

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 926 585**

51 Int. Cl.:

**C12N 9/64**

(2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **31.01.2017 PCT/US2017/015879**

87 Fecha y número de publicación internacional: **10.08.2017 WO17136358**

96 Fecha de presentación y número de la solicitud europea: **31.01.2017 E 17705236 (2)**

97 Fecha y número de publicación de la concesión europea: **08.06.2022 EP 3411478**

54 Título: **Genes de Factor VIII optimizados**

30 Prioridad:

**01.02.2016 US 201662289696 P**  
**18.10.2016 US 201662409739 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**27.10.2022**

73 Titular/es:

**BIOVERATIV THERAPEUTICS INC. (100.0%)**  
**225 Second Avenue**  
**Waltham, MA 02451, US**

72 Inventor/es:

**TAN, SIYUAN y**  
**LIU, TONGYAO**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

ES 2 926 585 T3

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

## DESCRIPCIÓN

Genes de Factor VIII optimizados

## ANTECEDENTES DE LA DIVULGACIÓN

La vía de la coagulación de la sangre implica, en parte, la formación de un complejo enzimático de Factor VIIIa (FVIIIa) y Factor IXa (FIXa) (complejo Xasa) sobre la superficie de las plaquetas. El FIXa es una serina proteasa con actividad catalítica relativamente débil sin su cofactor FVIIIa. El complejo Xasa escinde el Factor X (FX) en el Factor Xa (FXa), el cual interactúa, a su vez, con el Factor Va (FVa) para escindir la protrombina y generar trombina. La hemofilia A es un trastorno hemorrágico provocado por mutaciones y/o deleciones en el gen de FVIII (FVIII) que da como resultado una deficiencia de la actividad del FVIII (Peyvandi *et al.* 2006). En algunos casos, los pacientes tienen niveles reducidos de FVIII debido a la presencia de inhibidores de FVIII tales como los anticuerpos anti-FVIII.

La hemofilia A se caracteriza por una hemorragia espontánea y un sangrado excesivo. Con el tiempo, el sangrado repetido en los músculos y las articulaciones, que a menudo comienza en la primera infancia, da como resultado una artropatía hemofílica y una lesión articular irreversible. Esta lesión es progresiva y puede conducir a una movilidad severamente limitada de las articulaciones, atrofia muscular y dolor crónico (Rodríguez-Merchan, E.C., *Semin. Thromb. Hemost.* 29:87-96 (2003), que se incorpora en esta memoria como referencia en su totalidad).

La enfermedad se puede tratar mediante una terapia de reemplazo que fija como objetivo la restauración de la actividad de FVIII al 1 a 5 % de los niveles normales para prevenir el sangrado espontáneo (véase, p. ej., Mannucci, P.M., *et al.*, *N. Engl. J. Med.* 344:1773-9 (2001), incorporado en esta memoria como referencia en su totalidad). Hay disponibles productos del FVIII recombinantes y derivados de plasma para tratar episodios hemorrágicos a demanda o para prevenir que se produzcan episodios hemorrágicos mediante el tratamiento profiláctico. Basado en la semivida de estos productos (10-12 h) (White G.C., *et al.*, *Thromb. Haemost.* 77:660-7 (1997); Morfini, M., *Haemophilia* 9 (supl. 1):94-99; discusión 100 (2003)), los regímenes de tratamiento requieren una administración intravenosa frecuente, comúnmente de dos a tres veces por semana para la profilaxis y de una a tres veces al día para el tratamiento a demanda (Manco-Johnson, M.J., *et al.*, *N. Engl. J. Med.* 357:535-544 (2007)), cada uno de los cuales se incorpora en esta memoria como referencia en su totalidad. Una administración tan frecuente es inconveniente y costosa.

Un impedimento importante para proporcionar una proteína FVIII recombinante de bajo costo a los pacientes es el alto costo de la producción comercial. La proteína FVIII se expresa deficientemente en sistemas de expresión heterólogos, dos o tres órdenes de magnitud más bajos que las proteínas de tamaño similar. (Lynch *et al.*, *Hum. Gene. Ther.*; 4:259-72 (1993). La expresión deficiente de FVIII se debe en parte a la presencia de elementos que actúan en cis en la secuencia codificante de FVIII que inhiben la expresión de FVIII, tales como elementos silenciadores de la transcripción (Hoebe *et al.*, *Blood* 85:2447-2454 (1995)), secuencias similares a la unión a la matriz (MARs, por sus siglas en inglés) (Falux *et al.*, *Mol. Cell. Biol.* 16:4264-4272 (1996)) y elementos inhibidores de la elongación de la transcripción (Koeberl *et al.*, *Hum. Gene. Ther.*; 6:469-479 (1995)).

Avances en la comprensión de los autores de la invención de la biología de la expresión de FVIII han conducido al desarrollo de variantes de FVIII más potentes. Por ejemplo, estudios bioquímicos demostraron que el dominio B del FVIII era prescindible para la actividad del cofactor del FVIII. La delección del dominio B dio como resultado un aumento de 17 veces en los niveles de ARNm frente al FVIII de tipo salvaje de longitud completa y un aumento del 30% en la proteína secretada. (Toole *et al.*, *Proc Natl Acad Sci USA* 83:5939-42 (1986)). Esto condujo al desarrollo del concentrado de proteína FVIII con dominio B eliminado (BDD, por sus siglas en inglés), que ahora se utiliza ampliamente en la clínica. Estudios recientes, sin embargo, indican que el hFVIII de longitud completa y BDD se pliegan incorrectamente en la luz del ER, dando como resultado la activación de la respuesta proteica desplegada (UPR, por sus siglas en inglés) y la apoptosis de los hepatocitos murinos.

Por lo tanto, existe una necesidad en la técnica de secuencias de FVIII que se expresen de forma eficaz en sistemas heterólogos.

Los documentos WO 2014/127215 y WO 2011/005968 describen secuencias de Factor VIII optimizadas por codones. El documento WO 2016/004113 describe secuencias de Factor IX optimizadas por codones.

## SUMARIO DE LA INVENCION

La presente invención se define en las reivindicaciones independientes y determinadas características opcionales de la misma se definen en las reivindicaciones dependientes. En la medida en que se utilicen en esta memoria los términos "invención", "ejemplo" y "realización", esto se interpretará de tal manera que la única protección buscada es para la invención tal como se reivindica.

La presente invención proporciona una molécula de ácido nucleico aislada como se define en las reivindicaciones adjuntas.

La presente invención proporciona, además, un vector que comprende la molécula de ácido nucleico como se define en las reivindicaciones adjuntas. La presente invención también proporciona una célula huésped que comprende la molécula de ácido nucleico o el vector como se define en las reivindicaciones adjuntas.

Además, la presente invención proporciona un método para producir un polipéptido con actividad de FVIII, que comprende: cultivar la célula huésped como se define en las reivindicaciones adjuntas en condiciones en las que se produce un polipéptido con actividad de FVIII y recuperar el polipéptido con actividad de FVIII.

Además, la presente invención proporciona una molécula de ácido nucleico o un vector como se define en las reivindicaciones adjuntas para uso en un método para tratar un trastorno hemorrágico.

La información técnica expuesta más adelante puede, en algunos aspectos, ir más allá del alcance de la invención, que se define exclusivamente por las reivindicaciones adjuntas. La información técnica adicional se proporciona para ubicar la invención real en un contexto técnico más amplio y para ilustrar posibles desarrollos técnicos relacionados.

Además, las referencias incidentales a los métodos para el tratamiento del cuerpo humano o animal mediante cirugía o terapia y los métodos de diagnóstico practicados en el cuerpo humano o animal no deben interpretarse como que reclaman protección para dichos métodos como tales, sino que deben interpretarse como una referencia a productos, en particular sustancias o composiciones, para uso en cualquiera de estos métodos.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

Las Figs. 1A-1J proporcionan secuencias de nucleótidos optimizadas por codones que codifican el Factor VIII con el dominio B suprimido (SEQ ID NO: 17). La FIG. 1A muestra la secuencia de nucleótidos de coFVIII-3 (SEQ ID NO: 1). La FIG. 1B muestra la secuencia de nucleótidos de coFVIII-4 (SEQ ID NO: 2). La FIG. 1C muestra la secuencia de nucleótidos de coFVIII-5 (SEQ ID NO: 70). La FIG. 1D muestra la secuencia de nucleótidos de coFVIII-6 (SEQ ID NO: 71). La FIG. 1E muestra la secuencia de nucleótidos de coFVIII-52 (SEQ ID NO: 3). La FIG. 1F muestra la secuencia de nucleótidos de coFVIII-62 (SEQ ID NO: 4). La FIG. 1G muestra la secuencia de nucleótidos de coFVIII-25 (SEQ ID NO: 5). La FIG. 1H muestra la secuencia de nucleótidos de coFVIII-26 (SEQ ID NO: 6). Las FIGs. 1I y 1J muestran las secuencias de nucleótidos y aminoácidos, respectivamente, suprimidas en el dominio B (BDD-FVIII) (SEQ ID NOs: 16 y 17, respectivamente).

Las FIGs. 2A-2J muestran ajustes del sesgo de uso de codones en las secuencias de nucleótidos optimizadas por codones que codifican BDD-FVIII. La FIG. 2A muestra la frecuencia relativa de codones en la secuencia de nucleótidos de tipo salvaje (antes de la optimización por codones) que codifica BDD-FVIII, p. ej., BDD-FVIII no optimizado. El índice de adaptación de codones humanos (CAI, por sus siglas en inglés) de la secuencia BDD-FVIII no optimizada es del 74 %. La FIG. 2B muestra la frecuencia relativa de los codones en la secuencia variante de coFVIII-1, que tiene un CAI humano del 88 %. La FIG. 2C muestra la frecuencia relativa de los codones en la secuencia variante de coFVIII-3, que tiene un CAI humano del 91 %. La FIG. 2D muestra la frecuencia relativa de los codones en la secuencia variante de coFVIII-4, que tiene un CAI humano del 97 %. La FIG. 2E muestra la frecuencia relativa de los codones en la secuencia variante de coFVIII-5, que tiene un CAI humano del 83 %. La FIG. 2F muestra la frecuencia relativa de los codones en la secuencia variante de coFVIII-6, que tiene un CAI humano del 83 %. La FIG. 2G muestra la frecuencia relativa de codones en la secuencia variante de coFVIII-52, que tiene un CAI humano del 91 %. La FIG. 2H muestra la frecuencia relativa de codones en la secuencia variante de coFVIII-62, que tiene un CAI humano del 91 %. La FIG. 2I muestra la frecuencia relativa de codones en la secuencia variante de coFVIII-25, que tiene un CAI humano del 88 %. La FIG. 2J muestra la frecuencia relativa de codones en la secuencia variante de coFVIII-26, que tiene un CAI humano del 88 %.

