2018187325 A1, cada uno de los cuales se
30 incorpora mediante esta referencia en la presente en su totalidad.

[0192] En algunas modalidades, un solo componente se autoensambla en la nanoestructura

proteica. En algunas modalidades, una o más muestras purificadas del primer y segundo componente para su uso en la formación de una nanoestructura proteica se mezclan en una relación molar aproximadamente equimolar en condiciones acuosas (*por ejemplo*, una nanoestructura proteica icosaédrica I53-50A/B). El primer y segundo componentes (a través de los dominios de ensamblaje) interactúan entre sí para guiar el ensamblaje de la nanoestructura proteica diana. Es posible confirmar el ensamblaje exitoso de la nanoestructura proteica diana mediante el análisis de la reacción de ensamblaje *in vitro* mediante métodos bioquímicos o biofísicos usados para evaluar el tamaño físico de las proteínas o ensamblajes proteicos, lo que incluye, de modo no taxativo, la cromatografía de exclusión por tamaño, electroforesis en gel natural (no desnaturalizante), dispersión de luz dinámica, dispersión de luz multiangular, ultracentrifugación analítica, microscopia electrónica de tinción negativa, microscopia crioelectrónica o cristalografía de rayos X. Si fuera necesario, es posible purificar la nanoestructura proteica ensamblada de otras especies o moléculas presentes en la reacción de ensamblaje *in vitro* con técnicas preparativas comúnmente utilizadas para aislar proteínas por su tamaño físico, lo que incluye, de modo no taxativo, cromatografía de exclusión por tamaño, ultracentrifugación preparativa, filtración de flujo tangencial o electroforesis en gel preparativa. La presencia de la proteína antigénica en la nanoestructura proteica puede evaluarse mediante técnicas comúnmente utilizadas para determinar la identidad de moléculas proteicas en soluciones acuosas, incluyendo pero no limitándose a SDS-PAGE, espectrometría de masas, secuenciación de proteínas, ELISA, resonancia de plasmón superficial, interferometría de bicapa o análisis de aminoácidos. La accesibilidad de la proteína en el exterior de la nanoestructura proteica, así como su conformación o antigenicidad, pueden evaluarse mediante técnicas comúnmente utilizadas para detectar la presencia y conformación de un antígeno, que incluye, de modo no taxativo, la unión por anticuerpos monoclonales, anticuerpos monoclonales específicos de conformación, resonancia de plasmón superficial, interferometría de biocapa o antisueros específicos del antígeno.

[0193] En diversas modalidades, las nanoestructuras proteicas de la descripción comprenden dos o más primeros polipéptidos distintos que llevan diferentes proteínas antigénicas como fusiones genéticas; estas nanoestructuras proteicas coexhiben múltiples proteínas diferentes (*por ejemplo*, diferentes variantes de polipéptido de coronavirus que incluye un dominio de unión al receptor (RBD) y un subdominio 1 adyacente a RBD (SD1) de la glucoproteína Spike) en la

misma nanoestructura proteica. Estas nanoestructuras proteicas multiantigénicas se producen mediante el ensamblaje *in vitro* con mezclas de dos o más antígenos, cada uno de los cuales comprende un dominio de ensamblaje. La fracción de cada antígeno en la mezcla determina la valencia media de cada proteína antigénica en las nanoestructuras proteicas resultantes. La presencia y la valencia media de cada antígeno en una muestra dada pueden evaluarse mediante análisis cuantitativo utilizando las técnicas descritas anteriormente para evaluar la presencia de proteínas antigénicas en nanoestructuras proteicas de valencia completa.

[0194] Las nanoestructuras proteicas de la presente invención pueden comprender ensamblajes de proteínas multiméricas adaptadas para la visualización de polipéptidos de coronavirus que incluyen un dominio de unión al receptor (RBD) y un subdominio 1 adyacente a RBD (SD1) de la glucoproteína Spike, o variantes antigénicas o fragmentos de estos. Las nanoestructuras proteicas de la presente invención comprenden al menos un primer componente y, opcionalmente, un segundo componente. El primer componente puede incluir la sustitución de al menos un residuo de aminoácido o la adición en el extremo N o C de uno o más residuos de aminoácido. En algunas modalidades, la proteína de fusión comprende una secuencia proteica determinada por métodos computacionales. Este primer componente puede formar todo el núcleo de la nanoestructura proteica; o el núcleo de la nanoestructura proteica puede incluir un segundo componente o un tercer, cuarto, quinto componente, etcétera. En algunas modalidades, el primer componente es un componente trimérico en el que los dominios de ensamblaje forman trímeros relacionados por simetría rotacional triple, y/o el segundo componente es un componente pentamérico, en el que los dominios de ensamblaje forman pentámeros relacionados por simetría rotacional quíntuple. En algunas modalidades, la combinación de los dos componentes forma una "partícula icosaédrica" con simetría en I53. Juntos, estos componentes pueden disponerse de forma tal que los miembros de cada componente estén relacionados entre sí mediante operadores de simetría. Un método computacional general para diseñar materiales proteicos autoensamblables, que implica el acoplamiento simétrico de bloques de construcción de proteínas en una arquitectura simétrica objetivo, se describe en la pub. de patente de los EE. UU. 2015/0356240 A1.

[0195] El "núcleo" de la nanoestructura proteica se usa en la presente para describir la porción central de la nanoestructura proteica que une las diversas copias de polipéptidos que incluyen un fragmento antigénico de la glucoproteína Spike del coronavirus, visualizados en la

nanoestructura proteica. En algunas modalidades, la proteína de fusión comprende un primer polipéptido que comprende un fragmento antigénico de la glucoproteína Spike de un coronavirus, un enlazador y un primer dominio de ensamblaje. En algunas modalidades, el fragmento antigénico de la glucoproteína Spike de un coronavirus, está enlazado de forma no covalente o covalente al dominio de ensamblaje. Por ejemplo, un anticuerpo o fragmento antigénico de este se puede fusionar al primer componente y configurarse para unirse a una parte del primer componente o a una etiqueta química en el primer componente. Por ejemplo, puede emplearse un enlazador de estreptavidina-biotina (o neuravidina-biotina). Alternativamente, se pueden utilizar varios enlaces bioconjugados. En algunas modalidades de la presente descripción, el antígeno comprende otras secuencias polipeptídicas además de los polipéptidos de coronavirus.

[0196] Un ejemplo no taxativo de una modalidad es un fragmento antigénico de la glucoproteína Spike de un coronavirus fusionado genéticamente a un componente de la nanoestructura proteica, que opcionalmente se expresa de forma recombinante en una célula hospedadora (*por ejemplo*, células de *E. coli*); junto con un ensamblaje proteico pentamérico como un segundo componente, que se expresa opcionalmente de forma recombinante en la misma célula hospedadora o una célula hospedadora diferente. Estos dos componentes se ensamblan en una nanoestructura proteica que exhibe 60 monómeros de fragmento antigénico alrededor de un núcleo icosaédrico. En algunas modalidades, un fragmento antigénico de la glucoproteína Spike de un coronavirus mezclado con otra proteína antigénica en la misma nanoestructura proteica, tal como dos variantes diferentes de la glucoproteína Spike de un coronavirus. En algunas modalidades, la nanoestructura proteica comprende, además de uno o más fragmentos antigénicos de la glucoproteína Spike de un coronavirus que incluye el dominio de unión al receptor (RBD) y, opcionalmente, un subdominio 1 adyacente a RBD (SD1), antígenos de otros organismos patógenos y, por lo tanto, puede usarse como una vacuna combinada. En algunas modalidades, la nanoestructura proteica está además unida a polipéptidos u otros agentes capaces de actuar como adyuvantes. En algunas modalidades, el primer componente y/o el segundo componente comprenden uno o más epítomos de linfocitos T, opcionalmente un epítomo de linfocitos T de origen heterólogo.

[0197] Otras posibles disposiciones de los componentes de la presente descripción incluyen la nanoestructura proteica adaptada para la visualización de hasta 12, 24 o 60 monómeros. En

algunas modalidades, un componente puede comprender un polipéptido enlazado a diversos antígenos, de manera que la nanoestructura proteica exhiba antígenos diferentes en la misma nanoestructura. En algunas modalidades, se exhiben 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 o más antígenos diferentes. Se proporcionan nanoestructuras proteicas ilustrativas y no limitantes en Bale et al. *Science* 353:389-94 (2016); Heinze et ál. *J. Phys. Chem B*. 120:5945-5952 (2016); King et al. *Nature* 510:103-108 (2014); y King et al. *Science* 336:1171-71 (2012).

Modalidades de unión

[0198] Las nanoestructuras proteicas de la presente descripción presentan proteínas antigénicas de diversas maneras, incluida la fusión de genes o por otros medios descritos en la presente. Tal como se utiliza en la presente, "enlazado a" o "unido a" denota cualquier medio conocido en la técnica para hacer que dos polipéptidos se asocien. La asociación puede ser directa o indirecta, reversible o irreversible, débil o fuerte, covalente o no covalente, y selectiva o no selectiva.

[0199] En algunas modalidades, la unión se consigue mediante modificación genética para crear una fusión en el extremo N o C de polipéptidos potencialmente antigénicos de la nanoestructura proteica.

[0200] En algunas modalidades, la unión se logra mediante la unión covalente postraducciona de una o más pluralidades de proteína antigénica. En algunas modalidades, la reticulación química se usa para unir de forma no específica el antígeno a una nanoestructura proteica. En algunas modalidades, se usa reticulación química para unir específicamente la proteína antigénica a una nanoestructura proteica (*por ejemplo*, al primer polipéptido o al segundo polipéptido). Se conocen en la técnica diversas reticulaciones químicas específicas y no específicas, como la química Click y otros métodos. En general, cualquier bioconjugado/reticulación química utilizada para unir dos proteínas puede adaptarse para su uso en las nanoestructuras proteicas descritas en la presente. En particular, pueden utilizarse las químicas empleadas en la creación de inmunoconjugados o conjugados anticuerpo-fármaco. En algunas modalidades, se crea una nanoestructura proteica usando un enlazador escindible o no escindible. Los procesos y métodos para la conjugación de antígenos con portadores se proporcionan, por ejemplo, en la pub. de patente de los EE. UU. núm. 2008/0145373 A1.

[0201] En algunas modalidades, la unión se consigue mediante unión no covalente entre un

componente y el antígeno. En algunas modalidades, la proteína antigénica se diseña para estar cargada negativamente en al menos una superficie y el polipéptido del núcleo se diseña para estar cargado positivamente en al menos una superficie, o cargado positiva y negativamente, respectivamente. Esto puede promover la asociación intermolecular entre la proteína antigénica y el polipéptido nuclear componente por fuerza electrostática. En algunas modalidades, se emplea la complementariedad de forma para provocar la unión de la proteína antigénica al núcleo del componente. La complementariedad de formas puede ser preexistente o diseñada racionalmente. En algunas modalidades, se usa el diseño computacional de interfaces de proteína-proteína para lograr la unión. En una modalidad, el antígeno está marcado con biotina y el polipéptido comprende una estreptavidina, o viceversa. En una modalidad, la estreptavidina se muestra por fusión génica o de otro modo como un tetrámero en un eje cuádruple del núcleo componente y el antígeno marcado con biotina es monomérico, dimérico o tetramérico, lo que permite la asociación al núcleo componente en una configuración apropiada para la multimerización nativa del antígeno. En algunas modalidades, se usa un adaptador basado en proteína para capturar la proteína antigénica. En algunas modalidades, el polipéptido se fusiona a una proteína capaz de unirse a una proteína complementaria, que se fusiona a la proteína antigénica.

[0202] La inmunogenicidad de la proteína antigénica puede controlarse alterando la orientación de la proteína antigénica con respecto al núcleo componente. Dependiendo de cómo se fije la proteína antigénica al núcleo componente de la nanoestructura proteica, la proteína antigénica puede mostrarse en varias orientaciones. En algunas modalidades, la proteína antigénica se muestra de modo que uno o más epítopos conocidos estén orientados en o hacia el extremo distal de la proteína antigénica, de modo que estos epítopos sean accesibles preferentemente al sistema inmunológico. En algunas modalidades, la orientación recapitulará la orientación de la proteína de glucoproteína Spike con respecto al virus. La elección de la orientación puede dirigir el sistema inmune hacia uno u otro epítipo.

[0203] En algunas modalidades, la preferencia del epítipo se controla por otros medios, tales como el posicionamiento de los glicanos en la nanoestructura proteica mediante la adición o sustracción del motivo de secuencia de glicano ligado a N N-X[T/S] en posiciones predeterminadas en la secuencia de aminoácidos de un componente incluido en la secuencia de aminoácidos de la proteína antigénica.

[0204] En algunas modalidades, los epítomos que se encuentran a distancias intermedias desde el extremo proximal al distal serán los preferidos sobre los epítomos situados más distalmente, dependiendo de varias consideraciones que incluyen, entre otros, la geometría general de la nanoestructura proteica, la hidrofobia de la superficie, la carga de la superficie y la unión competitiva de proteínas presentes endógenamente en el sujeto o proteínas proporcionadas exógenamente en la composición de la vacuna. La presente descripción abarca todos los métodos conocidos de diseño racional de la estructura de proteínas y lo anterior no pretende ser limitante.

Secuencias polipeptídicas

[0205] La pub. de patente de los EE. UU núm. 2015/0356240 A1 describe diversos métodos para diseñar ensamblajes de proteínas. Como se describe en la publicación de patente estadounidense n.º US 2016/0122392 A1 y en la publicación de patente internacional n.º WO 2014/124301 A1, se diseñaron los polipéptidos aislados mostrados en la **Tabla 12** por su capacidad de autoensamblarse en pares para formar nanoestructuras proteicas, tales como partículas icosaédricas. El diseño implicó el diseño de residuos de interfaz adecuados para cada miembro del par polipeptídico que puede ensamblarse para formar la nanoestructura proteica. La nanoestructura proteica formada de ese modo incluye interfaces de polipéptido-polipéptido no naturales, no covalentes y repetidas simétricamente que orientan un primer dominio de ensamblaje y un segundo dominio de ensamblaje en nanoestructuras proteicas, tal como una con simetría icosaédrica. Así, en una modalidad, un primer elemento de repetición y un segundo elemento de repetición del componente se seleccionan del grupo que consiste en las SEQ ID NO: 23-73. En cada caso, se incluye un residuo de metionina en el extremo N presente en la proteína de longitud completa, pero puede eliminarse para realizar una fusión que no esté incluida en la secuencia. Los residuos identificados en la **Tabla 12** están numerados comenzando por una metionina de extremo N (no mostrada). En diversas modalidades, se eliminan uno o más residuos adicionales del extremo N y/o se agregan residuos adicionales al extremo N (por ejemplo, para formar una extensión helicoidal).

Tabla 12

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
I53-34A SEQ ID NO:23	trímero	EGMDPLAVLAESRLLPLLTVRGGEDL AGLATVLELMGVGALEITLRTEKGLEA LKALRKSGLLLGAGTVRSPKEAEAAL EAGAAFLVSPGLLEEVAALAQARGVP YLPGLVTPTEVERALALGLSALKFFPA EPFQGVRLRAYAEVFPEVRFLPTGG IKEEHLPHYAALPNLLAVGGSWLLQG DLAAVMKKVKAALKALLSPQAPG	I53-34A: 28,32,36,37,1 86,188,191,19 2,195
I53-34B SEQ ID NO:24	pentámero	TKKVGIVDTTFARVDMAEAAIRTLKAL SPNIKIIRKTVPGIKDLPVACKKLLEEE GCDIVMALGMPGKAEKDKVCAHEAS LGLMLAQLMTNKHIIIEVVFVHEDEAKD DDELDILALVRAIEHAANVYYLLFKPE YLTRMAGKGLRQGREDA GPARE	I53-34B: 19,20,23,24,2 7,109,113,116 ,117,120,124, 148
I53-40A SEQ ID NO:25	pentámero	TKKVGIVDTTFARVDMASAAITLKME SPNIKIIRKTVPGIKDLPVACKKLLEEE GCDIVMALGMPGKAEKDKVCAHEAS LGLMLAQLMTNKHIIIEVVFVHEDEAKD DAELKILARRAIEHALNVYYLLFKPE YLTRMAGKGLRQGFEDAGPARE	I53-40A: 20,23,24,27,2 8,109,112,113 ,116,120,124
I53-40B SEQ ID NO:26	trímero	STINNQLKALKVIPVIAIDNAEDIPLGK VLAENGLPAAEITFRSSAAVKAIMLLR SAQPEMLIGAGTILNGVQALAAKEAG ATFVVSPGFNPNTVRACQIIGIDIVPG VNNPSTVEAALEMGLTTLKFFPAEAS GGISMVKS LVGPYGD IRLMPTGGITP SNIDNYLAIPQVLACGGTWMVDKKLV TNGEWDEIARLTREIVEQVNP	I53-40B: 47,51,54,58,7 4,102

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
I53-47A SEQ ID NO:27	trímero	PIFTLNTNIKATDVPSDFLSLTSRLVGL ILSKPGSYVAVHINTDQQLSFGGSTN PAAFGTLMSIGGIEPSKNRDHSAVL DHLNAMLGIPKNRMYIHVNLNGDDV GWNGTTF	I53-47A: 22,25,29,72,7 9,86,87
I53-47B SEQ ID NO:28	pentámero	NQSHSKDYETVRIAVVRARWHADIVD ACVEAFEIAMAAIGGDRFAVDVFDVP GAYEIPLHARTLAETGRYGAVLGTA FVVNGGIYRHEFVASAVIDGMMNVQLS TGVPVLSAVLTPHRYRDSA EHHRRFFA AHFAVKGVEAARACIEILAAREKIAA	I53-47B: 28,31,35,36,3 9,131,132,135 ,139,146
I53-50A SEQ ID NO:29	trímero	EELFKKKHIVAVLRANSVEEAIEKAVA VFAGGVHLIEITFTVPDADTVIKALSVL KEKGAIIGAGTVTSVEQCRKAVESGA EFIVSPHLDEEISQFCKEKGVFYMPG VMTPTLVKAMKLGHTILKLFPGEVV GPQFVKAMKGPFNPVKFVPTGGVNL DNVCEWFKAGVLAVGVGSALVKGTP DEVREKAKAFVEKIRGCTE	I53-50A: 25,29,33,54,5 7
I53-50B SEQ ID NO:30	pentámero	NQSHSKDYETVRIAVVRARWHAEIVD ACVSAFEAAMADIGGDRFAVDVFDVP GAYEIPLHARTLAETGRYGAVLGTA FVVNGGIYRHEFVASAVIDGMMNVQLS TGVPVLSAVLTPHRYRDSAHTLLFL ALFAVKGMEAAARACVEILAAREKIAA	I53-50B: 24,28,36,124, 125,127,128,1 29, 131,132,133,1 35,139
I53-51A SEQ ID NO:31	trímero	FTKSGDDGNTNVINKRVGKDSPLVNF LGDLDELNSFIGFAISKIPWEDMKKDL ERVQVELFEIGEDLSTQSSKKKIDESY	I53-51A: 80,83,86,87,8 8,90,91,94,16

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		VLWLLAATAIYRIESGPVKLFVIPGGS EEASVLHVTRSVARRVERNAVKYTKE LPEINRMIIVYLNRLSSLLFAMALVANK RRNQSEKIYEIGKSW	6,172,176
I53-51B SEQ ID NO:32	pentámero	NQHSHKDYETVRIAVVRARWHADIVD QCVRAFEEAMADAGGDRFAVDVFDV PGAYEIPLHARTLAETGRYGAVLGTA FVVNGGIYRHEFVASAVIDGMMNVQL STGVPVLSAVLTPHRYRSSREHHEFF REHFMVKGVEAAAACITILAAREKIAA	I53-51B: 31,35,36,40,1 22,124,128,13 1,135,139,143 ,146,147
I52-03A SEQ ID NO:33	pentámero	GHTKGPTPQQHDGSALRIGIVHARW NKTIIMPLLIGTIAKLLECGVKASNIVVQ SVPGSWELPIAVQRLYSASQLQTPSS GPSLSAGDLLGSSTTDLTALPTTTAS STGPFDALIAIGVLIKGETMHFEYIADS VSHGLMRVQLDTGVPVIFGVLTVLTD DQAKARAGVIEGSHNHGEDWGAAV EMGVRRRDWAAGKTE	I52-03A: 28,32,36,39,4 4,49
I52-03B SEQ ID NO:34	dímero	YEVDHADVYDLFYLGKDYAAEAS DIADLVRSTPEASSLLDVACGTGTH LEHFTKEFGDTAGLELSEDMLTARK RLPDATLHQGDMRDFQLGRKFSVV SMFSSVGYLKTVAELGAASFAEHL EPGGVVVVEPWWFPETFADGWVSA DVVRRDGRTVARVSHSVREGNATRM EVHFTVADPGKGVRHFSDVHLITLFH QREYEAAFMAAGLRVEYLEGGPSGR GLFVGVPA	I52-03B: 94,115,116,20 6,213

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
I52-32A SEQ ID NO:35	dímero	GMKEKFVLIITHGDFGKGLLSGAEVII GKQENVHTVGLNLGDNIEKVAKEVM RIIIAKLAEDKEIIIVVDLFGGSPFNIALE MMKTFDVKVITGINMPMLVELLTSINV YDTTELLENISKIGKDGIKVIEKSSLKM	I52-32A: 47,49,53,54,5 7,58,61,83,87, 88
I52-32B SEQ ID NO:36	pentámero	KYDGSKLRIGILHARWNLEIIAALVAG AIKRLQEFVKAENIIIVTVPGSFELPY GSKLFVEKQKRLGKPLDAIPIGVLIKG STMHFEYICDSTTHQLMKLNFELGIPV IFGVLTCLTDEQAEARAGLIEGKMHN HGEDWGAAAVEMATKFN	I52-32B: 19,20,23,30,4 0
I52-33A SEQ ID NO:37	pentámero	AVKGLGEVDQKYDGSKLRIGILHARW NRKIILALVAGAVLRLLFEGVKAENIIIE TVPGSFELPYGSKLFVEKQKRLGKPL DAIPIGVLIKGSTMHFEYICDSTTHQL MKLNFELGIPVIFGVLTCLTDEQAEAR AGLIEGKMHNHGEDWGAAAVEMATK FN	I52-33A: 33,41,44,50
I52-33B SEQ ID NO:38	dímero	GANWYLDNESSRLSFTSTKNADIAEV HRFLVLHGKVDPKGLAEVEVETESIS TGIPLRDMLLRVLVFQVSKFPVAQINA QLDMRPINNLAGAQLRLPLTVSL RGKSHSYNAELLATRLDERRFQVVTL EPLVIHAQDFDMVRAFNALRLVAGLS AVSLSVPVGAVLIFTAR	I52-33B: 61,63,66,67,7 2,147,148,154 ,155
I32-06A SEQ ID NO:39	dímero	TDYIRDGSAIKALSFAIILAEADLRHIPQ DLQRLAVRVIHACGMVDVANDLAFSE GAGKAGRNALLAGAPILCDARMVAE	I32-06A: 9,12,13,14,20, 30,33,34

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		GITRSRLPADNRVIYTLSDPSVPELAK KIGNTRSAAALDLWLPHIEGSIVAIGN APTALFRLFELLDAGAPKPALIIGMPV GFVGAAESKDELAANSRGVPYVIVRG RRGGSAMTAAAVNALASERE	
I32-06B SEQ ID NO:40	trímero	ITVFGLKSKLAPRREKLAEVIYSSLHL GLDIPKGKHAIRFLCLEKEDFYYPFDR SDDYTVIEINLMAGRSEETKMLLIFLLF IALERKLGIRAH DVEITIKEQPAHCWG FRGRTGDSARDLDYDIYV	I32-06B: 24,71,73,76,7 7,80,81,84,85, 88,114,118
I32-19A SEQ ID NO:41	trímero	GSDLQKLQRFSTCDISDGLLN VYNIPT GGYFPNLTAISPPQNSSIVGTAYTVLF APIDDPRPAVNYIDSVPPNSILVLALEP HLQSQFHFPFIKITQAMYGGLMSTRAQ YLKSNGTVVFGRIRDVDEHRTL NHPV FAYGVGSCAPKAVVKAVGTNVQLKIL TSDGVTQTICPGDYIAGDNNGIVRIPV QETDISKLVTYIEKSIEVDRLVSEAIKN GLPAKAAQTARRMVLKDYI	I32-19A: 208,213,218,2 22,225,226,22 9,233
I32-19B SEQ ID NO:42	dímero	SGMRVYLGADHAGYELKQAIIFLKM TGHEPIDCGALRYDADDDYP AFCIAA ATRTVADPGSLGIVLGGSGNGEQIAA NKVPGARCALAWSVQTAALAREHNN AQLIGIGGRMHTLEEALRIVKAFV TTP WSKAQRHQRRIDILAEYERTHEAPPV PGAPA	I32-19B: 20,23,24,27,1 17,118,122,12 5
I32-28A SEQ ID NO:43	trímero	GDDARIAAIGDVDELNSQIGVLLAEPL PDDVRAALSAIQHDLFDLGGELCIPG	I32-28A: 60,61,64,67,6

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		HAAITEDHLLRLALWLVHYNGQLPPL EEFILPGGARGAALAHVCRTVCRRAE RSIKALGASEPLNIAPAAYVNLLSDLLF VLARVLNRAAGGADVLDWRTRAH	8,71,110,120, 123,124,128
I32-28B SEQ ID NO:44	dímero	ILSAEQSFTLRHPHGQAAALAFVREP AAALAGVQRLRGLDSDGEQVWGELL VRVPLLGEVDLPFRSEIVRTPQGAEL RPLTLTGERAWVAVSGQATAAEGGE MAFAFQFQAHLATPEAEGEGGAAFE VMVQAAAGVTLLLVAMALPQGLAAGL PPA	I32-28B: 35,36,54,122, 129,137,140,1 41,144,148
I53-40A.1 SEQ ID NO:45	pentámero	TKKVGIVDTTFARVDMASAAITLKME SPNIKIIRKTVPGIKDLPVACKKLEEE GCDIVMALGMPGKKEKDKVCAHEAS LGLMLAQLMTNKHIIIEVFVHEDEAKD DAELKILARRAIEHALNVYYLLFKPE YLTRMAGKGLRQGFEDAGPARE	I53-40A: 20,23,24,27,2 8,109,112,113 ,116,120,124
I53-40B.1 SEQ ID NO:46	trímero	DDINNQLKRLKVIPVIAIDNAEDIPLGK VLAENGLPAAEITFRSSAAVKAIMLLR SAQPEMLIGAGTILNGVQALAAKEAG ADFFVSPGFNPNTVRACQIIGIDIVPG VNNPSTVEQALEMGLTTLKFFPAEAS GGISMVKS LVGPYGDIRLMPTGGITP DNIDNYLAIPQVLACGGTWMVDKKLV RNGEWDEIARLTREIVEQVNP	I53-40B: 47,51,54,58,7 4,102
I53-47A.1 SEQ ID NO:47	trímero	PIFTLNTNIKADDVPSDFLSLTSRLVGL ILSKPGSYVAVHINTDQQLSFGGSTN PAAFGLMSIGGIEPDKNRDHSAVL	I53-47A: 22,25,29,72,7 9,86,87

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		DHLNAMLGIPKNRMYIHFVN LNGDDV GWNGTTF	
I53-47A.1NegT2 SEQ ID NO:48	trímero	PIFTLNTNIKADDVPSDFLSLTSRLVGL ILSEPGSYVAVHINTDQQLSFGGSTN PAAFGLMSIGGIEPDKNEDHSAVL DHLNAMLGIPKNRMYIHFVDLDGDDV GWNGTTF	I53-47A: 22,25,29,72,7 9,86,87
I53-47B.1 SEQ ID NO:49	pentámero	NQSHKDHETVRIAVVRARWHADIVD ACVEAFEIAMAAIGGDRFAVDVFDVP GAYEIP LHARTLAETGRYGAVLGTA FVVNGGIYRHEFVASAVIDGMMNVQLD TGVPVLSAVLTPHRYRDSDEHHRFFA AHFAVKGVEAARACIEILNAREKIAA	I53-47B: 28,31,35,36,3 9,131,132,135 ,139,146
I53-47B.1NegT2 SEQ ID NO:50	pentámero	NQSHKDHETVRIAVVRARWHADIVD ACVEAFEIAMAAIGGDRFAVDVFDVP GAYEIP LHARTLAETGRYGAVLGTA FVVDGGIYDHEFVASAVIDGMMNVQLD TGVPVLSAVLTPHEYEDSDEDHEFFA AHFAVKGVEAARACIEILNAREKIAA	I53-47B: 28,31,35,36,3 9,131,132,135 ,139,146
I53-50A.1 SEQ ID NO:51	trímero	EELFKKHKIVAVLRANSVEEAIEKAVA VFAGGVHLIEITFTVPDADTVIKALSVL KEKGAIIGAGTVTSVEQCRKAVESGA EFIVSPHLDEEISQFCKEKGVFYMPG VMTPTLVKAMKLGHDKLFPGEVV GPQFVKAMKGPFPNVKFVPTGGVNL DNVCEWFKAGVLAVGVGDALVKGDP DEVREKAKKFVEKIRGCTE	I53-50A: 25,29,33,54,5 7
I53-	trímero	EELFKKHKIVAVLRANSVEEAIEKAVA	I53-50A:

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
50A.1NegT2 SEQ ID NO:52		VFAGGVHLIEITFTVPDADTVIKALSVL KEKGAIIGAGTVTSVEQCRKAVESGA EFIVSPHLDEEISQFCKEKGVFYMPG VMTPTLVKAMKLGHDILKLPGEVV GPEFVEAMKGPFPNVKFVPTGGVDL DDVCEWFDAGVLAVGVGDALVEGDP DEVREDAKEFVEEIRGCTE	25,29,33,54,57
I53-50A.1PosT1 SEQ ID NO:53	trímero	EELFKKHKIVAVLRANSVEEAIEKAVA VFAGGVHLIEITFTVPDADTVIKALSVL KEKGAIIGAGTVTSVEQCRKAVESGA EFIVSPHLDEEISQFCKEKGVFYMPG VMTPTLVKAMKLGHDILKLPGEVV GPQFVKAMKGPFPNVKFVPTGGVNL DNVCKWFKAGVLAVGVGKALVKGKP DEVREKAKKFVKKIRGCTE	I53-50A: 25,29,33,54,57
I53-50B.1 SEQ ID NO:54	pentámero	NQHSHKDHETVRIAVVRARWHAEIVD ACVSAFEAAMRDIGGDRFAVDVFDV PGAYEIP LHARTLAETGRYGAVLGTA FVVNGGIYRHEFVASAVIDGMMNVQL DTGVPVLSAVLTPHRYRDS DAHTLLF LALFAVKGMEAAARACVEILAAREKIAA	I53-50B: 24,28,36,124, 125,127,128,1 29,131,132,13 3,135,139
I53-50B.1NegT2 SEQ ID NO:55	pentámero	NQHSHKDHETVRIAVVRARWHAEIVD ACVSAFEAAMRDIGGDRFAVDVFDV PGAYEIP LHARTLAETGRYGAVLGTA FVVDGGIYDHEFVASAVIDGMMNVQL DTGVPVLSAVLTPHEYEDSDADTLLF LALFAVKGMEAAARACVEILAAREKIAA	I53-50B: 24,28,36,124, 125,127,128,1 29,131,132,13 3,135,139
I53-50B.4PosT1	trímero	NQHSHKDHETVRIAVVRARWHAEIVD	I53-50B:

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
SEQ ID NO:56		ACVSAFEAAMRDIGGDRFAVDVFDV PGAYEIPLHARTLAETGRYGAVLGTA FVVNGGIYRHEFVASAVINGMMNVQL NTGVPVLSAVLTPHNYDKSKAHTLLF LALFAVKGMEAAARACVEILAAREKIAA	24,28,36,124, 125,127,128,1 29,131,132,13 3,135,139
Género I53-40A SEQ ID NO:57	pentámero	TKKVGIVDTTFARVDMASAILTLKME SPNIKIIRKTVPGIKDLPVACKKLLEEE GCDIVMALGMPGK(A/K)EKDKVCAHE ASLGLMLAQLMTNKHIIIEVFVHEDEAK DDAELKILAARRAIEHALNVYYLLFKP EYLTRMAGKGLRQGFEDAGPARE	
Género I53-40B SEQ ID NO:58	trímero	(S/D)(T/D)INNQLK(A/R)LKVIPVIAIDNA EDIIPLGKVLAEGLPAAEITFRSSAAV KAIMLLRSAQPEMLIGAGTILNGVQAL AAKEAGA(T/D)FVVSPGFNPNTVRAC QIIGIDIVPGVNNPSTVE(A/Q)ALEMGL TTLKFFPAEASGGISMVKS LVGPYGDI RLMPTGGITP(S/D)NIDNYLAIPQVLAC GGTWMVDKKLV(T/R)NGEWDEIARLT REIVEQVNP	
Género I53-47A SEQ ID NO:59	trímero	PIFTLNTNIKA(T/D)DVPSDFLSLTSRL VGLILS(K/E)PGSYVAVHINTDQQLSF GGSTNPAAFGLMSIGGIEP(S/D)KN(R/E)DHSAVLFDHLNAMLGIPKNRMYI HFV(N/D)L(N/D)GDDVGWNGTTF	
Género I53-47B SEQ ID NO:60	pentámero	NQHSHKD(Y/H)ETVRIAVVRARWHAD IVDACVEAFEIAMAAIGGDRFAVDVFD VPGAYEIPLHARTLAETGRYGAVLGTA	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		AFVV(N/D)GGIY(R/D)HEFVASAVIDG MMNVQL(S/D)TGVPVLSAVLTPH(R/E))Y(R/E)DS(A/D)E(H/D)H(R/E)FFAAHF AVKGVEAARACIEIL(A/N)AREKIAA	
Género I53-50A SEQ ID NO:61	trímero	EELFKKHKIVAVLRANSVEEAIEKAVA VFAGGVHLIEITFTVPDADTVIKALSVL KEKGAIIGAGTVTSVEQCRKAVESGA EFIVSPHLDEEISQFCKEKGVFYMPG VMTPTLVKAMKLGH(T/D)ILKLFPGE VVGP(Q/E)FV(K/E)AMKGPFPNVKFV PTGGV(N/D)LD(N/D)VC(E/K)WF(K/D) AGVLAVGVG(S/K/D)ALV(K/E)G(T/D/K))PDEVRE(K/D)AK(A/E/K)FV(E/K)(K/E)I RGCTE	
Género I53-50B SEQ ID NO:62	pentámero	NQSHSKD(Y/H)ETVRIAVVRARWHA IVDACVSAFEAAM(A/R)DIGGDRFAVD VFDVPGAYEIPHLARTLAETGRYGAV LGTAFFVV(N/D)GGIY(R/D)HEFVASAVI (D/N)GMMNVQL(S/D/N)TGVPVLSAVL TPH(R/E/N)Y(R/D/E)(D/K)S(D/K)A(H/D))TLLFLALFAVKGMEAAARACVEILAA RKIAA	
T32-28A SEQ ID NO:63	dímero	GEVPIGDPKELNGMEIAAVYLQPIEME PRGIDLAASLADIIHLEADIIHALKNNPN GFPEGFWMPYLTIAAYALANADTGAIK TGTLMPMVADDGPHYGANIAMEKDK KGGFGVGTYALTFLISNPEKQGFRH	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		VDEETGVGKWFEPFVVTYFFKYTGTPK	
T32-28B SEQ ID NO:64	trímero	SQAIGILELTSIAKGMELGDAMLKSAN VDLLVSKTISPGKFLMLGGDIGAIQQ AIETGTSQAGEMLVDSLVLANIHP SVL PAISGLNSVDKRQAVGIVETWSVAACI SAADLAVKGSNVTLV RVHMAFGIGGK CYMVVAGDVL DVAAAVATASLAAGA KGLLVYASIIPRPHEAMWRQMVEG	
T33-09A SEQ ID NO:65	trímero	EEVVLITVPSALVAVKIAHALVEERLAA CVNIVPGLTSIYRWQGSVSDHELLLL VKTTTHAFPKLKERV KALHPYTVPEIV ALPIAEGNREYLDWLRENTG	
T33-09B SEQ ID NO:66	trímero	VRGIRGAITVEEDTPAAILAATIELLLK MLEANGIQSYEELAAVIFTVTEDLTSA FPAEAARLIGMHRVP LLSAREVPVPG SLPRVIRVLALWNTDTPQDRVRHVYL NEAVRLRPDLESAQ	
T33-15A SEQ ID NO:67	trímero	SKAKIGIVTVSDRASAGITADISGKAIL ALNLYLTSEWEPIYQVIPDEQDVIETT LIKMADEQDCCLIVTTGGTGPAKRDV TPEATEAVCDRMMPGFGELMRAESL KEVPTAILSRQTAGLRGDSLIVNLP GD PASISDCLLAVFPAIPYCIDLMEGPYLE CNEAMIKPFRPKAK	
T33-15B SEQ ID NO:68	trímero	VRGIRGAITVNSDTPTSIIITILLLEKM LEANGIQSYEELAAVIFTVTEDLTSAF PAEAARQIGMHRVP LLSAREVPVPGS	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		LPRVIRVLALWNTDTPQDRVRHVYLS EAVRLRPDLESAQ	
T33-21A SEQ ID NO:69	trímero	RITTKVGDKGSTRLFGGEEVWKDSPII EANGTLDELTSFIGEAKHYVDEEMKG ILEEIQNDIYKIMGEIGSKGKIEGISEER IAWLLKLILRYMEMVNLKSFVLPGGTL ESAKLDVCRTIARRALRKVLTVTREF GIGAEAAAYLLALSDLLFLLARVIEIEK NKLKEVRS	
T33-21B SEQ ID NO:70	trímero	PHLVIEATANLRLETSPGELLEQANKA LFASGQFGEADIKSRFVTLEAYRQGT AAVERAYLHACLSILDGRDIATRLLG ASLCAVLAEAVAGGGEEGVQVSVEV REMERLSYAKRVVARQR	
T33-28A SEQ ID NO:71	trímero	ESVNTSFLSPSLVTIRDFDNGQFAVL RIGRTGFPADKGDIDLCLDKMIGVRA AQIFLGDDTEDGFKGPHIRIRCVDIDD KHTYNAMVYVDLIVGTGASEVERETA EEEAKLALRVALQVDIADEHSCVTQF EMKLREELLSSDSFHPDKDEYYKDFL	
T33-28B SEQ ID NO:72	trímero	PVIQTFVSTPLDHHKRLLLAIIYRIVTR VVLGKPEDLVMMTFHDSTPMHFFGS TDPVACVRVEALGGYGPSEPEKVTSI VTAAITAVCGIVADRIFVLYFSPLHCG WNGTNF	
T33-31A SEQ ID NO:73	trímero	EEVVLITVPSALVAVKIAHALVEERLAA CVNIVPGLTSIYREEGSVVSDHELLLL VKTTTDAFPKCLKERVKELHPYEVPEIV	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		ALPIAEGNREYLDWLRENTG	
I53-50A ΔCys SEQ ID NO: 74	trímero	EELFKKHKIVAVLRANSVEEAIEKAVA VFAGGVHLIEITFTVPDADTVIKALSVL KEKGAIIGAGTVTSVEQ <u>A</u> ARKAVESGA EFIVSPHLDEEISQF <u>A</u> KEKGVFYMPG VMTPTELVKAMKLGHTILKLFPGEVV GPQFVKAMKGPFPPNVKFVPTGGVNL DNV <u>A</u> EWFKAGVLAVGVGSALVKGTP DEVREKAKAFVEKIRG <u>A</u> TE	
T33_dn2A SEQ ID NO: 75		NLAEKMYKAGNAMYRKQYTI ^{AI} IAYT LALLKDPNNAEAWYNLGNAAYKKGE YDEAIEAYQKALELDPNNAEAWYNLG NAYYKQGDYDEAIEYYKKALRLDPRN VDAIENLIEAEEKQG	
T33_dn2B SEQ ID NO: 76		EEAELAYLLGELAYKLGEYRIAIRAYRI ALKRDPNNAEAWYNLGNAAYYKQGDY REAIRYYLRALKLDPENAEAWYNLGN ALYKQGGKYDL ^{AI} IAYQAAL ^{ED} PNNAE AKQNLGNAKQKQG	
T33_dn5A SEQ ID NO: 77		NSAEAMYKMGNAAYKQGDYIL ^{AI} IAYL LALEKDPNNAEAWYNLGNAAYKQGD YDEAIEYYQKALELDPNNAEAWYNLG NAYYKQGDYDEAIEYYEKALELDPNN AEALKNLLEAIAEQD	
T33_dn5A SEQ ID NO: 78		TDPLAVILYI ^{AI} LKAEKSIARAKAAEALG KIGDERAVEPLIKALKDEDALVRAAAA DALGQIGDERAVEPLIKALKEEGLVR	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		ASAAIALGQIGDERAVQPLIKALTDER DLVRVAAVALGRIGDEKAVRPLIIVL KDEEGEVREAAAIALGSIGGERVRAA MEKLAERGTGFARKVAVNYLETHK	
T33_dn10A SEQ ID NO: 79		EEAELAYLLGELAYKLGEYRIAIRAYRI ALKRDPNNAEAWYNLGNAYYKQGDY DEAIEYYQKALELDPNNAEAWYNLGN AYYKQGDYDEAIEYYEKALELDPENL EALQNLLNAMDKQG	
T33_dn10B SEQ ID NO: 80		IEEVVAEMIDILAESSKKSIEELARAAD NKTTEKAVAEAEIEIARLATAAIQLIEAL AKNLASEEFMARAIASIAELAKKAIEAI YRLADNHTTDTFMARAIAAIANLAVTA ILAIAALASNHTTEEFMARAIASIAELA KKAIEAIYRLADNHTTDKFMAAAIEAIA LLATLAILAIALASNHTTEKFMARAIM AAILAAKAIEAIYRLADNHTSPTYIEKA IEAIEKIARKAIIEMLAKNITTEEYKE KAKKIIDIIRKLAKMAIKKLEDNRT	
I53_dn5A SEQ ID NO: 81	pentámero	KYDGSKLRIGILHARWNAEIIILALVLGA LKRLQEFVGKRENIIEIVPGSFELPY GSKLFVEKQKRLGKPLDAIPIGVLIK STMHFEYICDSTTHQLMKLNFELGIPV IFGVLTCLTDEQAEARAGLIEGKMHN HGEDWGAAAVEMATKFN	
I53_dn5B SEQ ID NO: 82	trímero	EEAELAYLLGELAYKLGEYRIAIRAYRI ALKRDPNNAEAWYNLGNAYYKQGRY REAIEYYQKALELDPNNAEAWYNLGN	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		AYYERGEYEEAIEYYRKALRLDPNNA DAMQNLLNAKMREE	
I53_dn5A.1 SEQ ID NO: 83	pentámero	KYDGSKLRIGILHARGNAEII LALVLGA LKRLQEFQVKRENIIIE TVPGSFELPY GSKLFVEKQKRLGKPLDAIPIGVLIRG STPHFDYIADSTTHQLMKLNFELGIPV IFGVITADTDEQAEARAGLIEGKMHNH GEDWGAAAVEMATKFN	
I53_dn5A.2 SEQ ID NO: 84	pentámero	KYDGSKLRIGILHARGNAEII LELVLGA LKRLQEFQVKRENIIIE TVPGSFELPY GSKLFVEKQKRLGKPLDAIPIGVLIRG STAHFDYIADSTTHQLMKLNFELGIPV IFGVLTTESEDEQAEERAGTKAGNHGE DWGAAAVEMATKFN	
I3-01 SEQ ID NO: 209		MKMEELFKKHKIVAVLRANSVEEAKK KALAVFLGGVHLIEITFTVPDADTVIKE LSFLKEMGAIIGAGTVTSVEQCRKAV ESGAEFIVSPHLDEEISQFCKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPG EVVGPFQVKAMKGPFNPVKFVPTGG VNLDNVCEWFKAGVLAVGVGSALVK GTPVEVAEKAKAFVEKIRGCTE	ninguno
I3-01 (M31) SEQ ID NO: 210		MKIEELFKKHKIVAVLRANSVEEAKKK ALAVFLGGVHLIEITFTVPDADTVIKEL SFLKEMGAIIGAGTVTSVEQCRKAVE SGAEFIVSPHLDEEISQFCKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPG	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		EVVGPQFVKAMKGPPNVKVVPTGG VNLDNVCEWFKAGVLAVGVGSALVK GTPVEVAEKAKAFVEKIRGCTE	
1WA3-ref SEQ ID NO: 211		MKMEELFKKHKIVAVLRANSVEEAKE KALAVFEGGVHLIEITFTVPDADTVIKE LSFLKEKGAIIGAGTVTSVEQCRKAVE SGAEFIVSPHLDEEISQFCKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFP EVVGPQFVKAMKGPPNVKVVPTGG VNLDNVCEWFKAGVLAVGVGSALVK GTPDEVREKAKAFVEKIRGCTE	
1WA3-1 SEQ ID NO: 212		(MK)MEELFKKHKIVAVLRANSVEEAK KKALAVFLGGVHLIEITFTVPDADTVIK ELSLKEMGAIIGAGTVTSVEQCRKA VESGAEFIVSPHLDEEISQFCKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFP GEVVGPQFVKAMKGPPNVKVVPTG GVNLDNVCEWFKAGVLAVGVGSALV KGTPVEVAEKAKAFVEKIRGCTE	ninguno
1WA3-2 SEQ ID NO: 213		(MK)IEELFKKHKIVAVLRANSVEEAKK KALAVFLGGVHLIEITFTVPDADTVIKE LSFLKEMGAIIGAGTVTSVEQCRKAV ESGAEFIVSPHLDEEISQFCKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFP EVVGPQFVKAMKGPPNVKVVPTGG VNLDNVCEWFKAGVLAVGVGSALVK GTPVEVAEKAKAFVEKIRGCTE	E33L, K61M, D187V, R190A
1WA3-3		(MK)MEELFKKHKIVAVLRANSVEEAK	E33L, K61M,

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
SEQ ID NO: 214		KKALAVFLGGVHLIEITFTVPDADTVIK ELSFLKEMGAIIGAGTVTSVEQARKAV ESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFP EVVGPQFVKAMKGPPNVKVVPTGG VNLDNVCEWFKAGVLAVGVGSALVK GTPVEVAEKAKAFVEKIRGCTE	D187V, R190A
1WA3-4 SEQ ID NO: 215		(MK)MEELFKKKHKIVAVLRANSVEEAK MKALAVFVGGVHLIEITFTVPDADTVI KELSFLKELGAIIGAGTVTSVEQCRKA VESGAEFIVSPHLDEEISQFCKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFP GEVVGPQFVKAMKGPPNVKVVPTG GVNLDNVCEWFKAGVLAVGVGSALV KGTIAEVAAKAAAFVEKIRGCTE	E33M, K61L, D187A, R190A
1WA3-5 SEQ ID NO: 216		(MK)MEELFKKKHKIVAVLRANSVEEAK KKALAVFLGGVDLIEITFTVPDADTVIK ELSFLKEMGAIIGAGTVTSVEQCRKA VESGAEFIVSPHLDEEISQFCKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFP GEVVGPQFVKAMKGPPNVKVVPTG GVNLDNVCEWFKAGVQAVGVGEALN KGTTPVEVAEKAKAFVEKIRGCTE	E33L, K61M, D187V, R190A
1WA3-6 SEQ ID NO: 217		(MK)MEELFKKKHKIVAVLRANSVEEAK KKALAVFMGGVDLIEITFTVPDADTVI KELSFLKELGAIIGAGTVTSVEQCRKA VESGAEFIVSPHLDEEISQFCKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFP	E33M, K61L, D187A, R190A

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		GEVVGPPQFVKAMKGPPFPNVK FVPTG GVNLDNVCEWFKAGVQAVGVGEALN KGTPAEVAEKAKAFVEKIRGCTE	
I3-01 H35D SEQ ID NO: 218		(METDTLLLWVLLLWVPGSTGDYKDE K)MEELFKKHKIVAVLRANSVEEAKKK ALAVFLGGVDLIEITFTVPDADTVIKEL SFLKEMGAIIGAGTVTSVEQCRKAVE SGAEFIVSPHLDEEISQFCKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPG EVVGPQFVKAMKGPPFPNVK FVPTGG VNLDNVCEWFKAGVLAVGVGSALVK GTPVEVAEKAKAFVEKIRGCTE(QKLI SEEDLHHHHHH)	
I3-01 K25D SEQ ID NO: 219		(METDTLLLWVLLLWVPGSTGDYKDE K)MEELFKKHKIVAVLRANSVEEAKKD ALAVFLGGVHLIEITFTVPDADTVIKEL SFLKEMGAIIGAGTVTSVEQCRKAVE SGAEFIVSPHLDEEISQFCKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPG EVVGPQFVKAMKGPPFPNVK FVPTGG VNLDNVCEWFKAGVLAVGVGSALVK GTPVEVAEKAKAFVEKIRGCTE(QKLI SEEDLHHHHHH)	
I3-01 K25N SEQ ID NO: 220		(METDTLLLWVLLLWVPGSTGDYKDE K)MEELFKKHKIVAVLRANSVEEAKKN ALAVFLGGVHLIEITFTVPDADTVIKEL SFLKEMGAIIGAGTVTSVEQCRKAVE SGAEFIVSPHLDEEISQFCKEKGVFY	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
		MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPNVKVVPTGGVNLDNVCEWFKAGVLAVGVGSALVKGTPVEVAEKAKAFVEKIRGCTE(QKLI SEEDLHHHHHH)	
I3-01 L171Q SEQ ID NO: 221		(METDTLLLWVLLLWVPGSTGDYKDEK)MEELFKKHKIVAVLRANSVEEAKKKALAVFLGGVHLIEITFTVPDADTVIKEL SFLKEMGAIIGAGTVTSVEQCRKAVE SGAEFIVSPHLDEEISQFCKEKGVFYMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPNVKVVPTGGVNLDNVCEWFKAGVQAVGVGSALVKGTPVEVAEKAKAFVEKIRGCTE(QKLI SEEDLHHHHHH)	
I3-01 L171Q/S177E/V 180N SEQ ID NO: 222		(METDTLLLWVLLLWVPGSTGDYKDEK)MEELFKKHKIVAVLRANSVEEAKKKALAVFLGGVHLIEITFTVPDADTVIKEL SFLKEMGAIIGAGTVTSVEQCRKAVE SGAEFIVSPHLDEEISQFCKEKGVFYMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPNVKVVPTGGVNLDNVCEWFKAGVQAVGVGEALNKGTPVEVAEKAKAFVEKIRGCTE(QKLI SEEDLHHHHHH)	
I3-01 “mutaciones de secreción”		(METDTLLLWVLLLWVPGSTGDYKDEK)MEELFKKHKIVAVLRANSVEEAKKKALAVFLGGVDLIEITFTVPDADTVIKEL	

Nombre	Componente multímero	Secuencia de aminoácidos	Residuos de interfaz identificados
(H35D/L171Q/S 177E/V180N) SEQ ID NO: 223		SFLKEMGAIIGAGTVTSVEQCRKAVE SGAEFIVSPHLDEEISQFCKEKGVFY MPGVMTPTELVKAMKLGHTILKLFP EVVGPQFVKAMKGPFNPVKFVPTGG VNLDNVCEWFKAGVQAVGVGEALNK GTPVEVAEKAKAFVEKIRGCTE(QKLI SEEDLHHHHH)	
I3-01 “interior negativo” SEQ ID NO: 224		(METDTLLLWVLLLWVPGSTGD)MEE LFKEHKIVAVLRANSVEEAKKKALAVF LGGVHLIEITFTVPDADTVIKELSFLKE MGAIIGAGTVTSVEQAREAVESGAEFI VSPHLDEEISQFAKEEGVFYMPGVMT PTELVKAMKLGHTILKLFPGEVVGPQ FVEAMKGPFNPVKFVPTGGVNLCNV AEWFEAGVLAVGVGSALVEGTPVEV AEKAKAFVEKIEGATE(QKLISEEDLH HHHHH)	
I3-01 “interior negativo con mutaciones de secreción” SEQ ID NO: 225		(METDTLLLWVLLLWVPGSTGD)MEE LFKEHKIVAVLRANSVEEAKKKALAVF LGGVDLIEITFTVPDADTVIKELSFLKE MGAIIGAGTVTSVEQAREAVESGAEFI VSPHLDEEISQFAKEEGVFYMPGVMT PTELVKAMKLGHTILKLFPGEVVGPQ FVEAMKGPFNPVKFVPTGGVNLDNV AEWFEAGVQAVGVGEALNEGTPVEV AEKAKAFVEKIEGATE(QKLISEEDLH HHHHH)	

[0206] La **Tabla 12** proporciona la secuencia de aminoácidos de un primer dominio de repetición y un segundo dominio de repetición de modalidades de la presente descripción. En cada caso, los pares de secuencias forman juntos un multímero I53 con simetría icosaédrica. La columna de la derecha de la **Tabla 12** identifica los números de residuos en cada polipéptido ilustrativo

5 que se identificaron como presentes en la interfaz de las nanoestructuras proteicas ensambladas resultantes (es decir, «residuos de interfaz identificados»). Como puede observarse, el número de residuos de interfaz para los polipéptidos ilustrativos de las SEQ ID NO: 23-56 varía de 4-13. En diversas modalidades, un primer elemento de repetición y segundo elemento de repetición comprenden una secuencia de aminoácidos que es al menos 75 %, al

10 menos 80 %, al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % idéntica en toda su longitud, e idéntica al menos en 1, al menos en 2, al menos en 3, al menos en 4, al menos en 5, al menos en 6, al menos en 7, al menos en 8, al menos en 9, al menos en 10, al menos en 11, al menos en 12 o al menos en 13 posiciones de interfaz

15 identificadas (dependiendo del número de residuos de interfaz para un polipéptido dado), a la secuencia de aminoácidos de un polipéptido seleccionado del grupo que consiste en las SEQ ID NO: 23-56. Las SEQ ID NO: 57-84 representan otras secuencias de aminoácidos de un primer elemento de repetición y un segundo elemento de repetición de modalidades de la presente descripción. En otras modalidades, un primer elemento de repetición y/o segundo

20 elemento de repetición comprenden una secuencia de aminoácidos que es al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % idéntica en toda su longitud, e idéntica 20 %, al menos 25 %, al menos 33 %, al menos 40 %, al menos 50 %, al menos 60 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos

25 90 % o 100 % de las posiciones de interfaz identificadas, a la secuencia de aminoácidos de un polipéptido seleccionado del grupo que consiste en las SEQ ID NO: 23-84, 209-225.