La FIG. 3 proporciona un mapa de plásmido de FVIII-303, que comprende coFVIII-1 en una cadena principal de pcDNA3 bajo el control del promotor de transtiretina potenciado por ET, que está situado aguas arriba del sitio de inicio de la traducción de coFVIII-1 y que comprende un potenciador sintético, un potenciador de mTIR y un promotor de mTIR.

La FIG. 4 muestra una representación gráfica de la actividad plasmática de FVIII en ratones HemA después de la inyección hidrodinámica de 5 µg de FVIII-303 (coFVIII-1; círculos) o 5 µg de FVIII-311 (BDD-FVIII; cuadrados). La actividad plasmática de FVIII se determinó mediante un ensayo cromogénico específico para FVIII a las 24, 48 y 72 horas después de la inyección. Se muestran los niveles de actividad relativos a las 72 horas, normalizados al nivel de expresión de FVIII-311.

La FIG. 5 muestra un mapa de plásmido de pLV-coFVIII-52, que comprende coFVIII-52 en un plásmido lentiviral bajo el control de un promotor ET, que está situado aguas arriba del sitio de inicio de la traducción de coFVIII-52 y que comprende un potenciador sintético, un potenciador mTTR y un promotor mTTR.

Las FIGs. 6A-6C muestran representaciones gráficas de la actividad plasmática de FVIII en ratones HemA después de la inyección hidrodinámica de diversos nucleótidos que codifican FVIII. La actividad plasmática de FVIII se determinó mediante un ensayo cromogénico específico para FVIII a las 24, 48 y 72 horas después de la inyección. La FIG. 6A muestra la actividad plasmática de FVIII en ratones HemA después de la inyección hidrodinámica de 5 µg de LV-coFVIII-1 (círculos negros), 5 µg de LV-coFVIII-3 (triángulos), 5 µg de LV-coFVIII-4 (triángulos invertidos), 5 µg de LV-coFVIII-5 (rombos) o 5 µg de LV-coFVIII-6 (círculos en blanco). La FIG. 6B

muestra la actividad plasmática de FVIII en ratones HemA tras la inyección hidrodinámica de 5 µg de LV-coFVIII-1 (círculos), 5 µg de LV-coFVIII-25 (triángulos) o 5 µg de LV-coFVIII-26 (triángulos invertidos). La FIG. 6C muestra la actividad plasmática de FVIII en ratones HemA después de la inyección hidrodinámica de 20 µg de LV-2116 (secuencia de nucleótidos BDD-FVIII no optimizada por codones (WT); círculos en blanco), 20 µg de LV-coFVIII-1 (triángulos), 20 µg de LV-coFVIII-52 (cuadrados) o 20 µg de LV-coFVIII-62 (círculos negros). Se muestran los niveles de actividad relativa a las 72 horas para cada uno de los plásmidos, normalizados a los niveles de expresión de LV-coFVIII-1 (FIGs. 6A, 6B y 6C) y/o LV-2116 (FIG. 6C), como se indica.

La FIG. 7 muestra la actividad de FVIII en plasma en ratones HemA 24 días después de la inyección con 1E8 TU/vector lentiviral de ratón que comprende coFVIII-1, coFVIII-5, coFVIII-52, coFVIII-6 o coFVIII-62 en comparación con el control LV-2116 (BDD-FVIII), y según lo medido por un ensayo cromogénico específico para FVIII. Las barras de errores indican las desviaciones estándares.

Las FIGs. 8A-8C proporcionan las diversas secuencias de nucleótidos optimizadas por codones que codifican BDD-FVIII condensado con un XTEN. La FIG. 8A muestra la secuencia de nucleótidos de coFVIII-52-XTEN (SEQ ID NO: 19), en donde una secuencia de nucleótidos que codifica un XTEN que tiene 144 aminoácidos ("XTEN<sub>144</sub>"; SEQ ID NO: 18; subrayado) se inserta dentro de la secuencia de nucleótidos de coFVIII-52. La FIG. 8B muestra la secuencia de nucleótidos de coFVIII-1-XTEN (SEQ ID NO: 20), en la que una secuencia de nucleótidos que codifica un XTEN que tiene 144 aminoácidos ("XTEN<sub>144</sub>"; SEQ ID NO: 18; subrayado) se inserta dentro de la secuencia de nucleótidos de coFVIII-1. La FIG. 8C muestra la secuencia de nucleótidos de coFVIII-6-XTEN (SEQ ID NO: 72), en la que una secuencia de nucleótidos que codifica un XTEN que tiene 144 aminoácidos ("XTEN<sub>144</sub>"; SEQ ID NO: 18; subrayado) se inserta dentro de la secuencia de nucleótidos de coFVIII-6 (p. ej., el residuo de aminoácido 745 correspondiente a la secuencia de FVIII maduro).

La FIG. 9 proporciona un mapa de plásmido de pLV-coFVIII-52-XTEN, que comprende coFVIII-52-XTEN en un vector lentiviral bajo el control del promotor ET. Los vectores lentivirales que comprenden cada una de las moléculas de ácido nucleico optimizadas por codones restantes que codifican un polipéptido con actividad de FVIII, como se describe en esta memoria, se construyeron de la misma manera que pLV-coFVIII-52-XTEN, en que se insertó la misma secuencia de XTEN para reemplazar el dominio B de FVIII.

Las FIGs. 10A y 10B muestran la actividad de FVIII en ratones HemA después de la inyección con ADN plasmídico (FIG. 10A) o vector lentiviral (FIG. 10B) que comprende las diversas secuencias de nucleótidos optimizadas por codones que codifican BDD-FVIII. La FIG. 10A muestra una representación gráfica de la actividad plasmática de FVIII en ratones HemA después de la inyección hidrodinámica con 5 µg de FVIII-311 (secuencia de nucleótidos que codifica BDD-FVIII no optimizadas por codones; cuadrados), 5 µg de FVIII-303 (coFVIII-1; círculos pequeños) o FVIII-306 (coFVIII-1-XTEN<sub>144</sub>; círculos grandes). La actividad relativa a las 72 horas, normalizada a FVIII-311, se muestra para cada uno de los plásmidos. La FIG. 10B muestra la actividad de FVIII en plasma en ratones HemA 21 días después de la inyección con 1E8 TU/ratón de vector lentiviral que comprende coFVIII-52 o coFVIII-52-XTEN en comparación con el control LV-2116 (BDD-FVIII), y según lo medido por un ensayo cromogénico específico para FVIII. Las barras de error indican las desviaciones estándares.

La FIG. 11A muestra la secuencia de aminoácidos del factor VIII humano maduro de longitud completa. La FIG. 11B muestra la secuencia de aminoácidos del factor de von Willebrand humano de longitud completa (SEQ ID NO: 44). Las FIGs. 11C y 11D muestran las secuencias de aminoácidos y nucleótidos, respectivamente, de un polipéptido XTEN que tiene 42 aminoácidos (XTEN AE42-4; SEQ ID NOs: 46 y 47, respectivamente). Las secuencias de aminoácidos de diversos polipéptidos XTEN que tienen 144 aminoácidos se muestran en las FIGs. 11E, 11G, 11I, 11K, 11M, 11O, 11Q, 11S, 11U y 11W (SEQ ID NOs: 48, 50, 52, 54, 56, 58, 60, 62, 64 y 66, respectivamente), y las secuencias de nucleótidos correspondientes se muestran en las FIGs. 11F, 11H, 11J, 11L, 11N, 11P, 11R, 11T, 11V y 11X (SEQ ID Nos: 49, 51, 53, 55, 57, 59, 61, 63, 65 y 67, respectivamente). La FIG. 11Y muestra la secuencia de nucleótidos de un promotor ET (SEQ ID NO: 69). La FIG. 11Z muestra la secuencia de nucleótidos para coFVIII-1 (SEQ ID NO: 68).

La FIG. 12A es una representación gráfica de la actividad plasmática de FVIII (UI/mL) en ratones HemA de 14 días de edad después de la administración IV de aproximadamente 1,5 E10 TU/kg de LV-wtBDD-FVIII (círculos), LV-coFVIII-6 (cuadrados) o LV-coFVIII-6XTEN (triángulos). La FIG. 12B es una representación gráfica del número de copias del vector (VCN) 150 días después del tratamiento de ratones HemA de 14 días de edad a los que se administraron por vía IV aproximadamente 1,5 E10 TU/kg de vectores lentivirales que expresan wtBDD-FVIII, coFVIII-1, coFVIII-3, coFVIII-4, coFVIII-5, coFVIII-6, coFVIII-52, coFVIII-62, coFVIII-25 o coFVIII-26. La FIG. 12C es una representación gráfica de la actividad plasmática de FVIII (UI/mL) 21 días después del tratamiento de ratones HemA de 14 días de edad a los que se administraron por vía IV aproximadamente 1,5 E10 TU/kg de vectores lentivirales que expresan wtBDD-FVIII, coFVIII-1, coFVIII-3, coFVIII-4, coFVIII-5, coFVIII-6, coFVIII-52, coFVIII-62, coFVIII-25 o coFVIII-26.