[0207] Tal como sucede con las proteínas en general, se espera que los polipéptidos toleren alguna variación en las secuencias diseñadas sin interrumpir el ensamblaje posterior en nanoestructuras proteicas: particularmente, cuando dicha variación comprende sustituciones

30 conservadoras de aminoácidos. Tal como se utiliza en la presente, «sustitución conservadora de aminoácidos» significa que: los aminoácidos hidrófobos (Ala, Cys, Gly, Pro, Met, Val, Ile,

Leu) se sustituyen con otros aminoácidos hidrófobos; los aminoácidos hidrófobos con cadenas laterales voluminosas (Phe, Tyr, Trp) se sustituyen con otros aminoácidos hidrófobos con cadenas laterales voluminosas; los aminoácidos con cadenas laterales cargadas positivamente (Arg, His, Lys) se sustituyen con otros aminoácidos con cadenas laterales cargadas positivamente; los aminoácidos con cadenas laterales cargadas negativamente (Asp, Glu) se sustituyen con otros aminoácidos con cadenas laterales cargadas negativamente; y los aminoácidos polares con cadenas laterales sin carga (Ser, Thr, Asn, Gln) se sustituyen con otros aminoácidos polares con cadenas laterales sin carga.

[0208] En diversas modalidades de las nanoestructuras proteicas de la invención, un primer elemento de repetición y un segundo elemento de repetición, o viceversa, comprenden polipéptidos con la secuencia de aminoácidos seleccionada de los siguientes pares, o versiones modificadas de estos (*es decir*, modificaciones permisibles según se describe para los polipéptidos de la invención: polipéptidos aislados que comprenden una secuencia de aminoácidos que es al menos el 75 %, al menos el 80 %, al menos el 81 %, al menos el 82 %, al menos el 83 %, al menos el 84 %, al menos el 85 %, al menos el 86 %, al menos el 87 %, al menos el 88 %, al menos el 89 %, al menos el 90 %, al menos el 91 %, al menos el 92 %, al menos el 93 %, al menos el 94 %, al menos el 95 %, al menos el 96 %, al menos el 97 %, al menos el 98 %, al menos el 99 %, o 100 % idéntica en su longitud, y/o idéntica al menos en una posición de interfaz identificada, a la secuencia de aminoácidos indicada por la SEQ ID NO):

SEQ ID NO: 23 y SEQ ID NO: 24 (I53-34A y I53-34B);
 SEQ ID NO: 25 y SEQ ID NO: 26 (I53-40A y I53-40B);
 SEQ ID NO: 25 y SEQ ID NO: 46 (I53-40A y I53-40B.1);
 SEQ ID NO: 45 y SEQ ID NO: 26 (I53-40A.1 y I53-40B);
 SEQ ID NO: 57 y SEQ ID NO: 58 (género I53-40A y género I53-40B);
 SEQ ID NO: 27 y SEQ ID NO: 28 (I53-47A y I53-47B);
 SEQ ID NO: 27 y SEQ ID NO: 49 (I53-47A y I53-47B.1);
 SEQ ID NO: 27 y SEQ ID NO: 50 (I53-47A y I53-47B.1NegT2);
 SEQ ID NO: 47 y SEQ ID NO: 28 (I53-47A.1 y I53-47B);
 SEQ ID NO: 47 y SEQ ID NO: 49 (I53-47A.1 y I53-47B.1);
 SEQ ID NO: 47 y SEQ ID NO: 50 (I53-47A.1 y I53-47B.1NegT2);
 SEQ ID NO: 48 y SEQ ID NO: 28 (I53-47A.1NegT2 y I53-47B);

- SEQ ID NO: 48 y SEQ ID NO: 49 (I53-47A.1NegT2 y I53-47B.1);
 SEQ ID NO: 48 y SEQ ID NO: 50 (I53-47A.1NegT2 y I53-47B.1NegT2);
 SEQ ID NO: 59 y SEQ ID NO: 60 (género I53-47A y género I53-47B);
 SEQ ID NO: 29 y SEQ ID NO: 30 (I53-50A y I53-50B);
- 5 SEQ ID NO: 29 y SEQ ID NO: 54 (I53-50A y I53-50B.1);
 SEQ ID NO: 29 y SEQ ID NO: 55 (I53-50A y I53-50B.1NegT2);
 SEQ ID NO: 29 y SEQ ID NO: 56 (I53-50A y I53-50B.4PosT1);
 SEQ ID NO: 51 y SEQ ID NO: 30 (I53-50A.1 y I53-50B);
 SEQ ID NO: 51 y SEQ ID NO: 54 (I53-50A.1 y I53-50B.1);
- 10 SEQ ID NO: 51 y SEQ ID NO: 55 (I53-50A.1 y I53-50B.1NegT2);
 SEQ ID NO: 51 y SEQ ID NO: 56 (I53-50A.1 y I53-50B.4PosT1);
 SEQ ID NO: 52 y SEQ ID NO: 30 (I53-50A.1NegT2 y I53-50B);
 SEQ ID NO: 52 y SEQ ID NO: 54 (I53-50A.1NegT2 y I53-50B.1);
 SEQ ID NO: 52 y SEQ ID NO: 55 (I53-50A.1NegT2 y I53-50B.1NegT2);
- 15 SEQ ID NO: 52 y SEQ ID NO: 56 (I53-50A.1NegT2 y I53-50B.4PosT1);
 SEQ ID NO: 53 y SEQ ID NO: 30 (I53-50A.1PosT1 y I53-50B);
 SEQ ID NO: 53 y SEQ ID NO: 54 (I53-50A.1PosT1 y I53-50B.1);
 SEQ ID NO: 53 y SEQ ID NO: 55 (I53-50A.1PosT1 y I53-50B.1NegT2);
 SEQ ID NO: 53 y SEQ ID NO: 56 (I53-50A.1PosT1 y I53-50B.4PosT1);
- 20 SEQ ID NO: 61 y SEQ ID NO: 62 (género I53-50A y género I53-50B);
 SEQ ID NO: 31 y SEQ ID NO: 32 (I53-51A y I53-51B);
 SEQ ID NO: 33 y SEQ ID NO: 34 (I52-03A y I52-03B);
 SEQ ID NO: 35 y SEQ ID NO: 36 (I52-32A y I52-32B);
 SEQ ID NO: 37 y SEQ ID NO: 38 (I52-33A y I52-33B)
- 25 SEQ ID NO: 39 y SEQ ID NO: 40 (I32-06A y I32-06B);
 SEQ ID NO: 41 y SEQ ID NO: 42 (I32-19A y I32-19B);
 SEQ ID NO: 43 y SEQ ID NO: 44 (I32-28A y I32-28B);
 SEQ ID NO: 45 y SEQ ID NO: 46 (I53-40A.1 y I53-40B.1);
 SEQ ID NO: 63 y SEQ ID NO: 64 (T32-28A y T32-28B);
- 30 SEQ ID NO: 65 y SEQ ID NO: 66 (T33-09A y T33-09B);
 SEQ ID NO: 67 y SEQ ID NO: 68 (T33-15A y T33-15B);

SEQ ID NO: 69 y SEQ ID NO: 70 (T33-21A y T33-21B);

SEQ ID NO: 71 y SEQ ID NO: 72 (T33-28A y T32-28B); y

SEQ ID NO: 73 y SEQ ID NO: 66 (T33-31A y T33-09B (también denominado T33-31B)).

[0209] En algunas modalidades, los dominios de ensamblaje son I53-50A (trímero, opcionalmente unido a un fragmento antigénico de una glucoproteína Spike del coronavirus como se describió anteriormente) e I53-50B (pentámero). Las nanoestructuras I53_dn5 se describen en el documento US 2022/0072120 A1, cuyo contenido se incorpora mediante referencia. Las variantes de I53_dn5 pueden incluir una o más sustituciones de aminoácidos, como C94A, C119A, W18G, K84R, M88P, E91D, L117I o L120D (conjuntamente «I53_dn5A.1»; Ueda et ál. eLife 9:e57659 (2020) o A25E, M88A, C119T, L120E, A127E, L131T, I132K, E133A, o una delección de las posiciones 135-137 («I53_dn5A.2»; Wang et ál. bioRxiv 2022.08.04.502842).

[0210] En algunas modalidades, el uno o más fragmentos antigénicos de glucoproteína Spike de coronavirus, o fragmentos antigénicos de estas, se expresan como una proteína de fusión con un primer dominio de ensamblaje. En algunas modalidades, el primer dominio de ensamblaje y el fragmento antigénico de una glucoproteína Spike de coronavirus están unidos por una secuencia enlazadora.

[0211] Los ejemplos no taxativos de complejos de proteínas diseñados útiles en nanoestructuras proteicas de la presente descripción incluyen aquellos descritos en la patente estadounidense n.º 9.630.994; la publicación de patente internacional n.º WO2018187325A1; la publicación de patente estadounidense n.º 2018/0137234 A1; la publicación de patente estadounidense n.º 2019/0155988A2, cada una de las cuales se incorpora a la presente en su totalidad.

[0212] En diversas modalidades de las nanoestructuras proteicas de la descripción, los dominios de ensamblaje son polipéptidos con la secuencia de aminoácidos seleccionada de los siguientes pares, o versiones modificadas de estos (*es decir*, modificaciones permisibles según se describe para los polipéptidos de la invención: polipéptidos aislados que comprenden una secuencia de aminoácidos que es al menos el 75 %, al menos el 80 %, al menos el 85 %, al menos el 90 %, al menos el 95 %, al menos el 99 %, o 100 % en su longitud, y/o idéntica al menos en una posición de interfaz identificada, a la secuencia de aminoácidos indicada por la SEQ ID NO):

SEQ ID NO: 75 y SEQ ID NO: 76 (T33_dn2A y T33_dn2B);
 SEQ ID NO: 77 y SEQ ID NO: 78 (T33_dn5A y T33_dn5B);
 SEQ ID NO: 79 y SEQ ID NO: 80 (T33_dn10A y T33_dn10B); o
 SEQ ID NO: 81 y SEQ ID NO: 82 (I53_dn5A y I53_dn5B).

5 **[0213]** En otro aspecto, la descripción proporciona una nanoestructura de proteína autoensamblable que comprende un primer componente que comprende un polipéptido de coronavirus recombinante descrito en la presente y, opcionalmente, un segundo componente que comprende una segunda proteína.

10 **[0214]** En algunas modalidades, la nanoestructura de proteína del polipéptido de coronavirus recombinante comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

15 **[0215]** En algunas modalidades, la nanoestructura proteica del polipéptido de coronavirus recombinante comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en relación con una secuencia de referencia según la SEQ ID NO: 1.

[0216] En algunas modalidades, el componente polipeptídico de nanoestructura proteica y el polipéptido de coronavirus recombinante están acoplados de forma covalente.

20 **[0217]** En algunas modalidades, el componente polipeptídico de nanoestructura de proteína y el polipéptido de coronavirus recombinante están acoplados de forma covalente.

[0218] En algunas modalidades, la nanoestructura de proteína comprende una proteína de fusión que incluye el polipéptido de coronavirus recombinante, un enlazador y el polipéptido componente de la nanoestructura de proteína.

25 **[0219]** En algunas modalidades, la nanoestructura de proteína comprende una proteína de fusión que incluye, en el orden de extremo N a C, el polipéptido de coronavirus recombinante, un enlazador y el polipéptido componente de la nanoestructura de proteína.

30 **[0220]** En algunas modalidades, la nanoestructura de proteína del polipéptido de coronavirus recombinante comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

[0221] En algunas modalidades, la nanoestructura de proteína de la proteína de fusión comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

[0222] En algunas modalidades, el primer componente de la nanoestructura de proteína comprende I53-50A. En algunas modalidades, el primer componente de la nanoestructura de proteína comprende I53-34A. En algunas modalidades, el primer componente de la nanoestructura de proteína comprende I53-40B. En algunas modalidades, el primer componente de la nanoestructura de proteína comprende I53-47A. En algunas modalidades, el primer componente de la nanoestructura de proteína comprende I53-51A.

[0223] En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína I3-01/MI3. En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína I53-50AB.

[0224] En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de ferritina. En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de encapsulina. En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de cápside de bacteriófago CP3. En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de cápside de bacteriófago Q β . En algunas modalidades, la nanoestructura de proteína es una nanoestructura de proteína de cápside de bacteriófago AP205. En algunas modalidades, la nanoestructura proteica comprende un primer polipéptido y un segundo polipéptido seleccionados entre los descritos en WO 2021252688, que se incorpora mediante esta referencia en la presente en su totalidad.

[0225] En algunas modalidades, el polipéptido recombinante de la nanoestructura de proteína comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 11-19.

[0226] En algunas modalidades, el componente polipeptídico de la nanoestructura de proteína comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 20-21.

[0227] En algunas modalidades, la nanoestructura de proteína comprende un segundo componente polipeptídico.

[0228] En algunas modalidades, el segundo polipéptido componente es I53-50B. En algunas modalidades, el segundo polipéptido componente es I53-34B. En algunas modalidades, el segundo polipéptido componente es I53-40B. En algunas modalidades, el segundo polipéptido componente es I53-47B. En algunas modalidades, el segundo polipéptido componente es I53-51B. En algunas modalidades, el segundo componente de la nanoestructura de proteína comprende I53-50B.4PosT1.

[0229] En algunas modalidades, el segundo componente polipeptídico comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o 100 % idéntica a las SEQ ID NO: 22.

Nanoestructuras basadas en ferritina

[0230] En algunas modalidades, el dominio de ensamblaje es un polipéptido de ferritina. En algunas modalidades, el dominio de ensamblaje de una nanoestructura proteica de ferritina comprende una secuencia polipeptídica al menos 70 %, al menos 80 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % idéntica a cualquiera de las siguientes secuencias:

MLSKDIKLLNEQVNKEMNSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFL
NENNVVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAISKDHDATFNFLQWY

VAEQHEEEVLFKDILDKVELIGNENHGLYLADQYVKGIASRKS. (SEQ ID NO: 204)

MLKPEMIEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAFLRRHAQEEMTHMQRLFD
YLTDTGNLPRINTVESPFAYSSLDLDFQETYKHEQLITQKINELAHAAMTNQDYPTFNFLQ
WYVSEQHEEEKLFKSIIDKLSLAGKSGEGLYFIDKELSTLDAQN. (SEQ ID NO: 205)

NFHQDCEAGLNRTVNLKFHSSYVYLSMASYFNRDDVALSNFAKFFRERSEEEKEHAekli
EYQNQRGGRVFLQSVEKPERDDWANGLEALQTALKLQKSVNQALLDLHAVAADKSDPH
MTDFLESPYLSESVETIKKLGDHITSLKKLWSSHPGMAEYLFNKHTLG. (SEQ ID NO: 206)

QFSKDIEKLLNEQVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAekliIFL
NENNVVQLTSISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWY

VAEQHEEEVLFKDILDKIELIGNENHGLYLADQYVKGIASRKSGS. (SEQ ID NO: 207)

SGESQVRQNFKPEMEEKLNEQMNLELYSSLLYQQMSAWCSYHTFEGAAFLRRHAQEE
MTHMQRLFDYLTDTGNLPRINTVESPFAYSSLDLDFQETYKHEQLITQKINELAHAAMTN
QDYPTFNFLQWYVSEQHEEEKLFKSIIDKLSLAGKSGEGLYFIDKELSTLDGS. (SEQ ID
NO: 208)

Otras nanoestructuras o nanopartículas

[0231] En algunas modalidades, los ectodominios descritos en la presente se muestran en cualquier nanoestructura o nanopartícula conocida en la técnica. Las nanoestructuras y nanopartículas ilustrativas incluyen, entre otros, partículas similares a virus (VLP, por sus siglas en inglés) del virus del papiloma humano (HPV, por sus siglas en inglés), VLP de Chikungunya, VLP de proteína de cápside AP205, VLP de un fago (por ejemplo, bacteriófago), una nanoestructura proteica de encapsulina, una nanoestructura proteica de cápside de bacteriófago CP3, una nanoestructura proteica de cápside de bacteriófago Q β . La visualización en estas y otras plataformas puede realizarse creando una proteína de fusión del ectodominio con una proteína relevante del sistema, mediante química de bioconjugados (por ejemplo, SpyCatcher), u otros medios conocidos en la técnica. La nanoestructura proteica puede ser una nanopartícula de lumazina sintasa como se describe, *por ejemplo*, en Geng et al. PLoS Pathog. 17(9):e1009897 (2021). La nanoestructura proteica puede ser una nanopartícula de ferritina como se describe, *por ejemplo*, en Joyce et al. bioRxiv 2021.05.09.443331 y en la pub. de pat. de EE. UU. núm. 2019/0330279 A1.

Polinucleótidos

[0232] Se proporcionan en la presente polinucleótidos que codifican los polipéptidos o las

nanoestructuras de proteínas descritas en la presente. La secuencia de polinucleótidos puede comprender ARN o ADN. Los polinucleótidos aislados se refieren a aquellos que se han eliminado de sus secuencias de ácidos nucleicos circundantes normales en el genoma o en secuencias de ADNc. Tales secuencias de polinucleótidos pueden comprender secuencias adicionales útiles para promover la expresión y/o purificación de la proteína codificada, que incluyen, entre otros, secuencias poliA, secuencias Kozak modificadas, y secuencias que codifican etiquetas epitópicas, señales de exportación, y señales secretoras, señales de localización nuclear y señales de localización de la membrana plasmática. Será evidente para los expertos en la técnica, basándose en las enseñanzas del presente documento, qué secuencias de ácido nucleico codificarán las proteínas de la descripción.

[0233] En un aspecto adicional, la presente descripción proporciona vectores de expresión recombinantes que comprenden el ácido nucleico aislado de cualquier realización o combinación de realizaciones de la divulgación unido operativamente a una secuencia de control adecuada. "Vector de expresión recombinante" incluye vectores que unen operativamente una región codificante de ácido nucleico o un gen a cualquier secuencia de control capaz de efectuar la expresión del producto génico. Las "secuencias de control" unidas operativamente a las secuencias de ácidos nucleicos de la divulgación son secuencias de ácidos nucleicos capaces de efectuar la expresión de las moléculas de ácidos nucleicos. Las secuencias de control no necesitan ser contiguas a las secuencias de ácido nucleico, siempre que funcionen para dirigir su expresión. Así, por ejemplo, pueden estar presentes secuencias intermedias no traducidas pero transcritas entre una secuencia promotora y las secuencias de ácidos nucleicos, y la secuencia promotora todavía puede considerarse "operativamente unida" a la secuencia codificante. Otras secuencias de control de este tipo incluyen, entre otras, señales de poliadenilación, señales de terminación y sitios de unión de ribosomas. Dichos vectores de expresión pueden ser de cualquier tipo conocido en la técnica, incluidos, entre otros, vectores de expresión basados en plásmidos y virus. La secuencia de control utilizada para impulsar la expresión de las secuencias de ácidos nucleicos descritas en un sistema mamífero puede ser constitutiva (impulsada por cualquiera de una variedad de promotores, incluidos, entre otros, CMV, SV40, RSV, actina, EF) o inducible (impulsada por cualquiera de una cantidad de promotores inducibles, incluidos, entre otros, tetraciclina, ecdisona, esteroides sensibles). La construcción de vectores de expresión para su uso en la transfección de células procariotas

también es bien conocida en la técnica y, por lo tanto, puede lograrse mediante técnicas estándar. (Véase, por ejemplo, Sambrook, Fritsch y Maniatis, en: *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 1989; *Gene Transfer and Expression Protocols*, págs. 109-128, ed., E.J. Murray, The Science and Research Institute. E.J. Murray, The Humana Press Inc., Clifton, N.J.), y el Ambion 1998 Catalog (Ambion, Austin, TX). El vector de expresión puede replicarse en los organismos hospedadores como un episoma o por integración en el ADN cromosómico del hospedador. Los vectores de expresión y hospedadores adecuados son conocidos en la técnica. En algunas modalidades, el vector de expresión comprende un plásmido. Sin embargo, la descripción pretende incluir otros vectores de expresión que cumplen funciones equivalentes, como los vectores virales.

[0234] En otro aspecto, la presente descripción proporciona células huésped que han sido transfectadas o transducidas con los vectores de expresión recombinantes divulgados en este documento, en donde las células huésped pueden ser procariotas o eucariotas. Las células pueden ser transfectadas o transducidas de forma transitoria o estable. Dicha transfección o transducción de vectores de expresión en células procariotas y eucariotas se puede lograr a través de cualquier técnica conocida en la técnica, incluyendo, entre otras, transformaciones bacterianas estándar, coprecipitación con fosfato de calcio, electroporación o transfección mediada por liposomas, mediada por DEAE dextrano, mediada por policationicos o mediada por virus. (Véase, por ejemplo, *Molecular Cloning: A Laboratory Manual* (Sambrook et al., 1989, Cold Spring Harbor Laboratory Press; *Culture of Animal Cells: A Manual of Basic Technique*, 2.^a edición, R. I. Freshney, 1987. Liss, Inc. New York, NY).

[0235] En algunas modalidades, el polinucleótido es un ARN mensajero (ARNm). Los métodos para generar un polinucleótido ya sea por síntesis química o por transcripción *in vitro* (IVT) (para ARNm) son bien conocidos en la técnica.

[0236] La estructura básica de un ARNm codificante puede parecerse a un ARNm eucariota “maduro” y puede incluir algunas o todas las siguientes características que incluyen (i) un marco de lectura abierto (ORF) codificante de proteínas, flanqueado por (ii) regiones no traducidas (UTR) 5' y 3', y en los extremos (iii) una estructura de tapa 5' de 7-metil guanosina y (iv) una cola de poli(A) 3'. Las características estructurales no codificantes se pueden optimizar individualmente para modular la estabilidad del ARNm, la eficiencia de la traducción y la inmunogenicidad. Al incorporar nucleósidos modificados, se pueden producir transcripciones de

ARNm denominadas “ARNm modificado con nucleósidos” con una actividad inmunoestimulante reducida y, por lo tanto, se puede obtener un perfil de seguridad mejorado. Además, los nucleósidos modificados permiten el diseño de vacunas de ARNm con una estabilidad y capacidad de traducción fuertemente mejoradas, ya que pueden evitar las vías antivirales directas que son inducidas por los FN tipo Iy están programados para degradar e inhibir el ARNm invasor. Por ejemplo, la sustitución de uridina por pseudouridina reduce la actividad de la 2'-5'-oligoadenilato sintetasa, que regula la escisión del ARNm por la RNasa L. Además, se miden actividades más bajas para la proteína quinasa R, una enzima que está asociada con la inhibición del proceso de traducción del ARNm.

10 **[0237]** Los polinucleótidos de la descripción pueden incluir una o más nucleobases, nucleósidos, nucleótidos o combinaciones de estos modificados (*por ejemplo*, alterados o alternativos). Los polinucleótidos pueden incluir cualquier modificación o alteración útil, tal como en la nucleobase, el azúcar o el enlace internucleósido (*por ejemplo*, a un fosfato de enlace/a un enlace fosfodiéster/a la cadena principal fosfodiéster). En algunas modalidades, están presentes alteraciones (*por ejemplo*, una o más alteraciones) en cada una de la nucleobase, el azúcar y el enlace internucleósido. Las alteraciones de acuerdo con la presente descripción pueden ser alteraciones de polinucleótidos, *por ejemplo*, la sustitución del 2'—OH del ribofuranosilo de un anillo de ARN por 2'—H, ácidos nucleicos de treosa (TNA), ácidos nucleicos de glicol (GNA), ácidos nucleicos peptídicos (PNA), ácidos nucleicos bloqueados (LNA) o híbridos de estos.