Las FIGs. 13A y 13B son representaciones gráficas que ilustran los niveles de actividad plasmática de FVIII (FIG. 13A) y los niveles de anticuerpos anti-FVIII (FIG. 13B) en cinco ratones HemA tratados con un lentivirus que expresa la variante coFVIII-5. A compañeros de camada HemA de catorce días de edad se les administraron aproximadamente 1,5 E10 TU/kg de un lentivirus que expresa la variante coFVIII-5 mediante inyección intravenosa. Cada uno de los ratones se designa con un número (*es decir*, 1, 2, 3, 4 y 5; FIGs 13A y 13B).

La FIG. 14 es una representación gráfica de la correlación entre el nivel de expresión de LV-FVIII, como se evidencia por la actividad plasmática de FVIII, 21 días después del tratamiento lentiviral, y la presencia de anticuerpos anti-FVIII. Cada uno de los puntos de datos corresponde a un solo ratón HemA. Cada uno de los ratones recibió una dosis de 1,5 E10 TU/kg mediante inyección intravenosa de un lentivirus que expresaba una de las variantes de coFVIII descritas en esta memoria. Las líneas horizontales indican la actividad plasmática



media de FVIII.

La FIG. 15 es una representación gráfica de la correlación entre el número de copias del vector (VCN, por sus siglas en inglés) por célula 150 días después del tratamiento lentiviral y la presencia de anticuerpos anti-FVIII. Cada uno de los puntos de datos corresponde a un solo ratón HemA. Cada uno de los ratones recibió una dosis

de  $1,5 \times 10^9$  TU/kg mediante inyección intravenosa de un lentivirus que expresaba una de las variantes de coFVIII descritas en esta memoria. Las líneas horizontales indican el VCN medio.

Las FIGs. 16A y 16B son representaciones gráficas que ilustran los niveles de actividad plasmática de FVIII (FIG. 16A) y los niveles de anticuerpos anti-FVIII (FIG. 16B) en dos ratones HemA (coFVIII-52-A y coFVIII-52-B) tratados con un lentivirus que expresa la variante coFVIII-52. A compañeros de camada HemA de catorce días de edad se les administraron aproximadamente  $1,5 \times 10^9$  TU/kg de un lentivirus que expresa la variante coFVIII-52 mediante inyección intravenosa. Las FIGs. 16C y 16D son imágenes que muestran tinción de hibridación in situ de ARN para la expresión de FVIII (tinción oscura) en tejido hepático recogido de ratones coFVIII-52-A (FIG. 16C) y coFVIII-52-B (FIG. 16D) de las FIGs. 16A y 16B.

La FIG. 17 es una representación gráfica que muestra la expresión de FVIII a largo plazo en ratones neonatos HemA tratados con un lentivirus que expresa una variante de FVIII con dominio B de tipo salvaje eliminado (wtBDD-FVIII; triángulos), coFVIII-52XTEN (círculos) o coFVIII-6XTEN (triángulo invertido). Se administraron a ratones HemA neonatales mediante inyección intravenosa aproximadamente  $1,5 \times 10^9$  TU/kg de un lentivirus que expresaba wtBDD-FVIII, coFVIII-52XTEN o coFVIII-6XTEN. La actividad plasmática de FVIII se midió durante aproximadamente 16 semanas.

La FIG. 18 es una representación gráfica del nivel de FVIII circulante en neonatos de perro HemA (S3 o K4) después de la administración de  $1,3 \times 10^9$  unidades de transducción/kg de vector lentiviral que comprende un nucleótido que codifica Factor VIII condensado con XTEN (SEQ ID NO: 72; LV- coFVIII-6-XTEN). Los cuadrados conectados por una línea continua representan muestras de aPTT-S3 y los triángulos conectados por una línea discontinua representan muestras de aPTT-K4. El eje y muestra la actividad de FVIII en plasma como el porcentaje de lo normal, en donde la actividad de FVIII humano normal es del 100 %. El eje x muestra los días posteriores al tratamiento con lentivirus, en donde el tratamiento con lentivirus se administra el día 0.

Las FIGs. 19A-19C son representaciones gráficas de la hemostasia de sangre entera según se monitoriza por el ensayo de tromboelastometría rotacional (ROTEM, por sus siglas en inglés) para un perro HemA sin tratamiento previo (FIG. 19A), el perro S3 a las 2 semanas después del tratamiento con lentivirus (FIG. 19B) y el perro K4 a las 2 semanas posteriores al tratamiento con lentivirus (FIG. 19C). El tiempo de coagulación (CT, por sus siglas en inglés) se muestra en segundos (s), el tiempo de formación de coágulos (CFT, por sus siglas en inglés) se muestra en segundos (s), el ángulo alfa ( $\alpha$ ) se muestra en grados ( $^\circ$ ), la amplitud 5 minutos después de la CT (A5) se muestra como milímetros (mm), la amplitud 20 minutos después de la CT (A20) se muestra como milímetros (mm) y la firmeza máxima del coágulo (MCF, por sus siglas en inglés) se muestra como milímetros (mm) para cada una de las FIG. 19A-19C. La FIG. 19D es una tabla que resume el intervalo normal para cada uno de los parámetros CT, CFT,  $\alpha$ , A5, A2 y MCF mostrados en las FIGs. 19A-19C.

## DESCRIPCIÓN DETALLADA DE LA DIVULGACIÓN

La presente divulgación describe genes optimizados por codones que codifican polipéptidos con actividad de Factor VIII (FVIII). La presente divulgación está dirigida a moléculas de ácido nucleico optimizadas por codones que codifican polipéptidos con actividad de Factor VIII, vectores y células huésped que comprenden moléculas de ácidos nucleicos optimizadas, polipéptidos codificados por moléculas de ácidos nucleicos optimizadas y métodos para producir polipéptidos de este tipo. La presente divulgación también está dirigida a una secuencia de ácido nucleico del Factor VIII optimizada, a un vector que comprende la secuencia de ácido nucleico optimizada o el polipéptido codificado por el mismo para uso en un método de tratamiento de trastornos hemorrágicos tales como la hemofilia. La presente divulgación satisface una necesidad importante en la técnica al proporcionar secuencias de Factor VIII optimizadas que demuestran una expresión incrementada en las células huésped, mejoran el rendimiento de la proteína Factor VIII en métodos para producir Factor VIII recombinante y, potencialmente, dan como resultado una mayor eficacia terapéutica cuando se utilizan en métodos de terapia génica. En determinadas realizaciones, la divulgación describe una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que tiene una secuencia homóloga a una secuencia de nucleótidos de SEQ ID NO: 71.

Construcciones ejemplares de la divulgación se ilustran en las Figuras adjuntas y en el listado de secuencias. Con el fin de proporcionar una comprensión clara de la memoria descriptiva y las reivindicaciones, a continuación se proporcionan las siguientes definiciones.

### I. Definiciones

Ha de señalarse que el término una entidad "un" o "una" se refiere a una o más de esa entidad: por ejemplo, se entiende que "una secuencia de nucleótidos" representa una o más secuencias de nucleótidos. Como tal, el término "un" (o "una") y las expresiones "uno/una o más" y "al menos uno/una" se pueden utilizar indistintamente en esta memoria.

El término "aproximadamente" se utiliza en esta memoria para dar a entender aproximadamente, prácticamente, alrededor o en las regiones de. Cuando el término "aproximadamente" se utiliza junto con un intervalo numérico, modifica ese

intervalo al extender los límites por encima y por debajo de los valores numéricos establecidos. En general, el término "aproximadamente" se utiliza en esta memoria para modificar un valor numérico por encima y por debajo del valor establecido en una varianza del 10 por ciento, hacia arriba o hacia abajo (mayor o menor).

El término "aislado" para los fines de la presente divulgación designa un material biológico (célula, polipéptido, polinucleótido o un fragmento, variante o derivado del mismo) que ha sido eliminado de su entorno original (el entorno en el que está presente de forma natural). Por ejemplo, un polinucleótido presente en el estado natural en una planta o un animal no está aislado; sin embargo, el mismo polinucleótido separado de los ácidos nucleicos adyacentes en los que está presente de forma natural se considera "aislado". No se requiere nivel particular de purificación alguno. Polipéptidos producidos de forma recombinante y proteínas expresadas en células huésped se consideran aislados para los fines de la divulgación, al igual que los polipéptidos nativos o recombinantes que se han separado, fraccionado o purificado parcial o sustancialmente mediante cualquier técnica adecuada.

"Ácidos nucleicos", "moléculas de ácido nucleico", "oligonucleótido" y "polinucleótido" se utilizan indistintamente y se refieren a la forma polimérica de éster de fosfato de los ribonucleósidos (adenosina, guanosina, uridina o citidina; "moléculas de ARN") o desoxirribonucleósidos (desoxiadenosina, desoxiguanosina, desoxitimidina o desoxicitidina; "moléculas de ADN"), o cualquier análogo de fosfoéster de los mismos, tales como fosforotioatos y tioésteres, ya sea en forma de cadena sencilla o de hélice de doble cadena. Son posibles hélices de doble cadena de ADN-ADN, ADN-ARN y ARN-ARN. La expresión molécula de ácido nucleico y, en particular, molécula de ADN o ARN, se refiere únicamente a la estructura primaria y secundaria de la molécula, y no la limita a forma terciaria particular alguna. Así, esta expresión incluye el ADN de doble cadena que se encuentra, *entre otros*, en moléculas de ADN lineales o circulares (p. ej., fragmentos de restricción), plásmidos, ADN superenrollado y cromosomas. Al discutir la estructura de moléculas particulares de ADN de doble cadena, las secuencias se pueden describir en esta memoria de acuerdo con la convención normal de dar solo la secuencia en la dirección 5' a 3' a lo largo de la cadena de ADN no transcrita (*es decir*, la cadena que tiene una secuencia homóloga al ARNm). Una "molécula de ADN recombinante" es una molécula de ADN que ha sufrido una manipulación biológica molecular. El ADN incluye, pero no se limita a ADNc, ADN genómico, ADN plasmídico, ADN sintético y ADN semisintético. Una "composición de ácido nucleico" de la divulgación comprende uno o más ácidos nucleicos como se describe en esta memoria.