20 **[0238]** El polinucleótido puede incluir una estructura de tapa 5'. La estructura de la tapa 5' de un polinucleótido participa en la exportación nuclear y en el aumento de la estabilidad del polinucleótido y se une a la proteína de unión a la tapa del ARNm (CBP). Se puede proporcionar una 5'-UTR como región flanqueante al ARNm. En algunas modalidades, la secuencia de polinucleótidos está optimizada en codones. El polinucleótido puede incluir una “secuencia poliA” o una “señal de poliadenilación”, términos que se usan indistintamente.

[0239] En algunas modalidades, el ARNm comprende entre 100 bases y 5 kilobases (kb).

[0240] En algunas modalidades, el ARNm comprende al menos 100 bases, al menos 200 bases, al menos 400 bases, al menos 600 bases o al menos 800 bases.

25 **[0241]** En algunas modalidades, el ARNm comprende al menos 1 kb, al menos 2 kb, al menos 3 kb, al menos 4 kb o al menos 5 kb.

[0242] En algunas modalidades, el ARNm comprende entre 100 bases y 5 kb, entre 100 bases

y 3 kb, entre 100 bases y 2 kb, entre 100 bases y 1 kb.

[0243] En algunas modalidades, el ARNm comprende entre 1 kb y 5 kb, entre 1 kb y 4 kb, o entre 1 kb y 3 kb.

Secuencia de señal

- 5 **[0244]** En algunas modalidades, el polipéptido codificado incluye una región peptídica que es un péptido señal. Las secuencias señal son conocidas en la técnica. La secuencia señal puede ser la secuencia señal nativa o puede sustituirse por otra secuencia señal. La función de las secuencias señal es incitar a la célula a translocar la proteína, normalmente a la membrana celular. El núcleo de la secuencia señal suele contener un largo tramo de aminoácidos
- 10 hidrófobos (de aproximadamente 5 a 16 residuos de longitud) que tiende a formar una hélice alfa simple y que también se conoce como "región h". Además, muchas secuencias señal comienzan con un corto tramo de aminoácidos cargados positivamente, lo que puede ayudar a reforzar la topología adecuada del polipéptido durante la translocación mediante lo que se conoce como la regla del interior positivo. Debido a su proximidad al extremo N, se denomina
- 15 «región n». Al final de la secuencia señal suele haber un tramo de aminoácidos que es reconocido y escindido por la peptidasa señal.

[0245] En algunas modalidades, el polipéptido no incluye una secuencia señal.

[0246] Se proporcionan ejemplos no limitativos de péptidos señal en la **Tabla 13**

Tabla 13.

Secuencia	Longitud	SEQ ID NO:
MELLILKANAITTILTAVTFCFASG	25	104
MSWKVMIIISLLITPQHGL	19	105
MPISILLIITTMIMASHC	18	106
MKAILVVLLYTFTTANA	17	107
MFVFLVLLPLVSSQC	15	108
MGILPSPGMPALLSLVSLLSVLLMGCVAETGS	32	109
MGILPSPGMPALLSLVSLLSVLLMGCVAETGT	32	110

20

Dominios de trimerización

[0247] En algunas modalidades, el polipéptido puede comprender un dominio de trimerización, como FoldOn o una trimerización GCN4. En algunas modalidades, la secuencia enlazadora

comprende un FoldOn, en donde la secuencia FoldOn es GYIPEAPRDGQAYVRKDGWVLLSTFL (SEQ ID NO: 226).

[0248] En algunas modalidades, el polipéptido puede comprender un dominio de trimerización, en donde la secuencia del dominio de trimerización es DKIEEILSKIYHIENEIARIKKLIGE (GEN) (SEQ ID NO: 227). En algunas modalidades, el polipéptido puede comprender un dominio de trimerización, en donde la secuencia del dominio de trimerización es EKFHQIEKEFSEVEGRIQDLEK (HA) (SEQ ID NO: 228).

[0249] En algunas modalidades, el polipéptido puede comprender un dominio de trimerización, en donde la secuencia del dominio de trimerización es EDKIEEILSKIYHIENEIARIKKLIGEA (cierre de isoleucinas en espiral) (SEQ ID NO: 229).

[0250] En algunas modalidades, el polipéptido puede comprender un dominio de trimerización, en donde la secuencia de dominio de trimerización es GSGYIPEAPRDGQAYVRKDGWVLLSTFL (fibrina de bacteriófago T4) (SEQ ID NO: 230).

[0251] En algunas modalidades, una secuencia de trimerización es RMKQIEDKIEEILSKIYHIENEIARIKKLIGEA (GCN4) (SEQ ID NO: 231). En algunas modalidades, un dominio de trimerización es una variante de GCN4. En algunas modalidades, la secuencia variante de GCN4 es RMKQIEDKIEEILSKIYHIENEIARIKKLIGERGGR (SEQ ID NO: 232), RMKQIEDKIEEILSKIYHIENEIARIKKLIGNRTGGR (SEQ ID NO: 233), RMKQIEDKIENITSKIYHIENEIARIKKLIGNRTGGR (SEQ ID NO: 234), RMKQIEDKIEEILSKIYNITNEIARIKKLIGNRTGGR (SEQ ID NO: 235), o RMKQIEDKIENITSKIYNITNEIARIKKLIGNRTGGR (SEQ ID NO: 236).

Composiciones farmacéuticas

[0252] En otro aspecto, la descripción proporciona una composición farmacéutica que comprende los polipéptidos, las nanoestructuras de proteína, o los polinucleótidos descritos anteriormente.

[0253] Las composiciones farmacéuticas descritas en el presente documento pueden administrarse en combinación con uno o más agentes terapéuticos adicionales. Los agentes terapéuticos adicionales pueden incluir, pero no se limitan a, antibióticos o agentes antibacterianos, agentes antieméticos, agentes antifúngicos, agentes antiinflamatorios, agentes antivirales, agentes inmunomoduladores, citocinas, antidepresivos, hormonas, agentes alquilantes, antimetabolitos, antibióticos antitumorales, agentes antimitóticos, inhibidores de la

topoisomerasa, agentes citostáticos, agentes antiinvasión, agentes antiangiogénicos, inhibidores de la función del factor de crecimiento, inhibidores de la replicación viral, inhibidores de enzimas virales, agentes anticancerígenos, interferones α , interferón β , ribavirina, hormonas y otros moduladores del receptor tipo Toll, inmunoglobulinas (Ig) y anticuerpos que modulan la función de Ig (como anti-IgE (omalizumab)).

[0254] Una composición puede ser estéril y/o libre de pirógenos. Las composiciones farmacéuticas pueden ser isotónicas con respecto a los humanos.

[0255] Las composiciones farmacéuticas pueden incluir uno o más agentes antimicrobianos, particularmente cuando están envasadas en formato de dosis múltiples. Los agentes antimicrobianos incluyen, entre otros, tiomersal y 2-fenoxietanol. En algunas modalidades, puede ser deseable usar un conservante sin mercurio o ningún conservante.

[0256] En algunas modalidades, el amortiguador en la composición de la vacuna es un amortiguador Tris, un amortiguador de histidina, un amortiguador de fosfato, un amortiguador de citrato o un amortiguador de acetato. La composición farmacéutica también puede incluir un lioprotector, por ejemplo, sacarosa, sorbitol o trehalosa. En algunas modalidades, la composición incluye un conservante, por ejemplo, cloruro de benzalconio, bencetonio, clorhexidina, fenol, m-cresol, alcohol bencílico, metilparabeno, propilparabeno, clorobutanol, o-cresol, p-cresol, clorocresol, nitrato fenilmercúrico, timerosal, ácido benzoico y diversas mezclas de estos. En otras modalidades, la composición farmacéutica incluye un agente de carga, como glicina. En aun otras modalidades, la composición incluye un tensioactivo, por ejemplo, polisorbato-20, polisorbato-40, polisorbato-60, polisorbato-65, polisorbato-80, polisorbato-85, poloxámero-188, monolaurato de sorbitán, monopalmitato de sorbitán, monoestearato de sorbitán, monooleato de sorbitán, trilaurato de sorbitán, triestearato de sorbitán, trioleato de sorbitán o una combinación de estos. La composición también puede incluir un agente de ajuste de la tonicidad, por ejemplo, un compuesto que hace que la formulación sea sustancialmente isotónica o isoosmótica con la sangre humana. Los agentes de ajuste de la tonicidad ilustrativos incluyen sacarosa, sorbitol, glicina, metionina, manitol, dextrosa, inositol, cloruro de sodio, arginina y clorhidrato de arginina. En otras modalidades, la composición farmacéutica incluye además un estabilizador, por ejemplo, una molécula que previene o reduce sustancialmente la inestabilidad química y/o física de la nanoestructura proteica, en forma liofilizada o líquida. Los estabilizadores ilustrativos incluyen sacarosa, sorbitol, glicina, inositol, cloruro de sodio,

metionina, arginina y clorhidrato de arginina.

[0257] El pH de la composición está normalmente entre aproximadamente 4,5 y aproximadamente 11, tal como entre aproximadamente 5 y aproximadamente 11, entre aproximadamente 5,5 y aproximadamente 11, entre aproximadamente 6 y aproximadamente 11, entre aproximadamente 5 y aproximadamente 10,5, entre aproximadamente 5,5 y aproximadamente 10,5, entre aproximadamente 6 y aproximadamente 10,5, entre aproximadamente 5 y aproximadamente 10, entre aproximadamente 5,5 y aproximadamente 10, entre aproximadamente 6 y aproximadamente 10, entre aproximadamente 5 y aproximadamente 9,5, entre aproximadamente 5,5 y aproximadamente 9,5, entre aproximadamente 6 y aproximadamente 9,5, entre aproximadamente 5 y aproximadamente 9, entre aproximadamente 5,5 y aproximadamente 9, entre aproximadamente 6 y aproximadamente 9, entre aproximadamente 5 y aproximadamente 8,5, entre aproximadamente 5,5 y aproximadamente 8,5, entre aproximadamente 6 y aproximadamente 8,5, entre aproximadamente 5 y aproximadamente 8, entre aproximadamente 5,5 y aproximadamente 8, entre aproximadamente 6 y aproximadamente 8, aproximadamente 4,5, aproximadamente 5, aproximadamente 6,5, aproximadamente 6, aproximadamente 6,5, aproximadamente 7, aproximadamente 7,5, aproximadamente 8, aproximadamente 8,5, aproximadamente 9, aproximadamente 9,5, aproximadamente 10, aproximadamente 10,5, aproximadamente 11, etc. Se puede mantener un pH estable mediante el uso de un amortiguador, *por ejemplo*, un amortiguador Tris, un amortiguador citrato, un amortiguador fosfato o un amortiguador histidina. Por lo tanto, una composición generalmente incluirá un amortiguador.

[0258] En la presente se proporcionan composiciones para un medicamento, por ejemplo, para usar en la inducción o mejora de una respuesta inmune en un sujeto que lo necesita, tal como un mamífero.

[0259] En la presente se proporcionan composiciones para la fabricación de un medicamento para inducir o mejorar una respuesta inmune en un sujeto que lo necesita, tal como un mamífero.

Vacunas

[0260] En otro aspecto, la descripción proporciona vacunas que comprenden los polipéptidos, las nanoestructuras de proteína, o los polinucleótidos descritos anteriormente. Tales

composiciones se pueden usar para generar anticuerpos en un mamífero (*por ejemplo*, un ser humano). Las composiciones de vacunas de la descripción normalmente incluyen un portador farmacéuticamente aceptable, y una discusión exhaustiva de dichos portadores está disponible en Remington: The Science and Practice of Pharmacy.

5 **[0261]** Los antígenos podrían usarse alternativamente para vacunas que no contienen nanoestructura proteica con métodos de oligomerización más pequeños usando la etiqueta FoldOn, en donde la secuencia de la etiqueta FoldOn es EKAAKAEEAARK (SEQ ID NO: 85).

[0262] Las composiciones de vacunas pueden incluir un adyuvante inmunológico. Los adyuvantes ilustrativos incluyen los siguientes: 1. composiciones que contienen minerales; 2.
10 emulsiones oleosas; 3. formulaciones de saponina; 4. virosomas y partículas similares a virus; 5. derivados bacterianos o microbianos; 6. bioadhesivos y mucoadhesivos; 7. liposomas; 8. formulaciones de éter de polioxietileno y de éster de polioxietileno; 9. polifosfazeno (pcpp); 10. muramil péptidos; 11. compuestos de imidazoquinolona; 12. compuestos de tiosemicarbazona; 13. compuestos de triptantrina; 14. inmunomoduladores humanos; 15. lipopéptidos; 16.
15 benzonafitridinas; 17. micropartículas; 18. polinucleótidos inmunoestimuladores (tales como ARN o ADN; *por ejemplo*, oligonucleótidos que contengan cpg).

[0263] Por ejemplo, la composición puede incluir un adyuvante de sal de aluminio, una emulsión de aceite en agua (*por ejemplo*, una emulsión de aceite en agua que comprende escualeno, tal como MF59 o AS03), un agonista de TLR7 (tal como imidazoquinolina o imiquimod), o una
20 combinación de estos. Las sales de aluminio adecuadas incluyen hidróxidos (*por ejemplo*, oxihidróxidos), fosfatos (*por ejemplo*, hidroxifosfatos, ortofosfatos), (*por ejemplo*, consulte los capítulos 8 y 9 de Vaccine Design. (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum) o mezclas de estos. Las sales pueden adoptar cualquier forma adecuada (por ejemplo, gel, cristalina, amorfa, etc.), siendo un ejemplo la adsorción del antígeno a la sal. La concentración
25 de Al⁺⁺⁺ en una composición para administración a un paciente puede ser inferior a 5 mg/ml, *por ejemplo*, <4 mg/ml, <3 mg/ml, <2 mg/ml, <1 mg/ml, etc. Un rango preferido es entre 0,3 y 1 mg/ml. Se prefiere un máximo de 0,85 mg/dosis. Los adyuvantes de hidróxido de aluminio y fosfato de aluminio son adecuados para su uso con la descripción.

[0264] Los adyuvantes ejemplares que pueden usarse en una composición farmacéutica
30 proporcionada en el presente documento incluyen, pero no se limitan a, 3M-052, Adju-Phos™, Alhydrogel™, Adjumer™, micropartículas de albúmina-heparina, Algal Glucan, Algammulin,

Alum, Antigen Formulation, adyuvante AS-2, ASO1, ASO3, células dendríticas autólogas, PBMC autólogas, Avridine™, B7-2, BAK, BAY R1005, BECC agonistas TLR-4, Bupivacaína, Bupivacaína-HCl, BWZL, Calcitriol, Gel de fosfato cálcico, Péptidos CCR5, CFA, Holotoxina del cólera (CT) y subunidad B de la toxina del cólera (CTB), proteína de fusión subunidad A1-proteína A fragmento D de la toxina del cólera, CpG, CPG-1018, CRL1005, Liposomas que contienen citocinas, D-Murapalmitina, DDA, DHEA, Toxoide diftérico, DL-PGL, DMPC, DMPG, Complejo DOC/Alum, Fowlpox, adyuvante completo de Freund, gamma inulina, adyuvante de Gerbu, GM-CSF, GMDP, hGM-CSF, hIL-12 (N222L), hTNF-alfa, IFA, IFN-gamma en pcDNA3, ADN de IL-12, plásmido de IL-12, IL-12/GMCSF plásmido (Sykes), IL-2 en pcDNA3, IL-2/Ig plásmido, IL-2/Ig proteína, IL-4, IL-4 enpcDNA3, Imiquimod™, ImmTher™, Immunoliposomas que contienen anticuerpos contra moléculas coestimuladoras, Interferón-gamma, Interleucina-1 beta, Interleucina-12, Interleucina-2, Interleucina-7, ISCOM(s)™, Iscoprep 7,0·3™, hemocianina de lapa californiana, adyuvante basado en lípidos, liposomas, Loxoribina, LT(R192G), adyuvante oral LT o LT-OA, LT-R192G, LTK63, LTK72, adyuvante Matrix-M™, MF59, MONTANIDE ISA 51, MONTANIDE ISA 720, MPL.™., MPL-SE, MTP-PE, Liposomas MTP-PE, Murametide, Murapalmitine, NAGO, Toxina del cólera nativa nCT, Vesículas de surfactante no iónico, mutante no tóxico E112K de la toxina del cólera mCT-E112K, p-hidroxibenzoico metil éster, pCIL-10, pCIL12, pCMVmCAT1, pCMVN, Peptomer-NP, Pleuran, PLG, PLGA, PGA y PLA, Pluronic L121, PMMA, PODDS™, Poly rA: Poly rU, Polisorbato 80, Cocleatos de proteína, QS-21, Saponina Quadri A, Quil-A, Rehydragel HPA, Rehydragel LV, RIBI, Sistema adyuvante tipo Ribi (MPL, TMD, CWS), S-28463, SAF-1, péptido Sclavo, proteoliposomas de Sendai, matrices lipídicas que contienen Sendai, Span 85, Specol, escualano 1, escualeno 2, estearil tirosina, SWE, toxoide tetánico (TT), Theramide™, dipéptido de treonil muramilo (TMDP), partículas Ty y liposomas de Walter Reed.

25 **[0265]** En modalidades preferidas, el adyuvante es un gel de hidróxido de aluminio (por ejemplo, Alhydrogel™). En modalidades preferidas, el adyuvante es SWE. En modalidades preferidas, el adyuvante es MF59. MF59 es una emulsión de aceite en agua que contiene escualeno (4,3 %) en tampón de ácido cítrico con tensioactivos no iónicos estabilizadores Tween 80 (0,5 %) y Span 85 (0,5 %). Se ha demostrado que el MF59 es bien tolerado en humanos y se utiliza en vacunas contra la gripe estacional (véase Ko y Kang, Hum Vaccin Immunother. 2018; 14(12): 3041–3045; patente de los EE. UU. N.º 6.299.884).

[0266] Por ejemplo, la composición puede incluir un adyuvante de sal de aluminio, una emulsión de aceite en agua (por ejemplo, una emulsión de aceite en agua que comprende escualeno, tal como MF59, SWE o AS03), un agonista de TL4, un agonista de TLR9 (tal como oligodesoxinucleótidos CpG), un agonista de TLR7 (tal como imidazoquinolina o imiquimod), o una combinación de estos. En algunas modalidades, el adyuvante es una combinación de una sal de aluminio y CPG1018. Las sales de aluminio adecuadas incluyen hidróxidos (por ejemplo, oxihidróxidos), fosfatos (por ejemplo, hidroxifosfatos, ortofosfatos), (por ejemplo, consulte los capítulos 8 y 9 de *Vaccine Design*. (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum) o mezclas de estos. Las sales pueden adoptar cualquier forma adecuada (por ejemplo, gel, cristalina, amorfa, etc.), siendo un ejemplo la adsorción del antígeno a la sal. La concentración de Al+++ en una composición para administración a un paciente puede ser inferior a 5 mg/ml, por ejemplo, <4 mg/ml, <3 mg/ml, <2 mg/ml, <1 mg/ml, etc. Un rango preferido es entre 0,3 y 1 mg/ml. Se prefiere un máximo de 0,85 mg/dosis. Los adyuvantes de hidróxido de aluminio y fosfato de aluminio son adecuados para su uso con la descripción. En algunas modalidades, una composición farmacéutica proporcionada en la presente comprende hidróxido de aluminio como adyuvante.

[0267] En algunas modalidades, el adyuvante es una emulsión de escualeno. En algunas modalidades, el adyuvante es un inmunoestimulante TLR4 (por ejemplo, SLA, GLA), por ejemplo, como se describe en Van Hoesen et al. *PLoS One*. 11(2):e0149610 (2016). En algunas modalidades, el adyuvante es un inmunoestimulante TLR7/8 (por ejemplo, R848, IMQ, 3M-052), por ejemplo, como se describe en Dowling D. *ImmunoHorizons* (6):185-197 (2018). En algunas modalidades, el adyuvante es un inmunoestimulante TLR9 (CpG), por ejemplo, como se describe en Bode et al. *Expert Rev Vaccines*. 10(4):499–511 (2011). En algunas modalidades, el adyuvante es saponina (QS21), por ejemplo, como se describe en Zhu et al. *Nat Prod Chem Res*. 3(4):e113 (2016).

[0268] En algunas modalidades, la vacuna comprende una combinación de dos o más adyuvantes (por ejemplo, emulsión de escualeno y alumbre o un inmunoestimulante TLR4).

[0269] Una composición de vacuna comprende una cantidad inmunológicamente eficaz de su(s) antígeno(s). Una «cantidad inmunológicamente eficaz» o «cantidad eficaz» es una cantidad que, administrada a un sujeto, es eficaz para provocar una respuesta de anticuerpos contra el antígeno. Esta cantidad puede variar en función de la salud y el estado físico del individuo a

tratar, su edad, la capacidad del sistema inmunitario del individuo para sintetizar anticuerpos, el grado de protección deseado, la formulación de la vacuna, la evaluación de la situación médica por parte del médico tratante y otros factores pertinentes. Se espera que la cantidad se sitúe en un rango relativamente amplio que pueda determinarse mediante ensayos rutinarios. El contenido de antígeno de las composiciones de la divulgación se expresará generalmente en términos de masa de proteína por dosis. Puede ser útil una dosis de 10 a 500 µg (por ejemplo, 50 µg) por antígeno.

[0270] En algunas modalidades, la vacuna (composición inmunogénica) es una emulsión estable.

[0271] En algunas modalidades, la vacuna es una vacuna bivalente. En algunas modalidades, la vacuna bivalente comprende el polipéptido o la nanoestructura proteica descritos en la presente. La vacuna bivalente puede comprender polipéptidos, nanoestructuras proteicas o polinucleótidos de dos cepas diferentes de un coronavirus, en cualquier combinación. Por ejemplo, la vacuna bivalente podría contener un cóctel de dos inmunógenos diferentes que muestren por separado cada antígeno, o inmunógenos “en mosaico” que muestren conjuntamente ambos antígenos. Por ejemplo, la vacuna bivalente puede incluir un polipéptido de la cepa del virus original para proporcionar una amplia protección contra la COVID-19 y un componente de la variante ómicron para proporcionar una mejor protección contra la COVID-19 causada por la variante ómicron.

Métodos de uso

[0272] En otro aspecto, la descripción proporciona un método para prevenir o tratar una enfermedad relacionada con el coronavirus en un sujeto que lo necesita, que comprende administrar al sujeto los polipéptidos, las nanoestructuras de proteína, los polinucleótidos, el componente farmacéutico o la vacuna descritos anteriormente.

[0273] En otro aspecto, la descripción proporciona un método para inmunizar a un sujeto que lo necesita contra una infección por coronavirus, administrándole al sujeto la vacuna descrita anteriormente.

[0274] En algunas modalidades, tal método protege contra el desarrollo de una enfermedad asociada al coronavirus (*por ejemplo*, enfermedad grave), *por ejemplo*, COVID-19. El sujeto puede ser cualquier sujeto mamífero adecuado, incluido, entre otros, un sujeto humano. En algunas modalidades, un sujeto es un niño humano, *por ejemplo*, un niño de menos de 12

meses de edad. En algunas modalidades, un sujeto es un niño pequeño, por ejemplo, de aproximadamente 1 a aproximadamente 3 años de edad o de aproximadamente 1 a aproximadamente 5 años de edad. En algunas modalidades, el sujeto es un adulto de más de 60 años. En algunas modalidades, el sujeto es un adulto de más de 65 años. En modalidades particulares, el sujeto depende de la ayuda de otros o tiene problemas o riesgos de salud graves (*por ejemplo*, una persona mayor frágil). En algunas modalidades, el sujeto es un adulto sano de entre 18 y 60 años de edad. En algunas modalidades, el sujeto es un adulto sano de entre 18 y 45 años. En otra modalidad, el sujeto es una mujer embarazada. En algunas modalidades, el sujeto es un adulto humano inmunodeprimido. En algunas modalidades, el sujeto es un adulto humano que padece una enfermedad cardíaca y/o pulmonar crónica subyacente o una discapacidad funcional. En algunas modalidades, el sujeto corre riesgo de padecer enfermedad por VSR (*p. ej.*, LRTI o neumonía).

[0275] Las composiciones inmunogénicas proporcionadas en este documento pueden usarse para vacunar a un feto. Se recomienda la administración de ciertas vacunas inactivadas durante el embarazo para inducir inmunidad en el feto, por ejemplo, la vacuna contra el toxoide tetánico, la vacuna contra el toxoide diftérico reducido y la vacuna contra la tos ferina acelular (Tdap) y la vacuna contra la influenza. Por lo tanto, en algunas modalidades, se proporciona en el presente documento un método para generar una respuesta inmune en un feto, comprendiendo el método la administración de una cantidad eficaz de la composición inmunogénica proporcionada en el presente documento a la madre de dicho feto. La composición inmunogénica se puede administrar en cualquier momento adecuado durante el embarazo, por ejemplo, en el último trimestre del embarazo.

[0276] Una composición proporcionada en la presente puede coadministrarse con otros tratamientos, tales como otras vacunas. Por lo tanto, en algunas modalidades, a un sujeto tratado de acuerdo con un método proporcionado en la presente también se le puede administrar una o más vacunas estacionales o pandémicas, tal como una vacuna contra la influenza. En algunas modalidades, a un sujeto tratado de acuerdo con los métodos proporcionados en este documento también se le puede administrar una vacuna neumocócica, herpes zóster recombinante (culebrilla) o Tdap. Se pueden coadministrar una, dos o más vacunas con una composición inmunogénica proporcionada en este documento. La “coadministración” incluye tanto la administración simultánea como la posterior. Por ejemplo,

las una, dos o más vacunas y una composición inmunogénica proporcionadas en este documento pueden administrarse el mismo día. En algunas modalidades, las una, dos o más vacunas y una composición inmunogénica proporcionadas en este documento se administran dentro de 1 hora, dentro de 2 horas, dentro de 3 horas, dentro de 4 horas, dentro de 5 horas, dentro de 6 horas, dentro de 8 horas, dentro de 10 horas o dentro de 12 horas.