Tal como se utiliza en esta memoria, una "región codificante" o "secuencia codificante" es una porción de polinucleótido que consiste en codones traducibles en aminoácidos. Aunque un "codón de parada" (TAG, TGA o TAA) no se traduce típicamente en un aminoácido, puede considerarse parte de una región codificante, pero cualquier secuencia flanqueante, por ejemplo, promotores, sitios de unión a ribosomas, terminadores de la transcripción, intrones y similares, no son parte de una región codificante. Los límites de una región codificante están típicamente determinados por un codón de inicio en el extremo 5', que codifica el extremo amino del polipéptido resultante, y un codón de parada de la traducción en el extremo 3', que codifica el extremo carboxilo del polipéptido resultante. Dos o más regiones codificantes pueden estar presentes en una sola construcción de polinucleótidos, p. ej., en un solo vector, o en construcciones de polinucleótidos separadas, p. ej., en vectores separados (diferentes). Se cumple entonces que un solo vector puede contener solo una sola región codificante, o comprender dos o más regiones codificantes.

Determinadas proteínas secretadas por células de mamífero están asociadas con un péptido señal secretor que se escinde de la proteína madura una vez que se ha iniciado la exportación de la cadena proteica en crecimiento a través del retículo endoplásmico rugoso. Los expertos ordinarios en la técnica son conscientes que los péptidos señal están generalmente condensados con el extremo N del polipéptido y se escinden del polipéptido completo o de "longitud completa" para producir una forma secretada o "madura" del polipéptido. En determinadas realizaciones, un péptido señal nativo o un derivado funcional de esa secuencia que retiene la capacidad de dirigir la secreción del polipéptido está asociado operativamente con él. Alternativamente, puede utilizarse un péptido señal de mamífero heterólogo, p. ej., un activador del plasminógeno tisular humano (TPA, por sus siglas en inglés) o un péptido señal de  $\beta$ -glucuronidasa de ratón o un derivado funcional del mismo.

La expresión "aguas abajo" se refiere a una secuencia de nucleótidos que está situada en 3' con respecto a una secuencia de nucleótidos de referencia. En determinadas realizaciones, las secuencias de nucleótidos aguas abajo se refieren a secuencias que siguen al punto de inicio de la transcripción. Por ejemplo, el codón de inicio de la traducción de un gen está situado aguas abajo del sitio de inicio de la transcripción.

La expresión "aguas arriba" se refiere a una secuencia de nucleótidos que está situada en 5' con respecto a una secuencia de nucleótidos de referencia. En determinadas realizaciones, las secuencias de nucleótidos aguas arriba se refieren a secuencias que están situadas en el lado 5' de una región codificante o punto de partida de la transcripción. Por ejemplo, la mayoría de los promotores están situados aguas arriba del sitio de inicio de la transcripción.

Tal como se utiliza en esta memoria, la expresión "región reguladora de genes" o "región reguladora" se refiere a secuencias de nucleótidos situadas aguas arriba (secuencias 5' no codificantes), dentro o aguas abajo (secuencias 3' no codificantes) de una región codificante, y que influyen en la transcripción, el procesamiento del ARN, la estabilidad o la traducción de la región codificante asociada. Las regiones reguladoras pueden incluir promotores, secuencias conductoras de la traducción, intrones, secuencias de reconocimiento de poliadenilación, sitios de procesamiento de ARN, sitios de

unión de efectores y estructuras de bucle de tallo. Si se pretende que una región codificante se exprese en una célula eucariota, una señal de poliadenilación y una secuencia de terminación de la transcripción normalmente se localizarán en 3' con respecto a la secuencia codificante.

Un polinucleótido que codifica un producto génico, *p. ej.*, un polipéptido, puede incluir un promotor y/u otros elementos de control de la expresión (*p. ej.*, transcripción o traducción) asociados operativamente con una o más regiones codificantes. En una asociación operable, una región codificante para un producto génico, *p. ej.*, un polipéptido, se asocia con una o más regiones reguladoras de manera que coloque la expresión del producto génico bajo la influencia o el control de la o las regiones reguladoras. Por ejemplo, una región codificante y un promotor están "asociados operativamente" si la inducción de la función del promotor da como resultado la transcripción del ARNm que codifica el producto génico codificado por la región codificante, y si la naturaleza del enlace entre el promotor y la región codificante no interfiere con la capacidad del promotor de dirigir la expresión del producto génico o no interfiere con la capacidad de transcripción del molde de ADN. Otros elementos de control de la expresión, además de un promotor, por ejemplo potenciadores, operadores, represores y señales de terminación de la transcripción, también pueden asociarse operativamente con una región codificante para dirigir la expresión del producto génico.

"Secuencias de control de la transcripción" se refieren a secuencias reguladoras de ADN, tales como promotores, potenciadores, terminadores y similares, que proporcionan la expresión de una secuencia codificante en una célula huésped. Los expertos en la técnica conocen una variedad de regiones de control de la transcripción. Éstas incluyen, sin limitación, regiones de control de la transcripción que funcionan en células de vertebrados, tales como, pero no limitadas a segmentos promotores y potenciadores de citomegalovirus (el promotor temprano inmediato, junto con el intrón-A), virus simio 40 (el promotor temprano) y retrovirus (tal como el virus del sarcoma de Rous). Otras regiones de control de la transcripción incluyen las derivadas de genes de vertebrados tales como actina, proteína de choque térmico, hormona de crecimiento bovina y  $\beta$ -globina de conejo, así como otras secuencias capaces de controlar la expresión génica en células eucariotas. Regiones adecuadas de control de la transcripción adicionales incluyen promotores y potenciadores específicos para el tejido, así como promotores inducibles por linfoquinas (*p. ej.*, promotores inducibles por interferones o interleuquinas).

De manera similar, el experto ordinario en la técnica conoce una diversidad de elementos de control de la traducción. Estos incluyen, pero no se limitan a sitios de unión a ribosomas, codones de iniciación y terminación de la traducción y elementos derivados de picornavirus (particularmente un sitio de entrada al ribosoma interno, o IRES, al que también se alude como secuencia CITE).

El término "expresión", tal como se utiliza en esta memoria, se refiere a un proceso mediante el cual un polinucleótido produce un producto génico, por ejemplo, un ARN o un polipéptido. Incluye, sin limitación, la transcripción del polinucleótido en ARN mensajero (ARNm), ARN de transferencia (ARNt), ARN en horquilla pequeño (ARNhp), ARN pequeño de interferencia (ARNip) o cualquier otro producto de ARN, y la traducción de un ARNm en un polipéptido. La expresión produce un "producto genético". Tal como se utiliza en esta memoria, un producto génico puede ser un ácido nucleico, *p. ej.*, un ARN mensajero producido por la transcripción de un gen, o un polipéptido que se traduce a partir de un transcrito. Productos génicos descritos en esta memoria incluyen, además, ácidos nucleicos con modificaciones postranscripcionales, *p. ej.*, poliadenilación o corte y empalme, o polipéptidos con modificaciones postraduccionales, *p. ej.*, metilación, glicosilación, la adición de lípidos, asociación con otras subunidades proteicas o escisión proteolítica. El término "rendimiento", tal como se utiliza en esta memoria, se refiere a la cantidad de un polipéptido producido por la expresión de un gen.

Un "vector" se refiere a cualquier vehículo para la clonación y/o transferencia de un ácido nucleico a una célula huésped. Un vector puede ser un replicón al que se puede unir otro segmento de ácido nucleico para provocar la replicación del segmento fijado. Un "replicón" se refiere a cualquier elemento genético (*p. ej.*, plásmido, fago, cósmido, cromosoma, virus) que funciona como una unidad autónoma de replicación *in vivo*, *es decir*, es capaz de replicarse bajo su propio control. El término "vector" incluye vehículos tanto virales como no virales para introducir el ácido nucleico en una célula *in vitro*, *ex vivo* o *in vivo*. Se conocen y utilizan en la técnica un gran número de vectores que incluyen, por ejemplo, plásmidos, virus eucarióticos modificados o virus bacterianos modificados. La inserción de un polinucleótido en un vector adecuado se puede lograr ligando los fragmentos de polinucleótido apropiados en un vector elegido que tenga extremos cohesivos complementarios.

Los vectores se pueden diseñar para codificar marcadores seleccionables o indicadores que proporcionan la selección o identificación de células que han incorporado el vector. La expresión de marcadores o indicadores seleccionables permite la identificación y/o selección de células huésped que incorporan y expresan otras regiones codificantes contenidas en el vector. Ejemplos de genes marcadores seleccionables conocidos y utilizados en la técnica incluyen: genes que proporcionan resistencia a ampicilina, estreptomina, gentamicina, kanamicina, higromicina, herbicida bialafos, sulfonamida y similares; y genes que se utilizan como marcadores fenotípicos, *es decir*, genes reguladores de antocianina, gen isopentanol transferasa, y similares. Ejemplos de indicadores conocidos y utilizados en la técnica incluyen: luciferasa (Luc), proteína fluorescente verde (GFP, por sus siglas en inglés), cloranfenicol acetiltransferasa (CAT),  $\beta$ -galactosidasa (LacZ),  $\beta$ -glucuronidasa (Gus) y similares. Marcadores seleccionables también pueden considerarse indicadores.