[0277] En otro aspecto, se proporciona en la presente un método para tratar a un sujeto que padece una infección viral. Como se usa en la presente, “tratar” o “tratar” incluye, entre otros lograr uno o más de los siguientes: (a) reducir el título viral en el sujeto; (b) limitar cualquier aumento del título viral en el sujeto; (c) reducir la gravedad de los síntomas de la infección viral; (d) limitar o prevenir el desarrollo de síntomas después de la infección viral; (e) inhibir el empeoramiento de los síntomas de la infección viral; (f) limitar o prevenir la reaparición de los síntomas de la infección viral en sujetos que previamente presentaban síntomas de infección viral; y/o (e) aumentar la supervivencia. En algunas modalidades, un método de vacunación disminuye el riesgo del sujeto de infectarse con un virus. En algunas modalidades, un método de vacunación limita el desarrollo de una infección viral. En algunas modalidades, un método de vacunación disminuye la gravedad de los síntomas de la infección viral. Los síntomas de una infección por coronavirus incluyen, entre otros, fiebre o escalofríos, tos, dificultad para respirar, fatiga, dolores musculares o corporales, dolor de cabeza, pérdida reciente del gusto o del olfato, dolor de garganta, congestión o secreción nasal, náuseas o vómitos y diarrea.

[0278] En algunas modalidades, los métodos proporcionados en la presente pueden usarse para prevenir una infección o enfermedad viral (*por ejemplo*, enfermedad respiratoria aguda) en un sujeto. Como se usa en la presente, “prevenir” o “prevenir” incluye, entre otros lograr uno o más de los siguientes: (a) generar una respuesta inmune (anticuerpo y/o basada en células, *por ejemplo*, linfocitos T CD4, linfocitos B de memoria y/o linfocitos T CD8) a un virus en el sujeto que se espera que confiera protección contra la infección del tracto respiratorio inferior (LRTI) causada por o asociada con un virus en el sujeto; (b) generar anticuerpos neutralizantes contra un virus en el sujeto que se espera que reduzca la gravedad de la LRTI causada por o asociada con un Coronavirus en el sujeto; (c) prevenir la LRTI causada por o asociada con un virus en un sujeto, detectada como un aumento en el título del virus i del sujeto o por un aumento en uno o más síntomas de infección viral; (d) prevenir la LRTI grave causada por o asociada con el Coronavirus en un sujeto detectada como un aumento en el título del virus del sujeto o por

un aumento en uno o más síntomas graves de infección viral; (e) reducir el riesgo de LRTI o LRTI grave causada por o asociada con el Coronavirus dentro de una población de sujetos; o (f) causar una serorrespuesta (o seroconversión) de un sujeto, tal como generar anticuerpos neutralizantes contra un virus al menos 4 veces más alto que un nivel de anticuerpos basal en el sujeto. La prevención se puede evaluar comparando las respuestas inmunes, especialmente los correlatos de protección, en sujetos a los que se les administró una vacuna con los mismos sujetos antes de la administración (denominada línea base), sujetos a los que se les administró un placebo o sujetos a los que se les administró una vacuna de comparación.

[0279] Como se usa en la presente, "limitar" el desarrollo de una infección viral, se refiere a lograr uno o más de los siguientes: (a) generar una respuesta inmune (anticuerpo y/o basada en células, *por ejemplo*, linfocitos T CD4, linfocitos B de memoria y/o linfocitos T CD8) a un virus en el sujeto que se espera que limite un aumento en el título viral o los síntomas en el sujeto; (b) generar anticuerpos neutralizantes contra un coronavirus en el sujeto a un nivel que se espera que limite un aumento en el título viral o los síntomas en el sujeto; (c) causar títulos virales reducidos en el sujeto después de la exposición a un coronavirus en comparación con sujetos a los que no se les administró el complejo proteico; y (d) causar una incidencia o gravedad reducida de los síntomas después de la infección viral.

[0280] La eficacia clínica de una vacuna contra la infección por coronavirus se puede evaluar mediante diversos medios conocidos en la técnica, incluidos, entre otros, estudios de eficacia clínica controlados con placebo para medir la carga viral o los síntomas de la enfermedad viral en sujetos vacunados frente a sujetos de control. También se pueden definir correlatos de protección, tales como títulos de anticuerpos neutralizantes (generalmente expresados como títulos de media geométrica), aumentos por encima de los valores iniciales (generalmente expresados como aumento geométrico) y tasa de serorrespuesta (un porcentaje de sujetos que logran un aumento en los títulos de anticuerpos neutralizantes por encima de un umbral predeterminado).

[0281] En algunas modalidades, los métodos descritos en la presente generan una respuesta inmune en un sujeto que no se sabe que está infectado con un virus, en donde la respuesta inmune sirve para limitar el desarrollo de la infección y los síntomas de una infección viral. En algunas modalidades, la respuesta inmune comprende la generación de anticuerpos neutralizantes y/o respuestas basadas en células contra un virus. En algunas modalidades, la

respuesta inmune comprende la generación de respuestas específicas de la proteína glucoproteína Spike con un título geométrico medio de al menos 1×10^3 , al menos 1×10^4 , al menos 1×10^5 , al menos 1×10^6 , al menos 1×10^7 , al menos 1×10^8 o al menos 1×10^9 . En una modalidad adicional, la respuesta inmune comprende la generación de anticuerpos contra múltiples epítomos antigénicos en el trímero de proteína específico de la proteína glucoproteína Spike.

[0282] En un aspecto, los métodos proporcionados en la presente pueden dar como resultado un aumento en los títulos de anticuerpos en un sujeto, *por ejemplo*, en un aumento en los anticuerpos neutralizantes específicos del virus o en los anticuerpos de unión específicos del virus. Los títulos de anticuerpos se pueden determinar usando cualquier ensayo adecuado conocido en la técnica o descrito en la presente, incluidos, entre otros, ensayos inmunoabsorbentes ligados a enzimas de unión (ELISA), inmunoadsorción ligado a enzimas de puntos (ELISpot), ELISA de competencia, inmunoprecipitación, inmunotransferencia y ensayos de aglutinación.

[0283] En algunas modalidades, se puede usar un ensayo de neutralización o microneutralización (MN) para medir aumentos en los anticuerpos neutralizantes en un sujeto después de la administración de una composición como se describe en la presente. La microneutralización se refiere a una neutralización realizada en un formato miniaturizado, tal como una placa de 96 pocillos. Se usa un ensayo de (micro)neutralización para comprobar la inhibición de un virus por anticuerpos (*por ejemplo*, anticuerpos purificados, suero o plasma). El ensayo mide el nivel de anticuerpos presentes en una muestra que son capaces de neutralizar un virus *in vitro*. Generalmente, los ensayos de microneutralización para muestras clínicas se realizan con una dilución seriada de suero mezclado con una concentración fija de virus. Los métodos para realizar ensayos de (micro)neutralización son bien conocidos. Se describen ensayos ilustrativos de microneutralización, véase, *por ejemplo*, van Baalen et al. Vaccine 35 (2017) 46–52.

[0284] En algunas modalidades, los métodos proporcionados en la presente dan como resultado un aumento de células inmunes en un sujeto, (*por ejemplo*, un aumento de linfocitos B de memoria específicos del virus y/o linfocitos T específicos del virus). La cantidad de células inmunes en un sujeto se puede determinar utilizando cualquier ensayo adecuado conocido en la técnica o descrito en este documento, incluidos, entre otros, FACS y citometría de flujo.

[0285] En algunas modalidades, la composición se administra por cualquier vía adecuada, incluyendo por vía intranasal, sublingual, oral, parenteral, por pulverización de inhalación, rectal o tópica en formulaciones de unidades de dosificación que contienen vehículos, adyuvantes y portadores farmacéuticamente aceptables convencionales. El término parenteral como se utiliza en este documento incluye, subcutánea, intravenosa, intraarterial, intramuscular, intraesternal, intratendinosa, intraespinal, intracraneal, intratorácica, técnicas de infusión o intraperitoneal.

Kits

[0286] En otro aspecto, la descripción proporciona kits que pueden usarse para preparar y administrar polipéptidos, nanoestructuras de proteínas, polinucleótidos, composiciones farmacéuticas o vacunas como se describió anteriormente. En algunas modalidades, un kit proporcionado en la presente comprende polipéptidos, nanoestructuras de proteínas, polinucleótidos, composiciones farmacéuticas o vacunas como se describe en la presente, e instrucciones para su uso en un método como se describe anteriormente. En algunas modalidades, un kit comprende una o más dosis unitarias como se describe en la presente, e instrucciones para su uso en un método de la descripción. En algunas modalidades, el kit comprende un vial que comprende una dosis única de una composición farmacéutica proporcionada en la presente. En algunas modalidades, un kit comprende un vial que comprende múltiples dosis proporcionadas en este documento. En algunas modalidades, un kit comprende además instrucciones para el uso de la composición farmacéutica. En algunas modalidades, un kit comprende además un diluyente para preparar diluciones de la composición farmacéutica antes de la administración. En algunas modalidades, la vacuna comprende un adyuvante. En algunas modalidades, un kit comprende una composición farmacéutica y un adyuvante que deben mezclarse antes de la administración.

EJEMPLOS

Materiales y métodos

[0287] Modelado estructural. Se usó PyMol versión 2.5.2 o ChimeraX versión 1.3 para analizar los modelos estructurales y generar imágenes. La predicción basada en la estructura usando Colabfold (**Ref 11**) se realizó con configuraciones predeterminadas contra la región RDB-SD1 de cada secuencia diseñada.

[0288] Expresión HEK293. Un día antes de las transfecciones, se contaron las células Expi293F (Thermo Fisher Scientific) y se sembraron a $2,5 \times 10^6$ células/mL y posteriormente se

incubaron durante la noche a 36 °C con agitación a 120 rpm. Al día siguiente, se contaron las células y se diluyeron a 3×10^6 células/mL. Luego, las células se transfectaron transitoriamente con el plásmido deseado de acuerdo con las instrucciones del fabricante (Thermo Fisher Scientific), con la excepción de que se usó Transporter 5 (Polysciences) como reactivo de transfección en una proporción de 30 ug de Transporter 5 por 1 ug de ADN plasmídico. Luego, las células transfectadas se incubaron a 36 °C durante 4 días mientras se agitaban a 120 rpm.

[0289]Transferencias Western. Se realizaron transferencias Western contra sobrenadantes de células HEK293 tras la expresión a pequeña escala usando el anticuerpo conjugado anti-His HRP. El sobrenadante se diluyó 2 veces en amortiguador de carga Laemmli 2x (Bio-Rad) con 5 % (v/v) de β -mercaptoetanol y luego se calentó a 95°C durante 10 minutos. Se cargaron 10 μ l en un gel de proteína NuPAGE 4-12 % Bis-Tris junto con 7 μ l de estándar/marcador de proteína marcada con His (Invitrogen, n.º LC5606) y se ejecutó durante 30 minutos a 200 V. Las proteínas se transfirieron a una membrana Immuno-Blot PDVF (Bio-Rad). Después de la transferencia, la transferencia se bloqueó con 2 % de leche (p/v) en 1x PBS, 0,05 % Tween 20 durante 1 hora a temperatura ambiente con agitación. Se diluyó un anticuerpo anti-His, conjugado con HRP (R&D Systems MAB050H) 1:7000 en 2 % de leche (p/v) en 1x PBS, 0,05 % Tween 20 y se incubó con la membrana durante 1 hora a temperatura ambiente con agitación. Luego, la membrana se lavó tres veces durante cinco minutos usando 1x PBS, 0,05 % Tween 20 y se detectó la proteína marcada con His usando (BioRx TMB Enhanced One Component HRP Membrane Substrate, Surmodics) y se capturó en el sistema de imágenes BioRad ChemiDoc MP.

[0290]Purificación IMAC Las muestras de proteína NA se purificaron directamente del sobrenadante de células HEK Expi293F mediante cromatografía de afinidad por metal inmovilizado (IMAC). La resina Ni^{2+} -agarosa (Cube Biotech, Indigo) se lavó con 5 volúmenes de columna (CV) de agua seguido de 5 CV de amortiguador de equilibrio (20 mM Tris pH 8,0, 250 mM de NaCl, 5 % de glicerol, 30 mM de imidazol), luego se resuspendió en 1 CV de amortiguador de equilibrio. Los sobrenadantes se clarificaron mediante centrifugación a 4000 x g seguida de filtración usando unidades de filtración al vacío de 0,2 mm. Se añadió suspensión de resina a cada sobrenadante en una proporción de 4 μ L de resina por mL de sobrenadante. Las suspensiones de resina y sobrenadante se incubaron con agitación suave a 4 °C durante ~16-18 horas, luego la resina se recolectó mediante aplicación a columnas de gravedad vacías

(Biorad, 7321010). Los lechos de resina se lavaron con un amortiguador de equilibrio de 20 CV y las proteínas se eluyeron usando un amortiguador de elución de 6 CV (20 mM Tris pH 8,0, 250 mM de NaCl, 5 % de glicerol, 300 mM de imidazol) en tres incrementos de 2 CV con incubaciones de 30 minutos entre cada aplicación. Después de la purificación mediante IMAC, las muestras de proteínas se purificaron aún más mediante cromatografía de exclusión por tamaño (SEC) o se dializaron en “amortiguador RBD 1” (20 mM Tris pH 8,0, 150 mM de NaCl), “amortiguador RBD 2” (5 % de glicerol, 20 mM Tris pH 8,0, 250 mM de NaCl, 5 % de glicerol) o “amortiguador RBD 3” (50 mM de MOPS, 150 mM de NaCl, pH 7,4).

[0291] Purificación SEC Cada elución de IMAC para muestras de antígeno-CompA se inyectó en una columna Superdex 200 Increase 10/300 GL (Cytiva) equilibrada en amortiguador RBD 1, amortiguador RBD 2 o amortiguador RBD 3 y se eluyó con 1,2 volúmenes de columna del amortiguador de ejecución respectivo para purificar aún más los antígenos mediante cromatografía de exclusión por tamaño (SEC). Para las preparaciones de nanoestructuras de proteínas ensambladas, las muestras se aplicaron a una columna Superose 6 Increase 10/300 GL (Cytiva) equilibrada en 20 mM Tris Ph 8,0, 150 mM de NaCl, 100 mM de L-Arginina y 5 % de glicerol y se eluyeron con 1,2 volúmenes de columna del mismo amortiguador. Después del análisis, las muestras purificadas se congelaron rápidamente en nitrógeno líquido y se almacenaron a -80 °C.

[0292] Interferometría de biocapa. ACE2-Fc, CR3022 (**Ref 12**) o S309 (**Ref 16**) normalizados en concentración a 10 ug/mL en amortiguador de ensayo BLI (PBS, 0,5 % de BSA, 0,05 % de Tween 20, pH 7,4) en un volumen suficiente para cargar 200 µL por pocillo de una microplaca negra de 96 pocillos (Greiner, 655209). Brevemente, en un instrumento Octet Red96e, los biosensores de proteína G pre-hidratados (Sartorius, 18-5022) se sumergieron en amortiguador de ensayo durante 60 segundos para conseguir un valor de referencia. A continuación, los biosensores se sumergieron en ACE2-Fc, CR3022 o S309 durante 60 segundos o 120 segundos respectivamente para su inmovilización, seguido de una etapa de referencia adicional. Se permitió que los anticuerpos inmovilizados se asociaran con 200 µL de fusiones antígeno-CompA purificado (a 10 µg/mL en amortiguador de ensayo BLI) o sobrenadantes celulares durante 120 a 700 segundos, y luego se sumergieron los biosensores en amortiguador de ensayo durante 120 a 700 segundos para observar cualquier posible disociación. Se informan los valores de desplazamiento máximo durante la fase de asociación a menos que se

indique lo contrario.

[0293] NanoDSF usando ITF. La estabilidad térmica de las proteínas y las mediciones de la línea base fluorescente se determinaron usando la fluorescencia intrínseca de la proteína (ITF) en un instrumento nanoDSF (UNchained Laboratories, UNcle). Las concentraciones de antígeno se normalizaron a 0,5-1 mg/ml. Cada muestra se aplicó a un casete capilar de cuarzo (UNchained Laboratories, UNi) para realizar mediciones de ITF en una rampa de temperatura (15 – 95 °C, 0,5-1 °C/min) después de un equilibrio de 5 minutos a 15 °C. Las temperaturas de fusión se determinaron para muestras individuales representativas a partir de cálculos de media baricéntrica usando el software Uncle Analysis V6.0 con configuraciones predeterminadas o con un script de Python personalizado. El script de Python calculó la temperatura de fusión a partir de la segunda derivada de la media baricéntrica dependiente de la temperatura. La temperatura de inicio se calculó a partir de la intersección de un ajuste lineal del inicio previo a la transición y la transición de fusión.

[0294] NanoDSF usando SYPRO. Los perfiles de fusión térmica de las construcciones NA se determinaron normalizando las concentraciones de muestra a 0,5-1 mg/mL. Luego, el gel de proteína naranja SYPRO™ (Invitrogen S6651) se diluyó 1:20 en cada amortiguador de muestra y se agregaron 3,5 µL a 31,5 µL de construcciones de NA normalizadas. Cada muestra se aplicó a un casete capilar de cuarzo (UNchained Laboratories, UNi) para realizar mediciones de ITF en una rampa de temperatura (15 – 95 °C, 0,5-1 °C/min) después de un equilibrio de 5 minutos a 15 °C. Las temperaturas de fusión se determinaron usando mediciones de área fluorescente a partir de muestras únicas representativas usando el software Uncle Analysis V6.0 con configuraciones predeterminadas. La fluorescencia de referencia cercana a 30 °C se midió a partir de integraciones de intensidad fluorescente basadas en tres réplicas promediando puntos que se recolectaron a temperaturas más cercanas a 30 °C para cada réplica.

[0295] Ensamblaje de muestras de antígeno-CompA con CompB. Para cualquier proceso de ensamblaje de nanoestructuras de proteínas, se mezcló CompB purificado (I53-50B.4PosT1) con una muestra de antígeno-CompA mediante pipeteo, con una relación molar de 1,1 partes de antígeno-CompA por 1 parte de CompB. Para reacciones de ensamblaje a pequeña escala, el proceso de ensamblaje a veces se realizó con la adición de otras soluciones de amortiguación concentradas para agregar excipientes a la solución final, incluidas concentraciones finales de 50-100 mM de L-arginina, 0,75 % de CHAPS, 4 % de sacarosa o 5 % de glicerol. Los excipientes

usados en cada caso se detallan en tablas posteriores. Para los ensamblajes de ampliación antes de la SEC, el proceso de ensamblaje se realizó en un fondo de 20 mM Tris Ph 8,0, 150 mM de NaCl, 100 mM de L-arginina y 5 % de glicerol. Todos los procesos de ensamblaje se dejaron reposar a temperatura ambiente durante al menos 30 minutos antes de la SEC o el análisis por dispersión de luz dinámica (DLS) usando protocolos predeterminados en un instrumento DLS (UNchained Laboratories, UNcle).

[0296] Inmunizaciones animales. Se inmunizaron ratones BALB/c hembra por vía intramuscular los días 0 o 21 con 0,027-1,0 ug de proteína antígeno-CompA purificada o antígeno-VLP mezclada 1:1 con adyuvante Addavax™. Se recogieron muestras de suero el día 0 (antes de la primera inmunización), el día 21 (antes de la segunda inmunización) y el día 35. Se realizaron ensayos de unión y neutralización de antígenos en Nexelis contra los respectivos antígenos y pseudovirus para cepas ancestrales y BA.5.

Ejemplo 1: Diseño de mutaciones de SD1resurf y su mayor función hacia el ensamblaje de nanoestructuras

[0297] En una primera ronda del proceso de diseño, se probaron las características iniciales usando antígenos RBD de la cepa ancestral SARS-CoV-2. Todos los antígenos probados se fusionaron genéticamente al trímero “CompA” I53-50A, que es un componente basado en proteínas que se puede ensamblar *in vitro* en una nanoestructura de proteína cuando se agrega un pentámero “CompB” complementario (en este caso, I53-50B.4PosT1) (**Ref 2**). Se probaron por separado múltiples mutaciones estabilizadoras descritas previamente (Rpk2, Rpk4, Rpk8, Rpk9, Rpk11 y Rpk15) (**Ref. 5**) en el contexto de este antígeno RBD-SD1 (**Tabla 14**). Se diseñó un nuevo conjunto de mutaciones en el dominio SD1 (denominado “SD1resurf”) que se dirigían a los grupos hidrófobos expuestos en el dominio SD1, lo que hacía que estos grupos fueran más polares para reducir la agregación (L560Q, F562Y, I569S). L560Q e I569S se diseñaron basándose en otras secuencias SD1 de sarbecovirus que contienen estos residuos polares en estas posiciones, mientras que F562Y se diseñó racionalmente para hacer que este residuo sea más polar. Todos los diseños de RBD-SD1 se probaron individualmente con uno de cinco enlaces flexibles diferentes entre el antígeno RBD-SD1 y CompA (**Tabla 15**), que tenían una longitud que variaba de 16 a 24 aminoácidos. Se generaron secuencias de ADN optimizadas por codones para todos los diseños con una secuencia de señal de secreción agregada, que luego se clonaron en un vector para la expresión de células HEK293 mediante transfecciones

transitorias.

Tabla 14. Mutaciones probadas en la primera ronda de diseños.

Nombre del conjunto de mutaciones	Mutaciones
Rpk2	F338L, Y365W
Rpk4	F392W
Rpk8	Y365F, V395I
Rpk9	Y365F, F392W, V395I
Rpk11	F338L, A363L, Y365M
Rpk15	I358F, Y365F, V395I
SD1resurf	L560Q, F562Y, I569S

Tabla 15. Enlaces probados entre antígenos RBD-SD1 y CompA

Longitud del enlazador (residuos de aminoácidos)	Secuencia de aminoácidos
16	GGSGGSGSGSGSGSGS
18	SGGGSGGSGSGSGSGSGS
19	EPEGSGGSGSGSGSGSGS
20	YGGSGGSGGSGSGSGSGSGS
24	GSGGSGGSGGSGGSGSGSGSGS

5

[0298] Los diseños se evaluaron primero a partir de sobrenadantes a pequeña escala, tanto para expresión mediante transferencia Western (**Figura 2A**) como para antigenicidad mediante interferometría de biocapa (BLI) usando receptores ACE2-Fc inmovilizados (**Figura 2B y Figura 2C, Tablas 16 y 17**) después de la selección descendente. Sorprendentemente, la mayoría de las muestras analizadas mostraron una señal de unión a ACE2 más alta en comparación con el control “RBD01”, que presentaba el fragmento RBD fusionado a CompA (**Ref 14**), y estos datos significaron una expresión y/o antigenicidad mejorada. Se observó que los niveles más altos de unión estaban influenciados independientemente por la adición de mutaciones estabilizadoras, incluidas múltiples mutaciones de Rpk y mutaciones de SD1resurf, y por el uso de enlaces de 18 o 20 residuos. En función de estos datos, se ampliaron múltiples diseños que contenían el enlazador de 18 residuos a través de la transfección HEK293 y se purificaron mediante cromatografía de afinidad de iones metálicos inmovilizados (IMAC) para una caracterización

10

15

más detallada, y se compararon con un antígeno RBD silvestre sin dominio SD1 (residuos 328-531) fusionado genéticamente a CompA. Se realizó BLI contra ACE2-Fc inmovilizado o el anticuerpo monoclonal neutralizante S309 (**Ref. 12**) o el anticuerpo monoclonal no neutralizante CR3022 (**Ref. 16**), y los resultados de ACE2-Fc y S309 se usaron para la selección descendente dada su unión a epítomos a los que se dirigen los anticuerpos neutralizantes. BLI contra ACE2-Fc y S309 mostró resultados generalmente similares (**Figura 3, Tabla 18**), con algunas construcciones que conservaban la isoleucina nativa (I) en la posición 358 que parecían superiores según las mediciones de S309. La fluorimetría de barrido diferencial nano (NanoDSF) basada en la fluorescencia intrínseca del triptófano mostró temperaturas de fusión mejoradas a partir de diseños que contenían mutaciones estabilizadoras de Rpk4, Rpk8, Rpk9 y Rpk15 en relación con antígenos silvestres con y sin dominios SD1 (**Figura 4A a Figura 4G, Tabla 19**). Las temperaturas de fusión fueron similares para los antígenos RBD silvestre con y sin dominios SD1 (**Figura 4A a Figura 4B, Tabla 19**), y para el diseño que contenía mutaciones SD1resurf (**Figura 4G, Tabla 19**). Se observaron resultados similares para NanoDSF monitoreados usando el tinte SYPRO, aunque los diseños Rpk4 y Rpk9 mostraron inicios notablemente más bajos en comparación con otros diseños, lo que sugiere una mejora en el plegamiento que reduce la hidrofobia de estos diseños (**Figura 4H a Figura 4K, Tabla 20**). Finalmente, se probó el ensamblaje para diferentes diseños en diversas condiciones de excipientes mediante la adición del componente pentamérico complementario CompB y el monitoreo del tamaño de partículas usando dispersión de luz dinámica (DLS) en la cepa ancestral (**Figura 5A y Figura 5B, Tabla 21**). Cuando se ensamblaron en amortiguadores con excipientes más modestos (arginina (**Figura 5B, Tabla 21**), o arginina y sacarosa (**Figura 5A, Tabla 21**), se observó una agregación significativa para todos los diseños excepto aquellos que contenían mutaciones estabilizadoras de Rpk9 y mutaciones de SD1resurf, que proporcionaron señales DLS coherentes con un ensamblaje exitoso de la nanoestructura de la proteína (35-45 nm).

[0299] En total, esta primera ronda de diseño demostró que las mutaciones Rpk9 descritas previamente pueden mejorar con éxito la estabilidad térmica de los antígenos RBD-SD1 de la cepa ancestral y mejorar aún más su capacidad para ensamblarse en nanoestructuras de proteínas I53-50 con una agregación mínima en formulaciones de amortiguador más simples. Sorprendentemente, se identificaron nuevas mutaciones SD1resurf que ayudaron al

ensamblaje de nanoestructuras de proteínas I53-50 con agregación mínima. Ambos conjuntos de mutaciones mostraron al menos la antigenicidad de los epítomos clave a los que se dirigen los anticuerpos neutralizantes.

[0300] La Tabla 16 muestra una primera ronda de BLI usando ACE2-Fc inmovilizado frente a sobrenadantes de diseños ancestrales expresados a pequeña escala.

Tabla 16.

Construcciones	Desplazamiento máximo (nm)
RBDb001	0,7794
RBDb002	0,8548
RBDb003	0,8513
RBDb004	0,85
RBDb005	0,8797
RBDb006	0,8115
RBDb011	0,9095
RBDb016	1,2278
RBDb021	0,8613
RBDb026	0,783
RBDb031	0,8718
RBDb036	0,8456
RBDb041	0,8785
RBD01 réplica 1	0,7766
RBD01 réplica 2	0,7968
RBD01 réplica 3	0,7721

[0301] La Tabla 17 muestra una segunda ronda de BLI usando ACE2-Fc inmovilizado frente a sobrenadantes de diseños ancestrales expresados a pequeña escala.

Tabla 17:

Construcciones	Desplazamiento máximo (nm)
RBDb016	1,1147

RBDb017	0,8861
RBDb018	0,8502
RBDb019	0,8952
RBDb020	0,859
RBDb021	0,833
RBDb022	0,8914
RBDb023	0,8497
RBDb024	0,8868
RBDb025	0,8685
RBDb036	0,8031
RBDb037	0,8782
RBDb038	0,8251
RBDb039	0,8861
RBDb040	0,867
RBD01 réplica 1	0,7022
RBD01 réplica 2	0,7005
RBD01 réplica 3	0,6929

[0302] La Tabla 18 muestra BLI frente a proteínas antígeno-CompA purificadas usando ACE2-Fc inmovilizado, CR3022 o S309.

Tabla 18:

Construcciones	ACE2, desplazamiento máximo (nm)	CR3022, desplazamiento máximo (nm)	S309, desplazamiento máximo (nm)
RBDb001 réplica 1	0,3426	1,0786	0,8241
RBDb001 réplica 2	0,3416	1,0706	0,8127
RBDb002	0,3267	0,5924	0,7878
RBDb012	0,3355	1,4242	0,8286
RBDb016	0,3434	1,3782	0,7426
RBDb017	0,333	0,5919	0,696

RBDb022	0,329	1,3828	0,7234
RBDb032	0,3376	1,343	0,5848
RBDb037	0,3267	1,3211	0,7265
RBD01 réplica 1	0,3154	0,8185	0,7775
RBD01 réplica 2	0,3153	0,8147	0,7691
Amortiguador (control negativo)	0,0098	-0,0542	0,0281

[0303]Tabla 19. Análisis NanoDSF de muestras de antígeno ancestral purificado CompA usando fluorescencia de triptófano intrínseca. Cada medición se informa como un promedio de tres réplicas.

5 **Tabla 19:**

Construcciones	Temperatura media de fusión (°C)
RBD01	49,2
RBDb001	49,2
RBDb002	48,8
RBDb012	51,3
RBDb017	50,9
RBDb022	52,2
RBDb032	51,3
RBDb037	48,7

[0304]Tabla 20. Análisis NanoDSF de muestras de antígeno ancestral purificado CompA usando tinción SYPRO. Cada medición se informa como un promedio de tres réplicas.

Tabla 20:

Construcciones	Temperatura media de fusión (°C)	Área promedio de fluorescencia de SYPRO cerca de 30 °C (Recuentos × nm)
RBDb002	44,9	1508685,71
RBDb012	46,6	824854,98

RBDb017	47,4	1353775,56
RBDb022	48,8	902591,02
RBDb032	49,1	1536223,86

[0305] La Tabla 21 muestra los resultados de DLS para el ensamblaje de muestras de antígeno purificado CompA con CompB.

Tabla 21:

Construcciones	Formulación del amortiguador	Z-Ave. Dia. (nm)	PDI	Interpretación de DLS
RBDb001	50 mM de MOPS, 0,75 % de CHAPS, 150 mM de NaCl, pH 7,4	35,70	0,112	Ensamblaje con baja polidispersidad
RBDb001	50 mM de MOPS, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	>1000	0,283	
RBDb001	50 mM de MOPS, 150 mM de NaCl, pH 7,4	>1000	0,427	
RBDb001	50 mM de arginina, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	>1000	0,324	
RBDb001	50 mM de arginina, 150 mM de NaCl, pH 7,4	>1000	0,246	
RBDb002	50 mM de MOPS, 0,75 % de CHAPS, 150 mM de NaCl, pH 7,4	36,48	0,120	Ensamblaje con baja polidispersidad
RBDb002	50 mM de MOPS, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	479,98	0,158	
RBDb002	50 mM de MOPS, 150 mM de NaCl, pH 7,4	>1000	0,541	

RBDb002	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	318,83	0,176	
RBDb002	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, pH 7,4	>1000	0,361	
RBDb017	50 mM de MOPS, 0,75 % de CHAPS, 150 mM de NaCl, pH 7,4	141,95	0,111	
RBDb017	50 mM de MOPS, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	>1000	0,364	
RBDb017	50 mM de MOPS, 150 mM de NaCl, pH 7,4	>1000	0,340	
RBDb017	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	83,35	0,103	
RBDb017	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, pH 7,4	>1000	0,320	
RBDb022	50 mM de MOPS, 0,75 % de CHAPS, 150 mM de NaCl, pH 7,4	36,59	0,100	Ensamblaje con baja polidispersidad
RBDb022	50 mM de MOPS, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	>1000	0,235	
RBDb022	50 mM de MOPS, 150 mM de NaCl, pH 7,4	42,93	0,151	Ensamblaje con baja polidispersidad
RBDb022	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, 4	38,84	0,107	Ensamblaje con baja

	% de sacarosa, pH 7,4			polidispersidad
RBDb022	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, pH 7,4	39,23	0,142	Ensamblaje con baja polidispersidad
RBDb037	50 mM de MOPS, 0,75 % de CHAPS, 150 mM de NaCl, pH 7,4	37,31	0,099	Ensamblaje con baja polidispersidad
RBDb037	50 mM de MOPS, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	317,80	0,106	
RBDb037	50 mM de MOPS, 150 mM de NaCl, pH 7,4	33,91	0,199	Ensamblaje con baja polidispersidad
RBDb037	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	38,68	0,067	Ensamblaje con baja polidispersidad
RBDb037	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, pH 7,4	39,56	0,156	Ensamblaje con baja polidispersidad
RBD01	50 mM de MOPS, 0,75 % de CHAPS, 150 mM de NaCl, pH 7,4	33,57	0,151	Ensamblaje con baja polidispersidad
RBD01	50 mM de MOPS, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	>1000	0,198	
RBD01	50 mM de MOPS, 150 mM de NaCl, pH 7,4	>1000	0,429	
RBD01	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, 4 % de sacarosa, pH 7,4	>1000	0,398	

RBD01	50 mM de MOPS, 50 mM de arginina, 150 mM de NaCl, pH 7,4	>1000	0,309	
RBD01	Tris, 20 mM Tris, 0,75 % de CHAPS, 5 % de glicerol, 250 mM de NaCl, pH 8	32,50	0,164	Ensamblaje con baja polidispersidad
RBD01	Tris, 20 mM Tris, 5 % de glicerol, 250 mM de NaCl, 4 % de sacarosa, pH 8	>1000	0,437	
RBD01	Tris, 20 mM Tris, 5 % de glicerol, 250 mM de NaCl, pH 8	607,76	0,486	
RBD01	Tris, 20 mM Tris, 50 mM de arginina, 5 % de glicerol, 250 mM de NaCl, 4 % de sacarosa, pH 8	40,55	0,290	Ensamblaje con baja polidispersidad
RBD01	Tris, 20 mM Tris, 50 mM de arginina, 5 % de glicerol, 250 mM de NaCl, pH 8	180,49	0,185	

Ejemplo 2: Diseño de mutaciones estabilizadoras adicionales basadas en el análisis estructural de regiones altamente conservadas del antígeno RBD-SD1

[0306] Se generaron construcciones que contenían fragmentos del antígeno RBD-SD1, mutaciones Rpk9, mutaciones SD1resurf y el enlazador de 18 residuos, y estas combinaciones se probaron tanto con la cepa ancestral como con BA.5, una variante derivada de ómicron. La proteína purificada se probó para su ensamblaje con CompB, que funcionó como se esperaba para el antígeno ancestral. (**Figura 6, Tabla 22**) (**Refs 3, 13**).

[0307] Con base en la ubicación estructural de las mutaciones estabilizadoras agregadas y la proximidad de F392W a la interfaz RBD-SD1, se crearon diseños que se basaron en la combinación de mutaciones Rpk8 (igual que Rpk9 pero sin F392W) y mutaciones SD1resurf, mientras que agregaron más mutaciones estabilizadoras para mejorar la estabilidad y el

ensamblaje. Las secuencias RBD que son menos estables térmicamente, tales como las variantes ómicron, fueron la base de otros diseños. Esta menor estabilidad supone un riesgo para la fabricación confiable de antígenos RBD y vacunas basadas en nanoestructuras de proteínas derivadas, y permite realizar pruebas más rigurosas de nuevas mutaciones estabilizadoras. Debido a la relevancia global de la variante BA.5 derivada de ómicron y su representación en vacunas de ARNm bivalentes en 2022, se seleccionó BA.5 para probar nuevas mutaciones estabilizadoras en antígenos RBD-SD1.

[0308] Para seleccionar mejor las mutaciones estabilizadoras que podrían ayudar al máximo a los antígenos RBD-SD1 de diversas cepas de SARS-CoV-2 u otros sarbecovirus, se analizaron datos estructurales de regiones altamente conservadas del antígeno RBD-SD1 que parecían carecer de estabilidad local, tales como cavidades estructurales u otras características no ideales. La interfaz entre el RBD y SD1 se destacó como una región que parecía localmente inestable (**Figura 1C**). Se observaron múltiples cavidades y grupos polares mal formados entre los dos dominios, y una región de bucle en el RBD de los residuos 517-522 parecía mantenerse en tensión contra el dominio SD1 y además contenía grupos hidrófobos expuestos. El análisis estructural de las estructuras de la glucoproteína Spike de múltiples variantes del SARS-CoV-2 y del SARS-CoV-1 mostró una conservación tanto estructural como a nivel de residuos de esta región de interfaz (PDB 6VXX, 7UB5, 5XLR). Para confirmar aún más esta conservación, se alinearon y analizaron secuencias de diversas variantes representativas del SARS-CoV-2 y sarbecovirus. La alineación mostró una fuerte conservación de residuos en esta interfaz (**Figura 7**). Además, no se pudieron identificar anticuerpos neutralizantes potentes caracterizados dirigidos a residuos dentro de esta región, lo que la hizo más susceptible a la mutación sin afectar las antigenicidades clave. Basándose en modelos en PyMol, Rosetta (**Ref 10**) y Colabfold (**Ref 11**), se diseñaron múltiples conjuntos de mutaciones que apuntaban a reforzar la estabilidad de la interfaz RBD-SD1. Algunos conjuntos de mutaciones tenían como objetivo fortalecer las interacciones entre el dominio RBD y SD1 mediante un mejor empaquetamiento y/o enlaces de hidrógeno, o la adición de un enlace disulfuro. Otros mutaron el bucle 517-522 para hacer que los grupos hidrófobos fueran más polares y/o eliminar la tensión en el bucle, como por ejemplo mediante la adición de mutaciones de glicina o la eliminación de prolinas nativas. También se evaluaron combinaciones selectas de diferentes estrategias de diseño. Además, también se probaron muchos diseños en otras partes del RBD que estaban próximas

al RBM (**Figura 1C**), lo que tenía como objetivo mejorar la estabilidad mejorando otras interacciones no covalentes. Se generaron secuencias de ADN para todos los diseños como fusiones genéticas a CompA usando un enlazador de 18 residuos. Como se mencionó anteriormente, se agregaron mutaciones Rpk8 a todos los diseños para respaldar aún más la estabilidad, y también se agregaron mutaciones SD1resurf a todos los diseños para ayudar al ensamblaje de la nanoestructura.

[0309] La **Tabla 22** muestra los resultados de DLS para el ensamblaje de muestras de antígeno purificado CompA con CompB.

Tabla 22:

Construcciones	Formulación del amortiguador	Z-Ave. Dia. (nm)	PDI	Interpretación de DLS
RBDdb092	20 mM Tris, 100 mM de arginina, 5 % de glicerol, 150 mM de NaCl, pH 7,4	38,17	0,083	Ensamblaje con baja polidispersidad
RBDdb093	20 mM Tris, 100 mM de arginina, 5 % de glicerol, 150 mM de NaCl, pH 7,4	548,60	0,179	

Ejemplo 3: Evaluación de mutaciones estabilizadoras y su mayor función sobre la antigenicidad y la estabilidad térmica

[0310] Los diseños fueron seleccionados a partir de transfecciones a pequeña escala en células HEK293, primero mediante transferencia Western (**Figura 8**), que eliminó algunos diseños con mutaciones proximales a RBM basadas en una secreción deficiente. Los diseños de buena expresión se seleccionaron aún más basándose en las señales BLI de los sobrenadantes contra ACE2-Fc (**Tabla 23**), para seleccionar diseños que tuvieran una alta expresión y/o una antigenicidad mejorada. Se seleccionaron múltiples diseños con señales BLI más altas para ampliar y purificar, y se compararon con una secuencia de antígeno silvestre que carecía de mutaciones Rpk8 y SD1resurf. El nanoDSF de diseños que usan fluorescencia intrínseca de triptófano mostró diversos niveles de mejoras en la estabilidad térmica (**Figura 9A, Tabla 24**), con temperaturas de fusión e inicio mejoradas de hasta 6-7°C en relación con RBDdb046 (diseño silvestre). El diseño silvestre mostró temperaturas de fusión e inicio respectivas de

aproximadamente 37°C y 40,5°C. Una estabilidad térmica tan baja genera riesgos para la fabricación debido a las temperaturas elevadas y el estrés asociado, y las mutaciones estabilizadoras proporcionan directamente una forma de mitigar dichos riesgos. Sorprendentemente, la BLI contra los receptores ACE2-Fc inmovilizados mostró señales de unión mejoradas a partir de diseños estabilizados en relación con el diseño silvestre (**Figura 9B** y **Figura 9C, Tabla 24**), particularmente en términos de tasas de unión. Todas las mutaciones estabilizadoras agregadas estaban estructuralmente distantes del sitio de unión de ACE2 en el RBD, lo que sugiere que las mejoras antigénicas se proporcionan a través de un plegamiento mejorado de los antígenos RBD-SD1. Finalmente, se descubrió que un subconjunto de estos diseños reducía la agregación durante el ensamblaje de nanoestructuras de proteínas I53-50 después de la adición de CompB en presencia de excipientes menores (**Figura 10, Tabla 24**), y el diseño silvestre mostró señales DLS que no eran coherentes con un ensamblaje adecuado de nanoestructuras de proteínas.

[0311] La Tabla 23 muestra los cambios en la unión de BLI para la detección de sobrenadante de la unión de ACE2 para los diseños BA.5.

Tabla 23:

Construcción	Desplazamiento	Construcción	Desplazamiento
RBDb001	0,229		
RBDb046	0,2865	RBDb074	0,3063
RBDb047	0,3069	RBDb075	0,3179
RBDb048	0,4857	RBDb076	0,3146
RBDb049	0,3398	RBDb077	0,3302
RBDb050	0,3056	RBDb078	0,3689
RBDb051	0,3252	RBDb079	0,2788
RBDb052	0,331	RBDb080	0,3165
RBDb062	0,3073	RBDb081	0,3815
RBDb063	0,3326	RBDb082	0,2954
RBDb065	0,326	RBDb083	0,2663
RBDb066	0,3235	RBDb084	0,3153
RBDb067	0,2971	RBDb085	0,2805

RBDb068	0,2964	RBDb086	0,2606
RBDb069	0,2858	RBDb087	0,3044
RBDb070	0,2968	RBDb088	0,3044
RBDb071	0,2932	RBDb089	0,2712
RBDb072	0,3163	RBDb090	0,349
RBDb073	0,3306	RBDb091	0,347

[0312] La Tabla 24 proporciona un resumen de los análisis *in vitro* para muestras de antígeno-CompA basado en BA.5 purificado. Las temperaturas de inicio y de fusión fueron promedios de tres réplicas, con valores calculados según un método de análisis personalizado descrito en

5 Materiales y métodos. “NC” significa “no recolectado” debido a que las construcciones etiquetadas posteriormente tenían otras propiedades inferiores. Los valores deseados de Z-Ave.d varían de 30 a 45 nm, y los valores más grandes sugieren agregación durante el ensamblaje de la nanoestructura.

Tabla 24:

Construcción	Temperatura de inicio (°C)	Temperatura de fusión (°C)	Rendimiento (g/L)	k_{eq}	Desplazamiento máximo (nm)	Z-Ave.d (nm)	PDI
RBDb046	36,9	40,5	0,066	0,032	0,271	58,3	0,164
RBDb047	38,9	43,2	0,042	0,106	0,277	32,67	0,25
RBDb052	35,9	40,1	0,099	0,043	0,282	N.C.	N.C.
RBDb063	39,4	42,5	0,092	0,047	0,279	N.C.	N.C.
RBDb066	39,7	43,7	0,049	0,097	0,282	36,33	0,183
RBDb072	39,4	42,9	0,076	0,04	0,280	N.C.	N.C.

				4			
RBDb073	40,9	44,9	0,077	0,08 9	0,288	38,16	0,16 1
RBDb075	43,1	47,3	0,089	0,06 7	0,270	36,67	0,15 7
RBDb076	40,5	44,2	0,072	0,05 4	0,288	38,53	0,15
RBDb078	41,9	46,1	0,114	0,10 9	0,287	>1000	0,11 3
RBDb081	43,8	47,7	0,093	0,11 6	0,293	>1000	0,25 2
RBDb084	40,3	43,8	0,095	0,02 3	0,293	N.C.	N.C.
RBDb087	42,2	46,2	0,091	0,06 5	0,282	40,65	0,10 1
RBDb090	42,5	46,2	0,239	0,03 8	0,271	43,6	0,13 5

Ejemplo 4: Inmunogenicidad de mutaciones estabilizadoras seleccionadas

- [0313] A continuación se probó la inmunogenicidad de un pequeño conjunto de nanoestructuras proteicas ensambladas y un conjunto más grande de CompAs triméricos. Para las
- 5 construcciones basadas en el RBD ancestral, se probaron nanoestructuras de proteínas y CompA para RBD01 (WT, sin SD1) y RBDb022 (que contiene mutaciones estabilizadoras, con SD1). Para las construcciones basadas en el RBD BA.5, solo se probó RBDb047 como una nanoestructura de proteína, mientras que RBDb046 (WT, con SD1), RBDb047, RBDb075 y RBDb078 (los tres contienen mutaciones estabilizadoras, con SD1) se probaron como CompAs.
- 10 Los ratones fueron inmunizados con un régimen de dos dosis los días 0 y 21, con dosis de nanoestructuras de proteína de 0,027-0,030 μg (adaptadas a moles equivalentes de antígeno entre cada nanoestructura de proteína) y dosis de CompA de 0,85-1,0 μg (adaptadas a moles equivalentes de antígeno entre cada CompA), con todos los grupos adyuvados con Addavax™. Los títulos de unión se midieron frente a antígenos de glucoproteína Spike coincidentes con la

cepa para cepas ancestrales y BA.5 (**Figura 11A**), y los títulos de neutralización de pseudovirus también se midieron frente a pseudovirus ancestrales y BA.5 coincidentes con la cepa (**Figura 11B**). Se realizaron sangrados para mediciones serológicas los días 21 (después de la inmunización primaria) y 35 (después del refuerzo). Dentro de los CompA ancestrales,

5 RBDb022 mantuvo títulos de unión ligeramente más altos que RBD01 tanto en el día 21 como en el 35. Si bien no se detectaron títulos neutralizantes el día 21, el RBDb022 CompA también mostró títulos neutralizantes más altos en comparación con el RBD01 CompA el día 35. Esto fue sorprendente, ya que las comparaciones previas entre inmunógenos similares a RBD01 CompAs y diferentes RBD estabilizados que carecían del dominio SD1 no habían mostrado

10 diferencias en los resultados publicados (**Ref 5**). Los títulos de unión y neutralización fueron generalmente similares entre los grupos de nanoestructuras de proteínas RBD01 y RBDb022. Sorprendentemente, se observaron diferencias sorprendentes entre los CompA para los antígenos BA.5. Entre los CompA, solo RBDb075 mostró una señal clara por encima del valor inicial en el día 21. En el día 35, los títulos de unión fueron más altos para CompAs con

15 antígenos estabilizados en comparación con RBDb046, y se demostró que RBDb075 tenía los títulos de unión más altos. Si bien no se detectó actividad neutralizante para BA.5 CompAs el día 21, se observaron diferencias dramáticas el día 35 con RBDb046 mostrando respuestas cercanas a las iniciales, mientras que todos los grupos CompA con mutaciones estabilizadoras mostraron respuestas más altas. Lo más sorprendente es que RBDb075 mostró títulos

20 neutralizantes más altos que RBDb047 y RBDb078. Estos datos muestran que el conjunto específico de mutaciones estabilizadoras agregadas proporcionó una ayuda única a la inmunogenicidad además de mejoras en la estabilidad térmica, la antigenicidad y el ensamblaje de la nanoestructura, cuya combinación no podría haberse predicho basándose en la secuencia de este antígeno usando los métodos actuales. La nanoestructura de la proteína RBDb047 fue

25 más inmunogénica que la CompAs en todas las mediciones.

Ejemplo 5: Antigenicidad, expresión y temperatura de fusión de mutaciones estabilizadoras seleccionadas aplicadas a diversas variantes de antígenos RBD-SD1

[0314] Para entender cómo funcionan las mutaciones estabilizadoras usadas en RBDb075 en

30 RBD de otras variantes, se generaron diseños equivalentes para otros RBD variantes usando las mismas mutaciones estabilizadoras, fragmento de antígeno (puntos de inicio y final) y

enlazador a CompA. Las variantes elegidas para el análisis incluyeron la variante ancestral, BA.1, BA.5, XBB y SARS-CoV-1. Como controles, estos diseños estabilizados se compararon con diseños equivalentes que carecían de mutaciones estabilizadoras, usándose RBDb002 y RBDb046 respectivamente para las variantes ancestrales y BA.5.

5 **[0315]** Los cinco diseños estabilizados y los cinco diseños silvestre equivalentes se secretaron primero a partir de células HEK293 a pequeña escala, y se evaluó la antigenicidad mediante BLI frente a ACE2, S309 y 16A8 (anti-CompA). Se compararon las proporciones de las señales de unión contra ACE2 y 16A8 para cada muestra. Para BA.1, BA.5 y XBB, la relación de unión ACE2/16A8 aumentó notablemente para las construcciones modificadas en relación con las
10 construcciones silvestres, lo que sugiere que las mutaciones estabilizadoras estaban mejorando el plegamiento del antígeno para permitir mejor la unión de ACE2. Las proporciones generalmente no cambiaron para las construcciones derivadas de la variante ancestral o SARS-CoV-1, lo que sugiere que estas variantes de RBD ya estaban bien plegadas en su forma silvestre (**Figura 12A y Tabla 25**). Se observaron resultados similares para las proporciones de
15 S309 a 16A8.

[0316] A continuación, se ampliaron y purificaron los diez diseños mediante IMAC, y se compararon los rendimientos de proteína total entre los diseños estabilizados y silvestres de cada variante. Para BA.1, BA.5 y XBB, los rendimientos mejoraron claramente con respecto a los diseños estabilizados, observándose mejoras especialmente significativas con BA.1 y BA.5.
20 Los rendimientos fueron generalmente similares entre los diseños silvestres y estabilizados de las variantes ancestrales y del SARS-CoV-1 (**Figura 12B y Tabla 25**).

[0317] Finalmente, se midieron las temperaturas de fusión de la proteína purificada mediante fluorescencia intrínseca del triptófano y se compararon entre los RBD de cada variante. La construcción BA.1 silvestre no se analizó debido a los bajos rendimientos que impidieron
25 realizar análisis adicionales. Para cada uno de los cuatro pares restantes de RBD variantes, se mejoró la temperatura de fusión para todos los antígenos estabilizados. Las mejoras variaron de 4,7 °C para SARS-CoV-1 a 6,8 °C para BA.5 (**Figura 12C y Tabla 25**).

[0318] En total, los elementos de diseño estabilizadores establecidos con RBDb075 se generalizaron con éxito a RBD de diversas variantes, no mostraron un impacto perjudicial en
30 cada RBD variante y mostraron mejoras en la antigenicidad, el rendimiento y/o la estabilidad térmica para cada RBD variante.

[0319] La Tabla 25 muestra la generalización de los diseños usados en RBDb002 (“silvestre”) y RBDb075 a diversas secuencias de cepas de SARS-CoV-2 y SARS-CoV-1. Las mediciones de unión de ACE2 normalizadas promedio se basaron en BLI y se calcularon a partir de dos mediciones replicadas de la señal de unión de ACE2 dividida por la señal de unión a un mAb anti-CompA. Los rendimientos se midieron después de la purificación con IMAC a partir de volúmenes equivalentes de cultivo de células transfectadas. Las temperaturas de fusión se midieron usando la fluorescencia intrínseca del triptófano y se midieron en función de los valores calculados por el software UNcle. “NC” significa “no recolectado”.