- La expresión "marcador seleccionable" se refiere a un factor de identificación, habitualmente un gen de resistencia a un antibiótico o a un producto químico, que sea capaz de seleccionarse en función del efecto del gen marcador, *es decir*, resistencia a un antibiótico, resistencia a un herbicida, marcadores colorimétricos, enzimas, marcadores fluorescentes y similares, en donde el efecto se utiliza para rastrear la herencia de un ácido nucleico de interés y/o para identificar una célula u organismo que ha heredado el ácido nucleico de interés. Ejemplos de genes marcadores seleccionables conocidos y utilizados en la técnica incluyen: genes que proporcionan resistencia a ampicilina, estreptomycin, gentamicina, kanamicina, higromicina, herbicida bialafos, sulfonamida y similares; y genes que se utilizan como marcadores fenotípicos, *es decir*, genes reguladores de antocianina, gen isopentanol transferasa, y similares.
- La expresión "gen informador" se refiere a un ácido nucleico que codifica un factor de identificación que sea capaz de identificarse en función del efecto del gen informador, en donde el efecto se utiliza para rastrear la herencia de un ácido nucleico de interés, para identificar una célula o un organismo que ha heredado el ácido nucleico de interés, y/o para medir la inducción o transcripción de la expresión génica. Ejemplos de genes indicadores conocidos y utilizados en la técnica incluyen: luciferasa (Luc), proteína verde fluorescente (GFP), cloranfenicol acetiltransferasa (CAT),  $\beta$ -galactosidasa (LacZ),  $\beta$ -glucuronidasa (Gus) y similares. Genes marcadores seleccionables también pueden considerarse genes indicadores.
- "Promotor" y "secuencia promotora" se utilizan indistintamente y se refieren a una secuencia de ADN capaz de controlar la expresión de una secuencia codificante o ARN funcional. En general, una secuencia codificante está situada en 3' con respecto a una secuencia promotora. Los promotores pueden derivarse en su totalidad de un gen nativo, o pueden estar compuestos por diferentes elementos derivados de diferentes promotores que se encuentran en la naturaleza, o incluso pueden comprender segmentos de ADN sintético. Los expertos en la técnica entienden que diferentes promotores pueden dirigir la expresión de un gen en diferentes tejidos o tipos de células, o en diferentes fases de desarrollo, o en respuesta a diferentes condiciones ambientales o fisiológicas. A promotores que hacen que un gen se exprese en la mayoría de los tipos de células la mayoría de las veces se les alude comúnmente como "promotores constitutivos". A los promotores que hacen que un gen se exprese en un tipo de célula específico se les alude comúnmente como "promotores específicos de células" o "promotores específicos de tejidos".
- A promotores que hacen que un gen se exprese en una fase específica de desarrollo o diferenciación celular se les alude comúnmente como "promotores específicos del desarrollo" o "promotores específicos de la diferenciación celular". A promotores que se inducen y provocan la expresión de un gen después de la exposición o el tratamiento de la célula con un agente, molécula biológica, producto químico, ligando, luz o similar que induce al promotor se les alude comúnmente como "promotores inducibles" o "promotores regulables". Se reconoce, además, que dado que en la mayoría de los casos los límites exactos de las secuencias reguladoras no se han definido completamente, los fragmentos de ADN de diferentes longitudes pueden tener una actividad promotora idéntica.
- La secuencia promotora está unida típicamente en su extremo 3' por el sitio de inicio de la transcripción y se extiende aguas arriba (dirección 5') para incluir el número mínimo de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima del fondo. Dentro de la secuencia promotora se encontrará un sitio de inicio de la transcripción (convenientemente definido, por ejemplo, por mapeo con la nucleasa S1), así como dominios de unión a proteínas (secuencias consenso) responsables de la unión de la ARN polimerasa.
- Las expresiones "endonucleasa de restricción" y "enzima de restricción" se utilizan indistintamente y se refieren a una enzima que se une y corta dentro de una secuencia de nucleótidos específica dentro del ADN de doble cadena.
- El término "plásmido" se refiere a un elemento extracromosómico que a menudo porta un gen que no forma parte del metabolismo central de la célula y habitualmente en forma de moléculas circulares de ADN de doble cadena. Elementos de este tipo pueden ser secuencias de replicación autónoma, secuencias de integración del genoma, secuencias de fagos o de nucleótidos, lineales, circulares o superenrolladas, de ADN o ARN de cadena sencilla o de doble cadena, derivadas de cualquier fuente, en las que se han unido o recombinado un cierto número de secuencias de nucleótidos en una construcción única que es capaz de introducir un fragmento promotor y una secuencia de ADN para un producto génico seleccionado junto con la secuencia no traducida 3' apropiada en una célula.
- Vectores virales eucarióticos que se pueden utilizar incluyen, pero no se limitan a vectores de adenovirus, vectores de retrovirus, vectores de virus adeno-asociados, poxvirus, *p. ej.*, vectores de virus vaccinia, vectores de baculovirus o vectores de virus herpes. Vectores no virales incluyen plásmidos, liposomas, lípidos cargados eléctricamente (citofectinas), complejos de proteína de ADN y biopolímeros.
- Un "vector de clonación" se refiere a un "replicón", que es una unidad de longitud de un ácido nucleico que se replica secuencialmente y que comprende un origen de replicación, tal como un plásmido, un fago o un cósmido, al que pertenece otro segmento de ácido nucleico se puede fijar para provocar la replicación del segmento fijado. Determinados vectores de clonación son capaces de replicarse en un tipo de célula, *p. ej.*, bacterias y expresarse en otro, *p. ej.*, células eucarióticas. Vectores de clonación comprenden típicamente una o más secuencias que pueden utilizarse para la selección de células que comprenden el vector y/o uno o más sitios de clonación múltiples para la inserción de secuencias de ácido nucleico de interés.

La expresión "vector de expresión" se refiere a un vehículo diseñado para permitir la expresión de una secuencia de ácido nucleico insertada después de la inserción en una célula huésped. La secuencia de ácido nucleico insertada se coloca en asociación operativa con regiones reguladoras como se describe arriba.

- 5 Vectores se introducen en células huésped mediante métodos bien conocidos en la técnica, *p. ej.*, transfección, electroporación, microinyección, transducción, fusión celular, DEAE dextrano, precipitación con fosfato de calcio, lipofección (fusión de lisosomas), uso de una pistola génica o un vector transportador de ADN.

"Cultivo", "cultivar" y "cultivando", tal como se utiliza en esta memoria, significa incubar células bajo condiciones *in vitro* que permiten el crecimiento o la división celular o mantener las células en un estado vivo. "Células cultivadas", tal como se utiliza en esta memoria, significa células que se propagan *in vitro*.

- 10 Tal como se utiliza en esta memoria, el término "polipéptido" pretende abarcar un "polipéptido" singular, así como "polipéptidos" plurales, y se refiere a una molécula compuesta de monómeros (aminoácidos) enlazados linealmente por enlaces amida (también conocidos como enlaces peptídicos). El término "polipéptido" se refiere a cualquier cadena o cadenas de dos o más aminoácidos y no se refiere a una longitud específica del producto. Así, péptidos, dipéptidos, tripéptidos, oligopéptidos, "proteína", "cadena de aminoácidos" o cualquier otro término o expresión utilizada para referirse a una cadena o cadenas de dos o más aminoácidos, se incluyen dentro de la definición de "polipéptido", y el término "polipéptido" se puede utilizar en lugar de cualquiera de estos términos o de forma intercambiable con ellos. El término "polipéptido" también pretende hacer referencia a los productos de modificaciones post-expresión del polipéptido, que incluyen, sin limitación, glicosilación, acetilación, fosforilación, amidación, derivatización por grupos protectores/bloqueadores conocidos, escisión proteolítica o modificación por aminoácidos que se producen de forma no natural. Un polipéptido puede derivarse de una fuente biológica natural o puede producirse mediante tecnología recombinante, pero no necesariamente se traduce a partir de una secuencia de ácido nucleico designada. Se puede generar de cualquier manera, incluso por síntesis química.

- 25 El término "aminoácido" incluye alanina (Ala o A); arginina (Arg o R); asparagina (Asn o N); ácido aspártico (Asp o D); cisteína (Cys o C); glutamina (Gln o Q); ácido glutámico (Glu o E); glicina (Gly o G); histidina (His o H); isoleucina (Ile o I); leucina (Leu o L); lisina (Lys o K); metionina (Met o M); fenilalanina (Phe o F); prolina (Pro o P); serina (Ser o S); treonina (Thr o T); triptófano (Trp o W); tirosina (Tyr o Y); y valina (Val o V). Aminoácidos no tradicionales también están dentro del alcance de la divulgación e incluyen norleucina, omitina, norvalina, homoserina y otros análogos de residuos de aminoácidos como los descritos en Ellman et al. Meth. Enzym. 202:301-336 (1991). Para generar residuos de aminoácidos que no se producen de forma natural, se pueden utilizar los procedimientos de Noren et al. Science 244:182 (1989) y Ellman *et al.*, supra. Brevemente, estos procedimientos implican la activación química de un ARNt supresor con un residuo de aminoácido que no se produce de forma natural, seguido de la transcripción y traducción *in vitro* del ARN. La introducción del aminoácido no tradicional también puede lograrse utilizando químicas peptídicas conocidas en la técnica. Tal como se utiliza en esta memoria, la expresión "aminoácido polar" incluye aminoácidos que tienen carga neta cero, pero que tienen cargas parciales distintas de cero en diferentes porciones de sus cadenas laterales (*p. ej.*, M, F, W, S, Y, N, Q, C). Estos aminoácidos pueden participar en interacciones hidrofóbicas e interacciones electrostáticas. Tal como se utiliza en esta memoria, la expresión "aminoácido cargado" incluye aminoácidos que pueden tener una carga neta distinta de cero en sus cadenas laterales (*p. ej.*, R, K, H, E, D). Estos aminoácidos pueden participar en interacciones hidrofóbicas e interacciones electrostáticas.

- 40 También se incluyen en la presente divulgación fragmentos o variantes de polipéptidos y cualquier combinación de los mismos. El término "fragmento" o "variante", cuando se refiere a dominios de unión a polipéptidos o moléculas de unión de la presente divulgación incluye cualquier polipéptido que conserve al menos algunas de las propiedades (*p. ej.*, afinidad de unión a FcRn por un dominio de unión a FcRn o variante de Fc, actividad de coagulación para una variante de FVIII, o actividad de unión de FVIII para el fragmento de VWF) del polipéptido de referencia. Fragmentos de polipéptidos incluyen fragmentos proteolíticos, así como fragmentos de delección, además de fragmentos de anticuerpos específicos comentados en otra parte en esta memoria, pero no incluyen el polipéptido que se produce de forma natural de longitud completa (o polipéptido maduro). Variantes de dominios de unión a polipéptidos o moléculas de unión de la presente divulgación incluyen fragmentos como se describe arriba, y también polipéptidos con secuencias de aminoácidos alteradas debido a sustituciones, delecciones o inserciones de aminoácidos. Las variantes pueden ser naturales o no naturales. Las variantes que no se producen de forma natural se pueden producir utilizando técnicas de mutagénesis conocidas en la técnica. Polipéptidos variantes pueden comprender sustituciones, delecciones o adiciones de aminoácidos conservativas o no conservativas.

- 55 Una "sustitución conservativa de aminoácidos" es aquella en la que el residuo de aminoácido se reemplaza con un residuo de aminoácido que tiene una cadena lateral similar. En la técnica se han definido familias de residuos de aminoácidos que tienen cadenas laterales similares, incluyendo cadenas laterales básicas (*p. ej.*, lisina, arginina, histidina), cadenas laterales ácidas (*p. ej.*, ácido aspártico, ácido glutámico), cadenas laterales polares sin carga (*p. ej.*, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (*p. ej.*, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales ramificadas beta (*p. ej.*, treonina, valina, isoleucina) y cadenas laterales aromáticas (*p. ej.*, tirosina, fenilalanina, triptófano, histidina). Por lo tanto, si un aminoácido en un polipéptido se reemplaza por otro aminoácido de la misma familia de cadenas laterales, la sustitución se considera conservativa. En otra

realización, una cadena de aminoácidos se puede reemplazar de manera conservadora con una cadena estructuralmente similar que difiere en el orden y/o la composición de los miembros de la familia de cadenas laterales.

La expresión "porcentaje de identidad", como se conoce en la técnica, es una relación entre dos o más secuencias polipeptídicas o dos o más secuencias polinucleotídicas, según se determina comparando las secuencias. En la técnica, "identidad" también significa el grado de relación de secuencia entre secuencias polipeptídicas o polinucleotídicas, según sea el caso, determinado por la coincidencia entre cadenas de secuencias de este tipo. La "identidad" se puede calcular fácilmente mediante métodos conocidos, incluyendo, pero no limitados a los descritos en: Computational Molecular Biology (Lesk, A.M., ed.) Oxford University Press, Nueva York (1988); Biocomputing Informatics and Genome Projects (Smith, D.W., ed.) Academic Press, Nueva York (1993); Computer Analysis of Sequence Data, Parte I (Griffin, A. M., y Griffin, H. G., eds.) Humana Press, New Jersey (1994); Sequence Analysis in Molecular Biology (von Heinje, G., ed.) Academic Press (1987); y Sequence Analysis Primer (Gribskov, M. y Devereux, J., eds.) Stockton Press, Nueva York (1991). Métodos preferidos para determinar la identidad están diseñados para dar la mejor coincidencia entre las secuencias analizadas. Métodos para determinar la identidad están codificados en programas informáticos disponibles públicamente. Los alineamientos de secuencias y los cálculos de porcentaje de identidad se pueden realizar utilizando software de análisis de secuencias, tales como el programa Megalign de la suite de computación bioinformática LASERGENE (DNASTAR Inc., Madison, WI), la suite de programas GCG (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI), BLASTP, BLASTN, BLASTX (Altschul et al., J. Mol. Biol. 215:403 (1990)) y DNASTAR (DNASTAR, Inc. 1228 S. Park St. Madison, WI 53715 EE.UU.). Dentro del contexto de esta solicitud, se entenderá que cuando se utiliza software de análisis de secuencias para el análisis, los resultados del análisis se basarán en los "valores predeterminados" del programa al que se hace referencia, a menos que se especifique lo contrario. Tal como se utiliza en esta memoria, "valores predeterminados" significará cualquier conjunto de valores o parámetros que originalmente se cargan con el software cuando se inicializa por primera vez. Con el fin de determinar el porcentaje de identidad entre una secuencia optimizada de BDD FVIII de la divulgación y una secuencia de referencia, solo se utilizan los nucleótidos de la secuencia de referencia correspondientes a los nucleótidos de la secuencia optimizada de BDD FVIII de la divulgación para calcular el porcentaje de identidad. Por ejemplo, cuando se compara una secuencia de nucleótidos de FVIII de longitud completa que contiene el dominio B con una secuencia de nucleótidos de FVIII optimizada con dominio B eliminado (BDD) de la divulgación, la parte del alineamiento que incluye los dominios A1, A2, A3, C1 y C2 se utilizará para calcular el porcentaje de identidad. Los nucleótidos en la porción de la secuencia de FVIII de longitud completa que codifica el dominio B (lo que dará como resultado un gran "hueco" en el alineamiento) no se contarán como un desajuste. Además, al determinar el porcentaje de identidad entre una secuencia optimizada de BDD FVIII de la divulgación, o una parte designada de la misma (p. ej., los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO : 3), y una secuencia de referencia, el porcentaje de identidad será calculado por alineamiento dividiendo el número de nucleótidos emparejados por el número total de nucleótidos en la secuencia completa de la secuencia BDD-FVIII optimizada, o una parte designada de la misma, como se menciona en esta memoria.

Tal como se utiliza en esta memoria, "nucleótidos correspondientes a nucleótidos en la secuencia optimizada de BDD FVIII de la divulgación" se identifican mediante el alineamiento de la secuencia optimizada de BDD FVIII de la divulgación para maximizar la identidad con la secuencia de referencia de FVIII. El número utilizado para identificar un aminoácido equivalente en una secuencia de FVIII de referencia se basa en el número utilizado para identificar el aminoácido correspondiente en la secuencia optimizada de BDD FVIII de la divulgación.

Una proteína de "fusión" o "quimérica" comprende una primera secuencia de aminoácidos enlazada a una segunda secuencia de aminoácidos con la que no está unida de forma natural en la naturaleza. Las secuencias de aminoácidos que normalmente existen en proteínas separadas se pueden unir en el polipéptido de fusión, o las secuencias de aminoácidos que normalmente existen en la misma proteína se pueden colocar en una nueva disposición en el polipéptido de fusión, p. ej., la fusión de un dominio de Factor VIII de la divulgación con un dominio Fc de Ig. Una proteína de fusión se crea, por ejemplo, mediante síntesis química o creando y traduciendo un polinucleótido en el que las regiones peptídicas se codifican en la relación deseada. Una proteína quimérica puede comprender, además, una segunda secuencia de aminoácidos asociada con la primera secuencia de aminoácidos mediante un enlace covalente no peptídico o un enlace no covalente.

Tal como se utiliza en esta memoria, la expresión "sitio de inserción" se refiere a una posición en un polipéptido de FVIII, o fragmento, variante o derivado del mismo, que está inmediatamente aguas arriba de la posición en la que se puede insertar un resto heterólogo. Un "sitio de inserción" se especifica como un número, siendo el número el número del aminoácido en el FVIII nativo maduro (SEQ ID NO: 15; FIG. 11A) al que corresponde el sitio de inserción, que es inmediatamente N-terminal con respecto a la posición de la inserción. Por ejemplo, la frase "a3 comprende un resto heterólogo en un sitio de inserción que corresponde al aminoácido 1656 de SEQ ID NO: 15" indica que el resto heterólogo está situado entre dos aminoácidos que corresponden al aminoácido 1656 y al aminoácido 1657 de SEQ ID NO: 15.

La frase "inmediatamente aguas abajo de un aminoácido", tal como se utiliza en esta memoria, se refiere a la posición justo al lado del grupo carboxilo terminal del aminoácido. De manera similar, la frase "inmediatamente aguas arriba de un aminoácido" se refiere a la posición justo al lado del grupo amino terminal del aminoácido.

El término "insertado" y las expresiones "se inserta", "insertado en" o términos y expresiones gramaticalmente relacionados, tal como se utilizan en esta memoria, se refieren a la posición de un resto heterólogo en un polipéptido de FVIII recombinante, en relación con la posición análoga en el FVIII humano nativo maduro. Tal como se utiliza en esta memoria, el término y las expresiones se refieren a las características del polipéptido de FVIII recombinante en relación con el FVIII humano nativo maduro, y no indican, implican ni infieren en método o proceso alguno mediante el cual se elaboró el polipéptido de FVIII recombinante.

Tal como se utiliza en esta memoria, el término "semivida" se refiere a una semivida biológica de un polipéptido particular *in vivo*. La semivida puede representarse por el tiempo requerido para que la mitad de la cantidad administrada a un sujeto se elimine de la circulación y/u otros tejidos del animal. Cuando se construye una curva de aclaramiento de un polipéptido dado en función del tiempo, la curva es habitualmente bifásica con una fase  $\alpha$  rápida y una fase  $\beta$  más larga. La fase  $\alpha$  representa típicamente un equilibrio del polipéptido Fc administrado entre el espacio intra-vascular y extra-vascular y está, en parte, determinada por el tamaño del polipéptido. La fase  $\beta$  representa típicamente el catabolismo del polipéptido en el espacio intravascular. En algunas realizaciones, el FVIII y las proteínas quiméricas que comprenden FVIII son monofásicos y, por lo tanto, no tienen una fase alfa, sino solo una fase beta única. Por lo tanto, en determinadas realizaciones, el término semivida, tal como se utiliza en esta memoria, se refiere a la semivida del polipéptido en la fase  $\beta$ .