Tabla 25:

Construcciones	Cepa	Silvestre o estabilizado	Unión ACE2 normalizada promedio (nm/nm)	Rendimiento (mg/L)	Temperatura de fusión (°C)
RBDb002	Ancestral	Silvestre	0,128	55	48,0
RBDb213	Ancestral	Estabilizado	0,137	53	52,9
RBDb094	BA.1	Silvestre	0,008	5	N.C.
RBDb155	BA.1	Estabilizado	0,096	36	43,4
RBDb046	BA.5	Silvestre	0,042	18	40,5
RBDb075	BA.5	Estabilizado	0,085	57	47,3
RBDb270	XBB	Silvestre	0,136	33	46,9
RBDb272	XBB	Estabilizado	0,212	45	53,0
RBDb329	SARS-CoV-1	Silvestre	0,137	36	56,9
RBDb331	SARS-CoV-1	Estabilizado	0,150	33	61,6

10

[0320] La Tabla 26 muestra las construcciones usadas en los experimentos.

Tabla 26:

Construcción	Cepa	Mutaciones	Enlazador
RBDb001	Ancestral	Ninguna	GGSGGSGSGGSGGSGS

RBDb002	Ancestral	Ninguna	SGGGSGGSGSGGGSGGSGS
RBDb003	Ancestral	Ninguna	EPEGGSGGSGSGGGSGGSGS
RBDb004	Ancestral	Ninguna	YGGSGGSGGSGSGGGSGGSGS
RBDb005	Ancestral	Ninguna	GSGGSGGSGGSGGSGSGGSG GSGS
RBDb006	Ancestral	F338L/Y365W	GGSGGSGSGGGSGGSGS
RBDb007	Ancestral	F338L/Y365W	SGGGSGGSGSGGGSGGSGS
RBDb008	Ancestral	F338L/Y365W	EPEGGSGGSGSGGGSGGSGS
RBDb009	Ancestral	F338L/Y365W	YGGSGGSGGSGSGGGSGGSGS
RBDb010	Ancestral	F338L/Y365W	GSGGSGGSGGSGGSGSGGSG GSGS
RBDb011	Ancestral	F392W	GGSGGSGSGGGSGGSGS
RBDb012	Ancestral	F392W	SGGGSGGSGSGGGSGGSGS
RBDb013	Ancestral	F392W	EPEGGSGGSGSGGGSGGSGS
RBDb014	Ancestral	F392W	YGGSGGSGGSGSGGGSGGSGS
RBDb015	Ancestral	F392W	GSGGSGGSGGSGGSGSGGSG GSGS
RBDb016	Ancestral	Y365F/V395I	GGSGGSGSGGGSGGSGS
RBDb017	Ancestral	Y365F/V395I	SGGGSGGSGSGGGSGGSGS
RBDb018	Ancestral	Y365F/V395I	EPEGGSGGSGSGGGSGGSGS
RBDb019	Ancestral	Y365F/V395I	YGGSGGSGGSGSGGGSGGSGS
RBDb020	Ancestral	Y365F/V395I	GSGGSGGSGGSGGSGSGGSG GSGS
RBDb021	Ancestral	Y365F/F392W/V3 95I	GGSGGSGSGGGSGGSGS
RBDb022	Ancestral	Y365F/F392W/V3 95I	SGGGSGGSGSGGGSGGSGS
RBDb023	Ancestral	Y365F/F392W/V3 95I	EPEGGSGGSGSGGGSGGSGS
RBDb024	Ancestral	Y365F/F392W/V3	YGGSGGSGGSGSGGGSGGSGS

		95I	
RBDb025	Ancestral	Y365F/F392W/V3 95I	GSGGSGGSGGSGGSGSGGSG GSGS
RBDb026	Ancestral	F338L/A363L/Y3 65M	GGSGGSGSGGSGGSGS
RBDb027	Ancestral	F338L/A363L/Y3 65M	SGGGSGGSGSGGSGGSGS
RBDb028	Ancestral	F338L/A363L/Y3 65M	EPEGGSGGSGSGGSGGSGS
RBDb029	Ancestral	F338L/A363L/Y3 65M	YGGSGGSGGSGSGGSGGSGS
RBDb030	Ancestral	F338L/A363L/Y3 65M	GSGGSGGSGGSGGSGSGGSG GSGS
RBDb031	Ancestral	I358F/Y365F/V39 5I	GGSGGSGSGGSGGSGS
RBDb032	Ancestral	I358F/Y365F/V39 5I	SGGGSGGSGSGGSGGSGS
RBDb033	Ancestral	I358F/Y365F/V39 5I	EPEGGSGGSGSGGSGGSGS
RBDb034	Ancestral	I358F/Y365F/V39 5I	YGGSGGSGGSGSGGSGGSGS
RBDb035	Ancestral	I358F/Y365F/V39 5I	GSGGSGGSGGSGGSGSGGSG GSGS
RBDb036	Ancestral	L560Q/F562Y/I56 9S	GGSGGSGSGGSGGSGS
RBDb037	Ancestral	L560Q/F562Y/I56 9S	SGGGSGGSGSGGSGGSGS
RBDb038	Ancestral	L560Q/F562Y/I56 9S	EPEGGSGGSGSGGSGGSGS
RBDb039	Ancestral	L560Q/F562Y/I56	YGGSGGSGGSGSGGSGGSGS

		9S	
RBDb040	Ancestral	L560Q/F562Y/I569S	GSGGSGGSGSGGSGGSGSGSGSGS
RBDb041	Ancestral	L452R	GGSGGSGSGSGSGSGS
RBDb042	Ancestral	L452R	SGGGSGGSGSGSGSGSGS
RBDb043	Ancestral	L452R	EPEGSGGSGSGSGSGSGS
RBDb044	Ancestral	L452R	YGGSGGSGGSGSGSGSGSGS
RBDb045	Ancestral	L452R	GSGGSGGSGSGSGSGSGSGSGS
RBDb046	BA.5	Ninguna	SGGGSGGSGSGSGSGSGS
RBDb047	BA.5	Y365F/V395I/L560Q/F562Y/I569S	SGGGSGGSGSGSGSGSGS
RBDb048	BA.2	Y365F/V395I/L560Q/F562Y/I569S	SGGGSGGSGSGSGSGSGS
RBDb049	BA.2.12.1	Y365F/V395I/L560Q/F562Y/I569S	SGGGSGGSGSGSGSGSGS
RBDb050	BA.2.75	Y365F/V395I/L560Q/F562Y/I569S	SGGGSGGSGSGSGSGSGS
RBDb051	BA.5	Y365F/V395I/V407L/L560Q/F562Y/I569S	SGGGSGGSGSGSGSGSGS
RBDb052	BA.5	Y365F/V395I/I402V/V407L/L560Q/F562Y/I569S	SGGGSGGSGSGSGSGSGS
RBDb053	BA.5	Y365F/V395I/I402V/ I410F/L560Q/F562Y/I569S	SGGGSGGSGSGSGSGSGS
RBDb054	BA.5	Y365F/V395I/I402V/ I410F/A435V/V512I/L560Q/F56	SGGGSGGSGSGSGSGSGS

		2Y/I569S	
RBDb055	BA.5	Y365F/V395I/I40 2V/V407L/ I410F/ A435V/V512I/L56 0Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb056	BA.5	Y365F/V395I/V35 0L/I402V/V407L/ I 410F/I418V/A435 V/V512I/L560Q/F 562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb057	BA.5	Y365F/V395I/V35 0L/I402V/V407L/ I 410F/I418V/A435 I/V510I/V512I/L56 0Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb058	BA.5	Y365F/V395I/V35 0L/I402V/V407L/ I 410F/I418V/F429 W/V433I/A435V/ V510I/V512F/L56 0Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb059	BA.5	Y365F/V395I/I40 2V/V407L/I410F/ L560Q/F562Y/I56 9S	SGGGSGGSGSGGGSGGSGS
RBDb060	BA.5	Y365F/V395I/I40 2V/V407L/I410F/ V512I/L560Q/F56 2Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb061	BA.5	Y365F/V395I/I40 2V/I410F/V512I/L	SGGGSGGSGSGGGSGGSGS

		560Q/F562Y/I569 S	
RBDb062	BA.5	Y365F/V395I/A34 8P/L560Q/F562Y/ I569S	SGGGSGGSGSGGGSGGSGS
RBDb063	BA.5	Y365F/V395I/I40 2V/L560Q/F562Y/ I569S	SGGGSGGSGSGGGSGGSGS
RBDb064	BA.5	Y365F/V395I/S51 4T/L560Q/F562Y/ I569S	SGGGSGGSGSGGGSGGSGS
RBDb065	BA.5	Y365F/V395I/I40 2V/F464Y/L560Q/ F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb066	BA.5	Y365F/V395I/F32 9Y/L560Q/F562Y/ I569S	SGGGSGGSGSGGGSGGSGS
RBDb067	BA.5	Y365F/V395I/G52 6S/P527N/L560Q /F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb068	BA.5	Y365F/V395I/G52 6S/P527N/K528T /L560Q/F562Y/I5 69S	SGGGSGGSGSGGGSGGSGS
RBDb069	BA.5	Y365F/V395I/F32 9Y/G526S/P527N /K528T/L560Q/F5 62Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb070	BA.5	Y365F/V395I/F32 9Y/G526S/P527N /K528I/L560Q/F5	SGGGSGGSGSGGGSGGSGS

		62Y/I569S	
RBDb071	BA.5	Y365F/V395I/F32 9K/K528Q/L560Q /F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb072	BA.5	Y365F/V395I/F32 9R/K528Q/L560Q /F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb073	BA.5	Y365F/V395I/L51 7T/H519S/L560Q /F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb074	BA.5	Y365F/V395I/L51 7A/H519S/L560Q /F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb075	BA.5	Y365F/V395I/L51 7T/H519S/A520G /L560Q/F562Y/I5 69S	SGGGSGGSGSGGGSGGSGS
RBDb076	BA.5	Y365F/V395I/L51 7T/H519S/N544L/ L560Q/F562Y/I56 9S	SGGGSGGSGSGGGSGGSGS
RBDb077	BA.5	Y365F/V395I/L51 7T/H519S/A520C /Q564C/L560Q/F 562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb078	BA.5	Y365F/V395I/L51 7T/H519G/A520C /Q564C/L560Q/F 562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb079	BA.5	Y365F/V395I/L51 7D/H519T/A520D	SGGGSGGSGSGGGSGGSGS

		/P512Q/N544L/L5 46V/Q564L/F562 T/L560Q/F562Y/I 569S	
RBDb080	BA.5	Y365F/V395I/L51 7S/H519T/A520D /P521Q/N544L/L5 46V/Q564L/F562 T/L560Q/F562Y/I 569S	SGGGSGGSGSGGGSGGSGS
RBDb081	BA.5	Y365F/V395I/L51 7S/H519R/A520H /P521Q/N544L/L5 46V/Q564W/F562 T/L560Q/F562Y/I 569S	SGGGSGGSGSGGGSGGSGS
RBDb082	BA.5	Y365F/V395I/L51 7S/H519T/A520D /N544L/Q564L/F5 62T/L560Q/F562 Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb083	BA.5	Y365F/V395I/F32 9Y/G526S/P527N /K528T/L517S/H5 19T/A520D/P521 Q/N544L/L546V/ Q564L/F562T/L5 60Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb084	BA.5	Y365F/V395I/F32 9Y/G526S/P527N /K528I/L517S/H5	SGGGSGGSGSGGGSGGSGS

		19T/A520D/P521 Q/N544L/L546V/ Q564L/F562T/L5 60Q/F562Y/I569S	
RBDb085	BA.5	Y365F/V395I/F32 9Y/G526S/P527N /K528T/L517D/H5 19T/A520D/P512 Q/N544L/L546V/ Q564L/F562T/L5 60Q/F562Y/I569S	SGGGSGGGSGGGSGGGSGS
RBDb086	BA.5	Y365F/V395I/F32 9Y/G526S/P527N /K528T/L517S/L5 17S/H519R/A520 H/P521Q/N544L/ L546V/Q564W/F5 62T/L560Q/F562 Y/I569S	SGGGSGGGSGGGSGGGSGS
RBDb087	BA.5	Y365F/V395I/F32 9Y/L517T/ H519S /L560Q/F562Y/I5 69S	SGGGSGGGSGGGSGGGSGS
RBDb088	BA.5	Y365F/V395I/F32 9K/ K528Q/ L517 T/ H519S/L560Q/ F562Y/I569S	SGGGSGGGSGGGSGGGSGS
RBDb089	BA.5	Y365F/V395I/F32 9Y/G526S/P527N /K528T/L517T/H5 19S/L560Q/F562	SGGGSGGGSGGGSGGGSGS

		Y/I569S	
RBDb090	BA.5	Y365F/V395I/V367F/L560Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb091	BA.5	Y365F/V395I/F375Y/L560Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb092	Ancestral	Y365F/F392W/V395I/L560Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb093	BA.5	Y365F/F392W/V395I/L560Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb094	BA.1	Ninguna	SGGGSGGSGSGGGSGGSGS
RBDb155	BA.1	Y365F/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb213	Ancestral	Y365F/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb270	XBB	Ninguna	SGGGSGGSGSGGGSGGSGS
RBDb272	XBB	Y365F/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S	SGGGSGGSGSGGGSGGSGS
RBDb329	SARS-CoV-1	Ninguna	SGGGSGGSGSGGGSGGSGS
RBDb331	SARS-CoV-	Y365F/V395I/L51	SGGGSGGSGSGGGSGGSGS

	1	7T/H519S/A520G /L560Q/F562Y/I5 69S	
--	---	--	--

[0321] En algunas modalidades, la secuencia señal se escinde postraduccionalmente cuando se expresa la proteína. En algunas modalidades, la secuencia señal subrayada/cursiva y etiqueta poli-His se incluyen opcionalmente en la secuencia.

- 5 **[0322]** La **Tabla 27** proporciona la secuencia de construcciones usadas en los experimentos (con la secuencia de señales y las etiquetas subrayadas/en cursiva).

Tabla 27:

Construcciones	Sec. AA (con señal sec. y etiquetas)
RBDb001	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLVNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGGS GGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRANS VEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIGA GTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPGV MTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFNPVKFVP TGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFVE KIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 111)
RBDb002	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLVNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFNPVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 112)

RBDb003	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSEPE GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPFQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 113)
RBDb004	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSYGG SGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPFQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 114)

RBDb005	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGSG GSGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKH KIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSV LKEKGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKE KGVFYMPGVMTPTELVKAMKLGHTILKLFPGEVVGPFQVKAMK GPFPNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDE VREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 115)
RBDb006	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVADWSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGGS GSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRANS VEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIGA GTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPGV MTPTELVKAMKLGHTILKLFPGEVVGPFQVKAMKGPFPNVKFVP TGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFVE KIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 116)

RBDb007	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVADWSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTTELVKAMKLGHTILKLFPGEVVGPFVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 117)
RBDb008	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVADWSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSEPE GGSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 118)

RBDb009	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVADWSVLYNSASF STFKCYGVSP TKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKIA DYN YKL PDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTE IYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QP YRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSYGG SGGSGGSGSGGSGGSGGSEKA AKAEEAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AII GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 119)
RBDb010	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVADWSVLYNSASF STFKCYGVSP TKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKIA DYN YKL PDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTE IYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QP YRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGSG GSGGSGGSGGSGGSGGSGGSGGSEKA AKAEEAARKMEELFKKH KIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSV LKEKGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKE KGVFYMPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMK GPF PNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDE VREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 120)

RBDb011	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCWTNVYADSFVIRGDEVQRQIAPGQTGKI ADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNL KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG YQPYRVVVLSEFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSG SGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRAN SVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIG AGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPG VMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPNVKFV PTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFV EKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 121)
RBDb012	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCWTNVYADSFVIRGDEVQRQIAPGQTGKI ADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNL KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG YQPYRVVVLSEFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 122)

RBDb013	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCWTNVYADSFVIRGDEVQRQIAPGQTGKI ADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNL KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG YQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSEP EGSGSGSGSGSGSGSGSGSEKAAKAEEAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIAGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 123)
RBDb014	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCWTNVYADSFVIRGDEVQRQIAPGQTGKI ADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNL KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG YQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSYG GSGSGSGSGSGSGSGSGSGSEKAAKAEEAARKMEELFKKHKIVA VLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKE KGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGV FYMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGP PNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVRE KAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 124)

RBDb015	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCWTNVYADSFVIRGDEVQRQIAPGQTGKI ADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNL KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG YQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGS GGSGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKK HKIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALS VLKEKGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAK EKGVFYMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAM KGPFPNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPD EVREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 125)
RBDb016	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGGS GGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRANS VEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIGA GTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPGV MTPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPNVKFVP TGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFVE KIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 126)

RBDb017	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEV RQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLEITFTVPDADTVIKALSVLKEKGAI I GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 127)
RBDb018	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEV RQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSEPE GGSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 128)

RBDb019	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSYGG SGGSGGSGSGGSGGSGGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPFQVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 129)
RBDb020	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGSG GSGGSGGSGGSGGSGGSGGSGGSEKAAKAEAAARKMEELFKKH KIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSV LKEKGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKE KGVFYMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPFQVKAMK GPFPNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDE VREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 130)

RBDb021	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCWTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGGS GGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRANS VEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIGA GTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPGV MTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPNVKFVP TGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFVE KIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 131)
RBDb022	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCWTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 132)

RBDb023	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCWTNIYADSFVIRGDEV RQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSEPE GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 133)
RBDb024	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCWTNIYADSFVIRGDEV RQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCS YGG SGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 134)

RBDb025	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCWTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGSG GSGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKH KIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSV LKEKGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKE KGVFYMPGVMTPTLVKAMKLGHTILKLFPGEVVGPFQFVKAMK GPFPNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDE VREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 135)
RBDb026	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVLDMSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGGS GSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRANS VEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIGA GTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPGV MTPTLVKAMKLGHTILKLFPGEVVGPFQFVKAMKGPFPNVKFVP TGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFVE KIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 136)

RBDb027	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVLDMSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTTELVKAMKLGHTILKLFPGEVVGPFVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 137)
RBDb028	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVLDMSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSEPE GGSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 138)

RBDb029	<u>MGILPSPGMPALLSLVSLLSVLLMGCAETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVLDMSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSYGG SGGSGGSGSGGSGGSGGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPFQVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 139)
RBDb030	<u>MGILPSPGMPALLSLVSLLSVLLMGCAETGSRVQPTESIVRFP</u> NITNLCPLGEVFNATRFASVYAWNRRKRISNCVLDMSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGY QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGSG SGGSGGSGGSGGSGGSGGSGGSGGSEKAAKAEAAARKMEELFKKH KIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSV LKEKGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKE KGVFYMPGVMTPTELVKAMKLGHTILKLFPGEVVGPFQVKAMK GPFPNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDE VREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 140)

RBDb031	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKFSNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRANS VEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIGA GTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPGV MTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPNVKFVP TGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFVE KIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 141)
RBDb032	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKFSNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 142)

RBDb033	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKFSNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEV RQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSEPE GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 143)
RBDb034	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKFSNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEV RQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCS YGG SGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 144)

RBDb035	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKFSNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGSG SGGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKH KIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSV LKEKGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKE KGVFYMPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMK GPPFNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDE VREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 145)
RBDb036	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSGG SGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRAN SVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIG AGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPG VMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPPFNVKFV PTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFV EKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 146)

RBDb037	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 147)
RBDb038	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSEP EGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 148)

RBDb039	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSYG GSGGSGGSGSGGSGGSGGSGSEKAAKAEAAARKMEELFKKHIVA VLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKE KGAIIAGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGV FYMPGVMPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGP PNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVRE KAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 149)
RBDb040	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSGS GGSGGSGGSGGSGGSGGSGGSGGSGSEKAAKAEAAARKMEELFKK HKIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALS VLKEKGAIIAGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAK EKGVFYMPGVMPTTELVKAMKLGHTILKLFPGEVVGPPQFVKAM KGPPPNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPD EVREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 150)

RBDb041	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRANS VEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIGA GTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPGV MTPTELVKAMKLGHTILKLFPGEVVGPFQVKAMKGPFPNVKFVP TGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFVE KIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 151)
RBDb042	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTELVKAMKLGHTILKLFPGEVVGPFQVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 152)

RBDb043	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSEPE GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 153)
RBDb044	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG Y QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCS YGG SGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 154)

RBDb045	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASF STFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSGSG GSGGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKH KIVAVLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSV LKEKGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKE KGVFYMPGVMTPTELVKAMKLGHTILKLFPGEVVGPFVKAMK GPFPNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDE VREKAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 155)
RBDb046	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADYSVLYNFAPF FAFKCYGVSPTKLNDLCFTNVYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTELVKAMKLGHTILKLFPGEVVGPFVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 156)

RBDb047	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDISKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 157)
RBDb048	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDISKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGFNCYFPLRSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 158)

RBDb049	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYQYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGFNCYFPLRSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 159)
RBDb050	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFHEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVSGNYNYLYRLFRKSKLKP FERDISTEIYQAGNKPCNGVAGFNCYFPLQSYGFRPTYGVGHQ PYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTG VLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSGG GSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 160)

RBDb051	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNELSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLD SKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 161)
RBDb052	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNELSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLD SKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 162)

RBDb053	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNEVSQFAPGQTGNI ADYNYKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNL KPFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVG HQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLT GTGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCS SGGGSGGSGSGGGSGGSGSEKAAKAEAAARKMEELFKKHKIVA VLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKE KGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGV FYMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPF PNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVRE KAKAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 163)
RBDb054	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNEVSQFAPGQTGNI ADYNYKLPDDFTGCVIVWNSNKLDSKVGGNYNYRYRLFRKSNL KPFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVG HQPYRVVILSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGGSGGSGSGGGSGGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPF NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 164)

RBDb055	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNELSQQFAPGQTGNI ADYNYKLPDDFTGCVIVWNSNKLDSKVGGNYNYRYRLFRKSNL KPFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVG HQPYRVVILSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGSGGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 165)
RBDb056	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASLYAWNRRKRISNCVADFSVLYNFAPFF AFKCYGVSPTKLNDLCFTNIYADSFVVRGNELSQQFAPGQTGNVA DYNKLPDDFTGCVIVWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVILSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 166)

RBDb057	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASLYAWNRRKRISNCVADFSVLYNFAPFF AFKCYGVSPTKLNDLCFTNIYADSFVVRGNELSQFAPGQTGNVA DYNKLPDDFTGCVIIWNSNKLDSKVGGNYNYRYRLFRKSNLKP FERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGHQ PYRIVILSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGV LTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSGGG SGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRAN SVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIG AGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPG VMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPNVKFV PTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFV EKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 167)
RBDb058	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASLYAWNRRKRISNCVADFSVLYNFAPFF AFKCYGVSPTKLNDLCFTNIYADSFVVRGNELSQFAPGQTGNVA DYNKLPDDWTGCIIVWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRIVFLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 168)

RBDb059	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNELSQFAPGQTGNI ADYNYKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNL KPFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVG HQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLT GTGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCS SGGSGGSGSGGSGGSGGSGSEKAAKAEAAARKMEELFKKHKIVA VLRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKE KGAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGV FYMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPF PNVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVRE KAKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 169)
RBDb060	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNELSQFAPGQTGNI ADYNYKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNL KPFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVG HQPYRVVILSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGGSGGSGSGGSGGSGGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPF NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 170)

RBDb061	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNEVSQFAPGQTGNI ADYNYKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNL KPFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVG HQPYRVVILSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGSGGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 171)
RBDb062	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFPSVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVVLVSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 172)

RBDb063	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 173)
RBDb064	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLTFFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 174)

RBDb065	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVVRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PYERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 175)
RBDb066	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 176)

RBDb067	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCSNKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 177)
RBDb068	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCSNTKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 178)

RBDb069	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCSNTKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 179)
RBDb070	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCSNIKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 180)

RBDb071	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRKP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPQKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGGSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 181)
RBDb072	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRRP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPQKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGGSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 182)

RBDb073	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFETLSAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 183)
RBDb074	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFEALSAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 184)

RBDb075	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLD SKVGGNYNYRYRLFRKSNLK PFERDISTEIIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFETLSGPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGSGGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPPF NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 185)
RBDb076	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLD SKVGGNYNYRYRLFRKSNLK PFERDISTEIIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFETLSAPATVCGPKKSTNLVKNKCVNFNGLGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPPFN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 186)

RBDb077	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFETLSCPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQCFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLPGEVVGPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 187)
RBDb078	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFETLGCPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQCFGRDSADTTDAVRDPQTLEILDITPCSS GGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 188)

RBDb079	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDISKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFEDLTDQATVCGPKKSTNLVKNKCVNFNFLGVTG TGVLTESNKKFQPTQLFGRDSADTTDAVRDPQTLEILDITPCSS GGGSGGSGSGGSGGSGGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 189)
RBDb080	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDISKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFESLTDQATVCGPKKSTNLVKNKCVNFNFLGVTG TGVLTESNKKFQPTQLFGRDSADTTDAVRDPQTLEILDITPCSS GGGSGGSGSGGSGGSGGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 190)

RBDb081	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFESLRHQATVCGPKKSTNLVKNKCVNFNGLGVTG TGVLTESNKKFQPTQWFGGRDSADTTDAVRDPQTLEILDITPCSS GGSGGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 191)
RBDb082	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFESLTDPATVCGPKKSTNLVKNKCVNFNGLGLTGT GVLTESNKKFQPTQLFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 192)

RBDb083	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFESLTDQATVCSNTKSTNLVKNKCVNFNFLGVTGT GVLTESNKKFQPTQLFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 193)
RBDb084	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFESLTDQATVCSNIKSTNLVKNKCVNFNFLGVTGT GVLTESNKKFQPTQLFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 194)

RBDb085	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFEDLTDQATVCSNTKSTNLVKNKCVNFNFLGVTGT GVLTESNKKFQPTQLFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 195)
RBDb086	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFESLRHQATVCSNTKSTNLVKNKCVNFNFLGVTG TGVLTESNKKFQPTQWFGGRDSADTTDAVRDPQTLEILDITPCSS GGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 196)

RBDb087	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLD SKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFETLSAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 197)
RBDb088	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRKP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLD SKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFETLSAPATVCGPQKSTNLVKNKCVNFNFNGLTGT TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTELVKAMKLGHTILKLPGEVVGPPQFVKAMKGPFP NVKFPVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 198)