El término "enlazado", tal como se utiliza en esta memoria, se refiere a una primera secuencia de aminoácidos o secuencia de nucleótidos unida de forma covalente o no covalente a una segunda secuencia de aminoácidos o secuencia de nucleótidos, respectivamente. La primera secuencia de aminoácidos o nucleótidos se puede unir o yuxtaponer directamente a la segunda secuencia de aminoácidos o nucleótidos o, alternativamente, una secuencia intermedia puede unir covalentemente la primera secuencia con la segunda secuencia. El término "enlazado" significa no solo una fusión de una primera secuencia de aminoácidos a una segunda secuencia de aminoácidos en el extremo C o el extremo N, sino que también incluye la inserción de la primera secuencia de aminoácidos completa (o la segunda secuencia de aminoácidos) en cualquiera de dos aminoácidos de la segunda secuencia de aminoácidos (o la primera secuencia de aminoácidos, respectivamente). En una realización, la primera secuencia de aminoácidos se puede unir a una segunda secuencia de aminoácidos mediante un enlace peptídico o un enlazador. La primera secuencia de nucleótidos se puede enlazar a una segunda secuencia de nucleótidos mediante un enlace fosfodiéster o un enlazador. El enlazador puede ser un péptido o un polipéptido (para cadenas de polipéptidos) o un nucleótido o una cadena de nucleótidos (para cadenas de nucleótidos) o cualquier resto químico (tanto para cadenas de polipéptidos como de polinucleótidos). El término "enlazado" también se indica con un guión (-).

Tal como se utiliza en esta memoria, la expresión "asociado con" se refiere a un enlace covalente o no covalente formado entre una primera cadena de aminoácidos y una segunda cadena de aminoácidos. En una realización, la expresión "asociado con" significa un enlace no peptídico covalente o un enlace no covalente. Esta asociación se puede indicar con dos puntos, *es decir*, (:). En otra realización, significa un enlace covalente excepto un enlace peptídico. Por ejemplo, el aminoácido cisteína comprende un grupo tiol que puede formar un enlace disulfuro o un puente con un grupo tiol en un segundo residuo de cisteína. En la mayoría de las moléculas de IgG que se producen de forma natural, las regiones CH1 y CL están asociadas por un enlace disulfuro y las dos cadenas pesadas están asociadas por dos enlaces disulfuro en las posiciones correspondientes a 239 y 242 utilizando el sistema de numeración de Kabat (posición 226 o 229, sistema de numeración de la UE). Ejemplos de enlaces covalentes incluyen, pero no se limitan a un enlace peptídico, un enlace metálico, un enlace de hidrógeno, un enlace disulfuro, un enlace sigma, un enlace pi, un enlace delta, un enlace glicosídico, un enlace agnóstico, un enlace doblado, un enlace dipolar, una cadena posterior pi, un enlace doble, un enlace triple, un enlace cuádruple, un enlace quíntuple, un enlace séxtuple, conjugación, hiperconjugación, aromaticidad, hapticidad o antienlace. Ejemplos no limitantes de enlace no covalente incluyen un enlace iónico (p. ej., enlace catiónico-pi o enlace de sal), un enlace metálico, un enlace de hidrógeno (p. ej., enlace de hidrógeno, complejo de dihidrógeno, enlace de hidrógeno de baja barrera o enlace de hidrógeno simétrico), fuerza de van der Waals, fuerza de dispersión de London, un enlace mecánico, un enlace halógeno, aurofilia, intercalación, apilamiento, fuerza entrópica o polaridad química.

La expresión "híbrido monómero-dímero", tal como se utiliza en esta memoria, se refiere a una proteína quimérica que comprende una primera cadena polipeptídica y una segunda cadena polipeptídica, que están asociadas entre sí mediante un enlace disulfuro, en donde la primera cadena comprende un factor de coagulación, p. ej., Factor VIII, y una primera región Fc y la segunda cadena comprende, consiste esencialmente en o consiste en una segunda región Fc sin el factor de coagulación. La construcción híbrida de monómero-dímero es, por lo tanto, un híbrido que comprende un aspecto de monómero que tiene solo un factor de coagulación y un aspecto de dímero que tiene dos regiones Fc.

La hemostasis, tal como se utiliza en esta memoria, significa detener o ralentizar el sangrado o la hemorragia; o la detención o disminución del flujo de sangre a través de un vaso sanguíneo o parte del cuerpo.

Trastorno hemostático, tal como se utiliza en esta memoria, significa una afección heredada o adquirida genéticamente caracterizada por una tendencia a la hemorragia, ya sea espontáneamente o como resultado de un traumatismo, debido a una capacidad alterada o incapacidad para formar un coágulo de fibrina. Ejemplos de trastornos de este tipo incluyen las hemofilias. Las tres formas principales son la hemofilia A (deficiencia de factor VIII), la hemofilia B (deficiencia de factor IX o "enfermedad de Christmas") y la hemofilia C (deficiencia de factor XI, leve tendencia al sangrado). Otros trastornos hemostáticos incluyen, p. ej., la enfermedad de von Willebrand, deficiencia de Factor XI (deficiencia de PTA), deficiencia

de Factor XII, deficiencias o anomalías estructurales en fibrinógeno, protrombina, Factor V, Factor VII, Factor X o Factor XIII, síndrome de Bernard-Soulier, que es un defecto o una deficiencia en GPIb. GPIb, el receptor de vWF, puede ser defectuoso y conducir a la falta de formación de coágulos primarios (hemostasia primaria) y de la tendencia incrementada al sangrado) y trombastenia de Glanzman y Naegeli (trombastenia de Glanzmann). En la insuficiencia hepática (formas aguda y crónica), hay una producción insuficiente de factores de coagulación por parte del hígado; esto puede aumentar el riesgo de sangrado.

Las moléculas de ácido nucleico aisladas, los polipéptidos aislados o los vectores que comprenden la molécula de ácido nucleico aislada de la divulgación se pueden utilizar de forma profiláctica. Tal como se utiliza en esta memoria, la expresión "tratamiento profiláctico" se refiere a la administración de una molécula antes de un episodio de sangrado. En una realización, el sujeto que necesita un agente hemostático general se somete o está a punto de someterse a cirugía. Un polinucleótido, polipéptido o vector de la divulgación se puede administrar antes o después de la cirugía como profiláctico. El polinucleótido, polipéptido o vector de la divulgación se puede administrar durante o después de la cirugía para controlar un episodio de sangrado agudo. La cirugía puede incluir, pero no se limita a trasplante de hígado, resección de hígado, procedimientos dentales o trasplante de células madre.

Las moléculas de ácido nucleico aisladas, los polipéptidos aislados o los vectores de la divulgación también se utilizan para el tratamiento a demanda. La expresión "tratamiento a demanda" se refiere a la administración de una molécula de ácido nucleico aislado, un polipéptido aislado o un vector en respuesta a los síntomas de un episodio de sangrado o antes de una actividad que pueda provocar el sangrado. En un aspecto, el tratamiento a demanda puede administrarse a un sujeto cuando comienza el sangrado tal como después de una lesión, o cuando se espera un sangrado, tal como antes de la cirugía. En otro aspecto, el tratamiento a demanda puede darse previo a actividades que aumentan el riesgo de sangrado tales como los deportes de contacto.

Tal como se utiliza en esta memoria, la expresión "sangrado agudo" se refiere a un episodio de sangrado independientemente de la causa subyacente. Por ejemplo, un sujeto puede tener un traumatismo, uremia, un trastorno hemorrágico hereditario (p. ej., deficiencia de factor VII), un trastorno plaquetario o resistencia debida al desarrollo de anticuerpos frente a factores de coagulación.

Tratar, tratamiento, tratando, tal como se utiliza en esta memoria, se refiere, p. ej., a la reducción de la gravedad de una enfermedad o afección; la reducción en la duración del curso de una enfermedad; la mejora de uno o más síntomas asociados con una enfermedad o afección; la provisión de efectos beneficiosos a un sujeto con una enfermedad o afección, sin curar necesariamente la enfermedad o afección, o la profilaxis de uno o más síntomas asociados con una enfermedad o afección. En una realización, el término "tratando" o "tratamiento" significa mantener un nivel valle de FVIII de al menos aproximadamente 1 UI/dL, 2 UI/dL, 3 UI/dL, 4 UI/dL, 5 UI/dL, 6 UI/dL, 7 UI/dL, 8 UI/dL, 9 UI/dL, 10 UI/dL, 11 UI/dL, 12 UI/dL, 13 UI/dL, 14 UI/dL, 15 UI/dL, 16 UI/dL, 17 UI/dL, 18 UI/dL, 19 UI/dL o 20 UI/dL en un sujeto mediante la administración de una molécula de ácido nucleico aislada, un polipéptido aislado o un vector de la divulgación. En otra realización, tratando o tratamiento significa mantener un nivel valle de FVIII entre aproximadamente 1 y aproximadamente 20 UI/dL, aproximadamente 2 y aproximadamente 20 UI/dL, aproximadamente 3 y aproximadamente 20 UI/dL, aproximadamente 4 y aproximadamente 20 UI/dL, aproximadamente 5 y aproximadamente 20 UI/dL, aproximadamente 6 y aproximadamente 20 UI/dL, aproximadamente 7 y aproximadamente 20 UI/dL, aproximadamente 8 y aproximadamente 20 UI/dL, aproximadamente 9 y aproximadamente 20 UI/dL, o aproximadamente 10 y aproximadamente 20 UI/dL. El tratamiento o tratar una enfermedad o afección también puede incluir el mantenimiento de la actividad del FVIII en un sujeto a un nivel equiparable a por lo menos aproximadamente el 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 11 %, 12 %, 13 %, 14 %, 15 %, 16 %, 17 %, 18 %, 19 % o 20 % de la actividad de FVIII en un sujeto no hemofílico. El nivel valle mínimo requerido para el tratamiento se puede medir mediante uno o más métodos conocidos y se puede ajustar (aumentar o disminuir) para cada persona.

"Administrar", tal como se utiliza en esta memoria, significa dar una molécula de ácido nucleico que codifica el Factor VIII farmacéuticamente aceptable, un polipéptido del Factor VIII o un vector que comprende una molécula de ácido nucleico que codifica el Factor VIII de la divulgación a un sujeto a través de una vía farmacéuticamente aceptable. Vías de administración pueden ser intravenosa, p. ej., inyección intravenosa e infusión intravenosa. Vías de administración adicionales incluyen, p. ej., administración subcutánea, intramuscular, oral, nasal y pulmonar. Las moléculas de ácido nucleico, los polipéptidos y los vectores se pueden administrar como parte de una composición farmacéutica que comprende al menos un excipiente.