RBDb089	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRYP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFETLSAPATVCSNTKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 199)
RBDb090	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSFLYNFAPF FAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 200)

RBDb091	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRKRISNCVADFSVLYNFAPF YAFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 201)
RBDb092	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCWTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 202)

RBDb093	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNFAPF FAFKCYGVSPTKLNDLCWTNIYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVGGNYNYRYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGVNCYFPLQSYGFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSG GGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVL RANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKG AIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFY MPGVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPN VKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKA KAFVEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 203)
RBDb094	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADYSVLYNLAPF FTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVSGNLYNYLYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGFNCYFPLRSYSFRPTYGVGH QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT GVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGG GSGSGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTELVKAMKLGHTILKLFPGEVVGPPQFVKAMKGPFPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE<u>LEHHHHHH</u> (SEQ ID NO: 265)

RBDb155	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFDEVFNATRFASVYAWNRRKRISNCVADFSVLYNLAPF FTFKCYGVSPTKLNDLCFTNIYADSFVIRGDEVQRQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSKVSGNYYLYRLFRKSNLK PFERDISTEIYQAGNKPCNGVAGFNCYFPLRSYSFRPTYGVGH QPYRVVLSFETLSGPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGGSGGSGSGGSGGSGGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 266)
RBDb213	<u>MGILPSPGMPALLSLVSLLSVLLMGCVAETGSRVQPTESIVRFP</u> NITNLCPFGEVFNATRFASVYAWNRRKRISNCVADFSVLYNSASF STFKCYGVSPTKLNDLCFTNIYADSFVIRGDEVQRQIAPGQTGKIA DYNKLPDDFTGCVIAWNSNNLDSKVGGNYYLYRLFRKSNLK PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVG QPYRVVLSFETLSGPATVCGPKKSTNLVKNKCVNFNFNGLTG TGVLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSS GGGSGGSGSGGSGGSGGSEKAAKAEAAARKMEELFKKHKIVAV LRANSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEK GAIIGAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVF YMPGVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFP NVKFVPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREK AKAFVEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 267)

RBDb270	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFHEVFNATTFASVYAWNRRKRISNCVADYSVIYNFAPFF AFKCYGVSPTKLNDLCFTNVYADSFVIRGNEVSQIAPGQTGNIA DYNKLPDDFTGCVIAWNSNKLDSPSGNYNYLYRLFRKSKLKP FERDISTEIQAGNKPCNGVAGSNCYSPLQSYGFRPTYGVGHQ PYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTG VLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSSGGG SGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRAN SVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIG AGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPG VMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPNVKFV PTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFV EKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 268)
RBDb272	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVQPTESIVRFP</u> NITNLCPFHEVFNATTFASVYAWNRRKRISNCVADFSVIYNFAPFF AFKCYGVSPTKLNDLCFTNIYADSFVIRGNEVSQIAPGQTGNIA YNYKLPDDFTGCVIAWNSNKLDSPSGNYNYLYRLFRKSKLKP ERDISTEIQAGNKPCNGVAGSNCYSPLQSYGFRPTYGVGHQP YRVVLSFETLSGPATVCGPKKSTNLVKNKCVNFNFNGLTGTG VLTESNKKFQPYQQFGRDSADTTDAVRDPQTLEILDITPCSSGG GSGGSGSGSGSGSGSGSEKAAKAEAAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTTELVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPPNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 269)

RBDb329	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVVP SGDVVRFP</u> NITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLNSTFF STFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVI ADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNKYRYLRHGKL RPFERDISNVPFSPDGKPC TPPALNCYWPLNDYGFYTTTGIGYQ PYRVVLSFELLNAPATVCGPKLSTD LIKNQCVNFNFNGLTGTG VLT PSSKRFQPFQQFGRDVSDFTDSVRDPKTSEILDISPCSSGG GSGGSGSGGSGGSGGSEKA AKAEEAARKMEELFKKHKIVAVLRA NSVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAI GAGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMP GVMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPF PNVKF VPTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAF VEKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 270)
RBDb331	<u>MGILPSPGMPALLSLVSLLSVLLMGCV AETGSRVVP SGDVVRFP</u> NITNLCPFGEVFNATKFPSVYAWERKKISNCVADFSVLNSTFFS TFKCYGVSATKLNDLCFSNIYADSFVVKGDDVRQIAPGQTGVIA DYNKLPDDFMGCVLAWNTRNIDATSTGNYNKYRYLRHGKLR PFERDISNVPFSPDGKPC TPPALNCYWPLNDYGFYTTTGIGYQP YRVVLSFETLSGPATVCGPKLSTD LIKNQCVNFNFNGLTGTGV LTPSSKRFQPYQQFGRDSSDFTDSVRDPKTSEILDISPCSSGGG SGGSGSGGSGGSGGSGSEKA AKAEEAARKMEELFKKHKIVAVLRAN SVEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIG AGTVTSVEQARKAVESGAEFIVSPHLDEEISQFAKEKGVFYMPG VMTPTELVKAMKLGHTILKLFPGEVVG PQFVKAMKGPF PNVKFV PTGGVNLDNVAEWFKAGVLAVGVGSALVKGTPDEVREKAKAFV EKIRGATE <u>LEHHHHHH</u> (SEQ ID NO: 271)

Referencias

1. Ahmed, S., et al., A Stabilized, Monomeric, Receptor Binding Domain Elicits High-Titer Neutralizing Antibodies Against All SARS-CoV-2 Variants of Concern, *Front. Immunol.*, 12:765211 (2021), disponible en <https://pubmed.ncbi.nlm.nih.gov/34956193/>.

2. Bale, J.B., et al., Accurate design of megadalton-scale two-component icosahedral protein complexes, *Science*, 353(6297):389-394 (2016), disponible en <https://pubmed.ncbi.nlm.nih.gov/27463675/>.
3. Cao, Y., et al., Characterization of the enhanced infectivity and antibody evasion of Omicron BA.2.75. *Cell Host Microbe.*, 30(11):1527-1539 (2022), disponible en <https://pubmed.ncbi.nlm.nih.gov/36270286/>.
4. Dalvie, N., et al., Engineered SARS-CoV-2 receptor binding domain improves manufacturability in yeast and immunogenicity in mice, *Proc. Natl., Acad. Sci.*, 118(38):e2106845118 (2021), disponible en <https://pubmed.ncbi.nlm.nih.gov/34493582/>.
5. Ellis, D., et al., Stabilization of the SARS-CoV-2 Spike glycoprotein Receptor-Binding Domain Using Deep Mutational Scanning and Structure-Based Design, *Front Immunol.*, 12:710263 (2021), disponible en <https://pubmed.ncbi.nlm.nih.gov/34267764/>.
6. Fougereous, C., et al., Capsid-like particles decorated with the SARS-CoV-2 receptor-binding domain elicit strong virus neutralization activity, *Nat. Commu.*, 12(1):324 (2021), disponible en <https://pubmed.ncbi.nlm.nih.gov/33436573/>.
7. Heinze, K., et al., Protein Nanocontainers from Nonviral Origin: Testing the Mechanics of Artificial and Natural Protein Cages by AFM, *J. Phys. Chem.*, 120(26):5945-5952 (2016), disponible en <https://pubmed.ncbi.nlm.nih.gov/27187612/>.
8. King, N., et al., Computational design of self-assembling protein nanomaterials with atomic level accuracy, *Science*, 336(6085):1171-1174 (2012), disponible en <https://pubmed.ncbi.nlm.nih.gov/22654060/>.
9. King, N., et al., Accurate design of co-assembling multi-component protein nanomaterials, *Nature*, 510(7503):103-108 (2014), disponible en <https://pubmed.ncbi.nlm.nih.gov/24870237/>.
10. Leman, J. K., et al., Macromolecular modeling and design in Rosetta: recent methods and frameworks, *Nat. Methods*, 17(7):665-680 (2020), disponible en <https://pubmed.ncbi.nlm.nih.gov/32483333/>.
11. Mirdita, M., et al., ColabFold: making protein folding accessible to all, *Nat. Methods*, 19(6):679-682 (2022), disponible en <https://pubmed.ncbi.nlm.nih.gov/35637307/>.
12. Pinto, D., et al., Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody, *Nature*, 583(7815):290-295 (2020), disponible en

<https://pubmed.ncbi.nlm.nih.gov/32422645/>.

13. Stalls, V., et al., Cryo-EM structures of SARS-CoV-2 Omicron BA.2 Spike glycoprotein, *Cell Rep.*, 39(13):111009 (2022), disponible en <https://pubmed.ncbi.nlm.nih.gov/35732171/>.

14. Walls, A.C., et al., Elicitation of Potent Neutralizing Antibody Responses by Designed Protein Nanoparticle Vaccines for SARS-CoV-2, *Cell*, 183(5):1367-1382 (2020), disponible en <https://pubmed.ncbi.nlm.nih.gov/33160446/>.

15. Wrapp, D., et al., Cryo-EM structure of the 2019-nCoV Spike glycoprotein in the prefusion conformation, *Science*, 367(6483):1260-1263 (2020), disponible en <https://pubmed.ncbi.nlm.nih.gov/32075877/>.

16. Yuan, M., et al., A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV, *Science*, 368(6491):630-633 (2020), disponible en <https://pubmed.ncbi.nlm.nih.gov/32245784/>.

MODALIDADES

15 La descripción proporciona además las siguientes modalidades enumeradas:

Modalidad 1. Un polipéptido recombinante, que comprende un fragmento antigénico de una glucoproteína Spike de coronavirus, dicho fragmento antigénico que comprende un dominio de unión al receptor (RBD) y, opcionalmente, un subdominio 1 de coronavirus (SD1).

20 Modalidad 2. El polipéptido de la modalidad 1, en donde el fragmento antigénico comprende RBD-SD1.

Modalidad 3. El polipéptido de la modalidad 1 o la modalidad 2, en donde el SD1 comprende al menos una sustitución de un residuo de aminoácido no polar expuesto en la superficie con un residuo de aminoácido polar.

25 Modalidad 4. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una sustitución de uno o más, dos o más, o tres aminoácidos en las posiciones 560, 562 o 569 con un residuo polar, en donde las sustituciones son relativas a una secuencia de referencia de acuerdo con la SEQ ID NO: 1.

30 Modalidad 5. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una sustitución de uno o más, dos o más, o tres sustituciones de aminoácidos L560Q, F562T, F562Y, o I569S, en relación con una secuencia de referencia de acuerdo con la SEQ ID NO: 1.

Modalidad 6. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una sustitución de uno o más, dos o más, o tres o más aminoácidos en las posiciones 544, 546, 560, 562, 564, 569, o 582 en relación con una secuencia de referencia de acuerdo con la SEQ ID NO: 1.

5 Modalidad 7. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones N544L, N544M, N544Q, L546V, L560Q, F562T, F562Y, Q564C, Q564L, Q564N, Q564W, I569S, o L582S en relación con una secuencia de referencia de acuerdo con la SEQ ID NO: 1.

10 Modalidad 8. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones 338, 358, 363, 365, 392, o 395 en relación con una secuencia de referencia de acuerdo con la SEQ ID NO: 1.

Modalidad 9. El polipéptido de una cualquiera de las modalidades anteriores, en donde
15 el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones F338L, I358F, A363L, Y365F, Y365M, Y365W, F392W, o V395I en relación con una secuencia de referencia de acuerdo con la SEQ ID NO: 1.

Modalidad 10. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende sustituciones de aminoácidos que comprenden F338L/Y365W;
20 F392W; Y365F/V395I; Y365F/ F392W /V395I; F338L/A363L/Y365M; y/o I358F/Y365F/V395I.

Modalidad 11. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones 329, 348, 350, 367, 375, 402, 407, 410, 418, 429, 433, 435, 452, 464, 510, 512, 514, 517, 518, 519, 520, 522, 527 o 528 con respecto a una secuencia de
25 referencia según la SEQ ID NO: 1.

Modalidad 12. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en las posiciones F329K, F329R, F329Y, A348P, V350L, V367F, F375Y, I402V, V407L, I410F, I418V, F429W, V433I, A435I, A435V, L452R, F464Y, P512Q, V510I, V512F,
30 V512I, S514T, L517A, L517D, L517S, L517T, L518N, L518Q, L518V, H519D, H519G, H519R, H519S, H519T, A520C, A520D, A520G, A520H, P521A, P521D, P521N, P521Q, P521S,

A522G, A522I, P527N, K528I, K528Q, K528T, L518G o L518S en relación con una secuencia de referencia según la SEQ ID NO: 1.

Modalidad 13. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 348, 402, 464, 514, 520, 526 y 527, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

Modalidad 14. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende las sustituciones de aminoácidos que comprenden A348P/Y365F/V395I/L560Q/F562Y/I569S; Y365F/V395I/I402V/L560Q/F562Y/I569S; Y365F/V395I/S514T/L560Q/F562Y/I569S; Y365F/V395I/I402V/F464Y/L560Q/F562Y/I569S; y/o Y365F/V395I/G526S/P527N/L560Q/F562Y/I569S.

Modalidad 15. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 329, 517, 519, 520, y 544, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

Modalidad 16. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende las sustituciones de aminoácidos que comprenden Y365F/V395I/L560Q/F562Y/I569S; Y365F/V395I/F329Y/L560Q/F562Y/I569S; Y365F/V395I/L517T/H519S/L560Q/F562Y/I569S; Y365F/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S; Y365F/V395I/L517T/H519S/N544L/L560Q/F562Y/I569S; y/o Y365F/V395I/F329Y/L517T/H519S/L560Q/F562Y/I569S.

Modalidad 17. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 392, 395, 560, 562 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 329, 517, 519, 520, y 544, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

Modalidad 18. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende las sustituciones de aminoácidos que comprenden

Y365F/F392W/V395I/L560Q/F562Y/I569S; Y365F/F392W/V395I/F329Y/L560Q/F562Y/I569S;
 Y365F/F392W/V395I/L517T/H519S/L560Q/F562Y/I569S;
 Y365F/F392W/V395I/L517T/H519S/A520G/L560Q/F562Y/I569S;
 Y365F/F392W/V395I/L517T/H519S/N544L/L560Q/F562Y/I569S; y/o

5 Y365F/F392W/V395I/F329Y/L517T/H519S/L560Q/F562Y/I569S.

Modalidad 19. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519, 520, 560, 562, 564 y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 521, 544 y 546, siendo las sustituciones de aminoácidos relativas a una
 10 secuencia de referencia según la SEQ ID NO: 1.

Modalidad 20. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende las sustituciones de aminoácidos que comprenden
 Y365F/V395I/L517T/H519G/A520C/Q564C/L560Q/F562Y/I569S; y/o
 Y365F/V395I/L517S/H519R/A520H/P521Q/N544L/L546V/Q564W/F562T/L560Q/F562Y/I569S

15 Modalidad 21. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519 y 520; y una o más, o dos o más sustituciones de aminoácidos en las posiciones 518 y 392, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la SEQ ID NO: 1.

20 Modalidad 22. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende las sustituciones de aminoácidos que comprenden
 Y365F/V395I/L517T/H519S/A520G; Y365F/V395I/L517T/L518G/H519D/A520G;
 Y365F/V395I/L517T/L518S/H519S/A520G; Y365F/F392W/V395I/L517T/H519S/A520G;
 Y365F/F392W/V395I/L517T/L518G/H519D/A520G; y/o
 25 Y365F/F392W/V395I/L517T/L518S/H519S/A520G.

Modalidad 23. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende sustituciones de aminoácidos en las posiciones 329, 365, 395, 517, 519, 520, 560, 562, y 569; y una o más, dos o más, o tres o más sustituciones de aminoácidos en las posiciones 392, 544, y 564, siendo las sustituciones de aminoácidos relativas a una
 30 secuencia de referencia según la SEQ ID NO: 1.

Modalidad 24. El polipéptido de una cualquiera de las modalidades anteriores, en donde

el polipéptido comprende las sustituciones de aminoácidos que comprenden
 Y365F/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S/F329Y;
 Y365F/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S/F329Y;
 Y365F/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S/F329Y;

- 5 Y365F/F392W/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S/F329Y;
 Y365F/F392W/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S/F329Y; y/o
 Y365F/F392W/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S/F329Y.

Modalidad 25. El polipéptido de una cualquiera de las modalidades anteriores, en donde
 el polipéptido comprende las sustituciones de aminoácidos que comprenden

- 10 Y365F/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S;
 Y365F/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S;
 Y365F/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S;
 Y365F/F392W/V395I/L517T/H519S/A520G/N544M/L560Q/F562Y/I569S;
 Y365F/F392W/V395I/L517T/H519S/A520G/N544Q/L560Q/F562Y/I569S; y/o
 15 Y365F/F392W/V395I/L517T/H519S/A520G/Q564N/L560Q/F562Y/I569S.

Modalidad 26. El polipéptido de una cualquiera de las modalidades anteriores, en donde
 el polipéptido comprende sustituciones de aminoácidos en las posiciones 365, 395, 517, 519,
 520, 521, 504, 560, 562, 569, y 329; y una o más sustituciones de aminoácidos en las posiciones
 392, siendo las sustituciones de aminoácidos relativas a una secuencia de referencia según la

- 20 SEQ ID NO: 1.

Modalidad 27. El polipéptido de una cualquiera de las modalidades anteriores, en donde
 el polipéptido comprende las sustituciones de aminoácidos que comprenden

- Y365F/V395I/L517T/H519G/P521A/A520C/Q564C/L560Q/F562Y/I569S/F329Y;
 Y365F/V395I/L517T/H519G/P521N/A520C/Q564C/L560Q/F562Y/I569S/F329Y;
 25 Y365F/F392W/V395I/L517T/H519G/P521A/A520C/Q564C/L560Q/F562Y/I569S/F329Y; y/o
 Y365F/F392W/V395I/L517T/H519G/P521N/A520C/Q564C/L560Q/F562Y/I569S/F329Y.

Modalidad 28. El polipéptido de una cualquiera de las modalidades anteriores, en donde
 el polipéptido comprende una pluralidad de sustituciones de aminoácidos en las posiciones 329,
 365, 395, 526, 527, 528, 560, 562 y 569.

- 30 Modalidad 29. El polipéptido de una cualquiera de las modalidades anteriores, en donde
 el polipéptido comprende las sustituciones de aminoácidos que comprenden

F329Y/Y365F/V395I/G526S/P527N/K528T/L560Q/F562Y/I569S,
F329Y/Y365F/V395I/G526S/P527N/K528I/L560Q/F562Y/I569S.

y/o

Modalidad 30. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una secuencia polipeptídica al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a una o más de las SEQ ID NO: 11-19.

Modalidad 31. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una secuencia polipeptídica idéntica a una cualquiera de las SEQ ID NO: 11-19.

Modalidad 32. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido se expresa a un nivel más alto que un polipéptido de referencia sin las sustituciones de aminoácidos, opcionalmente un polipéptido de referencia según las SEQ ID NO: 2 o 10.

Modalidad 33. El polipéptido de una cualquiera de las modalidades anteriores, en donde dicho polipéptido expresado es más estable térmicamente o más antigénico que las SEQ ID NO: 2 o 10.

Modalidad 34. El polipéptido de una cualquiera de las modalidades anteriores, en donde el polipéptido comprende una secuencia polipeptídica al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a las SEQ ID NO: 2-10.

Modalidad 35. Una nanoestructura de proteína autoensamblable que comprende un primer componente que comprende un polipéptido de coronavirus recombinante de acuerdo con una cualquiera de las modalidades 1 a 34; y opcionalmente un segundo componente que comprende una segunda proteína.

Modalidad 36. La nanoestructura de proteína de la modalidad 35, en donde el polipéptido de coronavirus recombinante comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a una o más de las SEQ ID NO: 11-19.

al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a las SEQ ID NO: 11-19.

Modalidad 37. La nanoestructura de proteína de la modalidad 35 o la modalidad 36, en donde el polipéptido de coronavirus recombinante comprende una sustitución de uno o más, dos o más, tres o más, cuatro o más aminoácidos en relación con una secuencia de referencia según la SEQ ID NO: 1.

Modalidad 38. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 37, en donde un componente polipeptídico de nanoestructura proteica y el polipéptido de coronavirus recombinante están acoplados de forma no covalente.

Modalidad 39. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 38, en donde un componente polipeptídico de nanoestructura proteica y el polipéptido de coronavirus recombinante están acoplados de forma covalente.

Modalidad 40. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 39, en donde la nanoestructura proteica comprende una proteína de fusión, la proteína de fusión que comprende el polipéptido de coronavirus recombinante, un enlazador, y un componente polipeptídico de nanoestructura proteica.

Modalidad 41. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 40, en donde la nanoestructura proteica comprende una proteína de fusión, la proteína de fusión que comprende, en orden de extremo N a C, el polipéptido de coronavirus recombinante, un enlazador, y el componente polipeptídico de nanoestructura proteica.

Modalidad 42. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 41, en donde el polipéptido de coronavirus recombinante comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a las SEQ ID NO: 11-19.

Modalidad 43. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 42, en donde la proteína de fusión comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al

menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a las SEQ ID NO: 11-19.

5 Modalidad 44. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 43, en donde el primer componente comprende I53-50A.

Modalidad 45. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 44, en donde la nanoestructura de proteína es una nanoestructura de proteína I53-50AB, o una nanoestructura de proteína I3-01/MI3.

10 Modalidad 46. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 45, en donde la nanoestructura de proteína es una nanoestructura de proteína ferritina; una nanoestructura de proteína encapsulina; una nanoestructura de proteína de cápside de bacteriófago CP3; una nanoestructura de proteína de cápside de bacteriófago Q β ; o una nanoestructura de proteína de cápside de bacteriófago AP205.

15 Modalidad 47. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 46, en donde el polipéptido recombinante comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a las SEQ ID NO: 11-19.

Modalidad 48. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 47, en donde el componente polipeptídico de nanoestructura proteica comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 25 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a las SEQ ID NO: 20-21.

Modalidad 49. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 30 48, en donde la nanoestructura de proteína comprende un segundo componente polipeptídico.

Modalidad 50. La nanoestructura de proteína de una cualquiera de las modalidades 35 a

49, en donde el segundo componente polipeptídico es I53-50B.

Modalidad 51. La nanoestructura de proteína de una cualquiera de las modalidades 35 a 50, en donde el segundo componente polipeptídico comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a las SEQ ID NO: 22.

Modalidad 52. Las nanoestructuras de proteína de una cualquiera de las modalidades 35 a 51, en donde el enlazador comprende entre 8 y 24 residuos de aminoácidos.

53. Las nanoestructuras de proteína de una cualquiera de las modalidades 35 a 52, en donde el enlazador se selecciona del grupo que consiste en GGSGGSGSGGSGGSGS; SGGGSGGSGSGGSGGSGS; EPEGGSGGSGSGGSGGSGS; YGGSGGSGGSGSGGSGGSGS; y GSGGSGGSGGSGGSGGSGGSGGSGS.

Modalidad 54. Las nanoestructuras de proteína de una cualquiera de las modalidades 35 a 53, en donde el polipéptido recombinante comprende un segmento polipeptídico que tiene una secuencia de polipéptidos al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, o al menos 99 % idéntica a la secuencia polipeptídica mostrada en la Tabla 27.

Modalidad 55. Un polinucleótido que codifica el polipéptido de una cualquiera de las modalidades 1 a 34 o la nanoestructura de proteína de una cualquiera de las modalidades 35 a 54.

Modalidad 56. El polinucleótido de la modalidad 55, en donde el polinucleótido es ARNm.

Modalidad 57. Una composición farmacéutica que comprende el polipéptido de una cualquiera de las modalidades 1 a 34, la nanoestructura de proteína de una cualquiera de las modalidades 35 a 54, o el polinucleótido de una cualquiera de la modalidad 55 a 56.

Modalidad 58. Una vacuna que comprende el polipéptido de una cualquiera de las modalidades 1 a 34, la nanoestructura de proteína de una cualquiera de las modalidades 35 a 54, o el polinucleótido de una cualquiera de la modalidad 55 a 56.

Modalidad 59. La vacuna de la modalidad 58, que comprende además un adyuvante.

Modalidad 60. La vacuna de la modalidad 58 o la modalidad 59, en donde el adyuvante comprende escualeno.

5 Modalidad 61. La vacuna de una cualquiera de las modalidades 58 a 60, en donde la vacuna es una vacuna bivalente.

Modalidad 62. La vacuna de una cualquiera de las modalidades 58 a 61, en donde la vacuna bivalente comprende el polipéptido de una cualquiera de las modalidades 1 a 34 o la nanoestructura de proteína de una cualquiera de las modalidades 35 a 54.

10 Modalidad 63. Un método para prevenir la enfermedad por coronavirus en un sujeto que lo necesita, comprende administrar al sujeto el polipéptido de una cualquiera de las modalidades 1 a 34, la nanoestructura proteica de una cualquiera de las modalidades 35 a 54, el polinucleótido de una cualquiera de las modalidades 55 a 56, el componente farmacéutico de una cualquiera de la modalidad 57, la vacuna de una cualquiera de las modalidades 58 a 62.

15 Modalidad 64. Un método para inmunizar a un sujeto que lo necesita contra una infección por coronavirus, el método que comprende administrar al sujeto la vacuna de una cualquiera de las modalidades 58 a 62.

20 Modalidad 65. Un kit que comprende el polipéptido de una cualquiera de las modalidades 1 a 34, la nanoestructura proteica de una cualquiera de las modalidades 35 a 54, el polinucleótido de una cualquiera de las modalidades 55 a 56, el componente farmacéutico de una cualquiera de la modalidad 57, la vacuna de una cualquiera de las modalidades 58 a 62, e instrucciones para el uso de los mismos.

25 Aunque la invención se ha descrito en relación con modalidades específicas propuestas de la misma, se entenderá que es capaz de modificaciones adicionales y esta solicitud pretende cubrir cualquier variación, usos o adaptaciones de la invención siguiendo, en general, los principios descritos en la presente, con las desviaciones que pueden aparecer dentro de la práctica conocida o habitual dentro de la técnica a la que pertenece la invención.