Tal como se utiliza en esta memoria, la frase "sujeto que lo necesita" incluye sujetos tales como mamíferos que se beneficiarían de la administración de una molécula de ácido nucleico, un polipéptido o un vector de la divulgación, p. ej., para mejorar la hemostasia. En una realización, los sujetos incluyen, pero no se limitan a individuos con hemofilia. En otra realización, los sujetos incluyen, pero no se limitan a los individuos que han desarrollado un inhibidor de FVIII y, por lo tanto, necesitan una terapia de derivación. El sujeto puede ser un adulto o un menor (p. ej., menor de 12 años).

Tal como se utiliza en esta memoria, la expresión "factor de coagulación" se refiere a moléculas, o análogos de las mismas, que se producen de forma natural o se producen de forma recombinante que previenen o reducen la duración de un episodio hemorrágico en un sujeto. En otras palabras, significa moléculas que tienen actividad pro-coagulación, es decir,



son responsables de la conversión de fibrinógeno en una malla de fibrina insoluble que hace que la sangre coagule o se coagule. Un "factor de coagulación activable" es un factor de coagulación en una forma inactiva (*p. ej.*, en su forma de zimógeno) que es capaz de convertirse en una forma activa.

La actividad de coagulación, tal como se utiliza en esta memoria, significa la capacidad de participar en una cascada de reacciones bioquímicas que culmina en la formación de un coágulo de fibrina y/o reduce la gravedad, duración o frecuencia de hemorragia o episodio de sangrado.

Tal como se utiliza en esta memoria, los términos "heterólogo" o "exógeno" se refieren a moléculas que normalmente no se encuentran en un contexto dado, *p. ej.*, en una célula o en un polipéptido. Por ejemplo, una molécula exógena o heteróloga se puede introducir en una célula y solo está presente después de la manipulación de la célula, *p. ej.*, mediante transfección u otras formas de ingeniería genética, o una secuencia de aminoácidos heteróloga puede estar presente en una proteína en la que no se encuentra de forma natural.

Tal como se utiliza en esta memoria, la expresión "secuencia de nucleótidos heteróloga" se refiere a una secuencia de nucleótidos que no se produce de forma natural con una secuencia de polinucleótidos dada. En una realización, la secuencia de nucleótidos heteróloga codifica un polipéptido capaz de extender la semivida de FVIII. En otra realización, la secuencia de nucleótidos heteróloga codifica un polipéptido que aumenta el radio hidrodinámico de FVIII. En otras realizaciones, la secuencia de nucleótidos heteróloga codifica un polipéptido que mejora una o más propiedades farmacocinéticas de FVIII sin afectar significativamente a su actividad o función biológica (*p. ej.*, su actividad procoagulante). En algunas realizaciones, el FVIII está enlazado o conectado al polipéptido codificado por la secuencia de nucleótidos heteróloga mediante un enlazador. Ejemplos no limitantes de restos polipeptídicos codificados por secuencias de nucleótidos heterólogas incluyen una región constante de inmunoglobulina o una porción de la misma, albúmina o un fragmento de la misma, un resto de unión a albúmina, una transferrina, los polipéptidos PAS de la Solicitud de Pat. de EE.UU. Nº 20100292130, una secuencia HAP, transferrina o un fragmento de la misma, el péptido C-terminal (CTP) de la subunidad  $\beta$  de la gonadotropina coriónica humana, una molécula pequeña que se une a la albúmina, una secuencia de XTEN, restos de unión a FcRn (*p. ej.*, regiones Fc completas o partes de las mismas que se unen a FcRn), regiones Fc de cadena sencilla (regiones ScFc, *p. ej.*, como se describe en los documentos US 2008/0260738, WO 2008/012543 o WO 2008/1439545), enlazadores de poliglicina, enlazadores de poliserina, péptidos y polipéptidos cortos de 6-40 aminoácidos de dos tipos de aminoácidos seleccionados de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P) con diversos grados de estructura secundaria desde menos del 50 % hasta más del 50 %, entre otros, o dos o más combinaciones de los mismos. En algunas realizaciones, el polipéptido codificado por la secuencia de nucleótidos heteróloga está enlazado a un resto no polipeptídico. Ejemplos no limitativos de restos no polipeptídicos incluyen polietilenglicol (PEG), moléculas pequeñas que se unen a albúmina, ácido polisialico, hidroxietil almidón (HES, por sus siglas en inglés), un derivado del mismo o cualquier combinación de los mismos.

Tal como se utiliza en esta memoria, la expresión "región Fc" se define como la porción de un polipéptido que corresponde a la región Fc de la Ig nativa, *es decir*, como se forma por la asociación dimérica de los respectivos dominios Fc de sus dos cadenas pesadas. Una región Fc nativa forma un homodímero con otra región Fc. Por el contrario, la expresión "región Fc fusionada genéticamente" o "región Fc de cadena única" (región scFc), tal como se utiliza en esta memoria, se refiere a una región Fc dimérica sintética compuesta por dominios Fc enlazados genéticamente dentro de una cadena polipeptídica única (*es decir*, codificada en una única secuencia genética contigua).

En una realización, la "región Fc" se refiere a la porción de una sola cadena pesada de Ig que comienza en la región bisagra justo aguas arriba del sitio de escisión de papaína (*es decir*, el residuo 216 en IgG, tomando el primer residuo de la región constante de la cadena pesada como 114) y terminando en el extremo C del anticuerpo. Por consiguiente, un dominio Fc completo comprende al menos un dominio bisagra, un dominio CH2 y un dominio CH3.

La región Fc de una región constante de Ig, dependiendo del isotipo de Ig, puede incluir los dominios CH2, CH3 y CH4, así como la región bisagra. Proteínas quiméricas que comprenden una región Fc de una Ig otorgan varias propiedades deseables a una proteína quimérica, incluyendo una estabilidad incrementada, una semivida en suero incrementada (véase Capon et al., 1989, Nature 337:525), así como la unión a receptores Fc tales como el receptor de Fc neonatal (FcRn) (Patentes de EE.UU. N°s 6.086.875, 6.485.726, 6.030.613; documentos WO 03/077834; US2003-0235536A1).

Una "secuencia de nucleótidos de referencia", cuando se utiliza en esta memoria como una comparación con una secuencia de nucleótidos de la divulgación, es una secuencia de polinucleótidos esencialmente idéntica a la secuencia de nucleótidos de la divulgación, excepto que las partes correspondientes a la secuencia de FVIII no están optimizadas. Por ejemplo, la secuencia de nucleótidos de referencia para una molécula de ácido nucleico que consiste en BDD FVIII optimizado por codones de SEQ ID NO: 1 y una secuencia de nucleótidos heteróloga que codifica una región Fc de cadena sencilla enlazada a SEQ ID NO: 1 en su extremo 3' es una molécula de ácido nucleico que consiste en el BDD FVIII original (o "progenitor") de SEQ ID NO: 16 (FIG. 1I) y la secuencia de nucleótidos heteróloga idéntica que codifica una región Fc de cadena sencilla enlazada a SEQ ID NO: 16 en su extremo 3'.

Un "índice de adaptación de codones", tal como se utiliza en esta memoria, se refiere a una medida del sesgo de uso de codones. Un índice de adaptación de codones (CAI, por sus siglas en inglés) mide la desviación de una secuencia de

genes codificantes de proteínas dada con respecto a un conjunto de genes de referencia (Sharp PM y Li WH, Nucleic Acids Res. 15(3):1281-95 (1987)). El CAI se calcula determinando la media geométrica del peso asociado a cada uno de los codones a lo largo de la secuencia del gen (medido en codones):

$$CAI = \exp\left(\frac{1}{L} \sum_{l=1}^L \ln(w_i(l))\right), \quad (I)$$

Para cada uno de los aminoácidos, el peso de cada uno de sus codones, en el CAI, se computa como la relación entre la frecuencia observada del codón ( $f_i$ ) y la frecuencia del codón sinónimo ( $f_j$ ) para ese aminoácido:

$$w_i = \frac{f_i}{\max(f_j)} \quad i, j \in [\text{codones sinónimos para aminoácidos}] \quad (II)$$

Tal como se utiliza en esta memoria, el término "optimizado" con respecto a las secuencias de nucleótidos, se refiere a una secuencia de polinucleótidos que codifica un polipéptido, en donde la secuencia de polinucleótidos ha sido mutada para potenciar una propiedad de esa secuencia de polinucleótidos. En algunas realizaciones, la optimización se realiza para aumentar los niveles de transcripción, aumentar los niveles de traducción, aumentar los niveles de ARNm en estado estacionario, aumentar o disminuir la unión de proteínas reguladoras tales como factores de transcripción generales, aumentar o disminuir el corte y empalme o aumentar el rendimiento del polipéptido producido por la secuencia de polinucleótidos. Ejemplos de cambios que se pueden realizar en una secuencia de polinucleótidos para optimizarla incluyen la optimización por codones, la optimización del contenido de G/C, la eliminación de secuencias repetidas, la eliminación de elementos ricos en AT, la eliminación de sitios de corte y empalme crípticos, la eliminación de elementos que actúan en cis que reprimen la transcripción o traducción, añadiendo o eliminando secuencias poli-T o poli-A, añadiendo secuencias alrededor del sitio de inicio de la transcripción que potencian la transcripción, tales como las secuencias de consenso de Kozak, la eliminación de secuencias que podrían formar estructuras de bucle del tallo, la eliminación de secuencias desestabilizadoras y dos o más combinaciones de los mismos.

#### Secuencia de Polinucleótidos que Codifica la Proteína FVIII

La presente invención describe un ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende al menos un 95 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 tal como se define en las reivindicaciones independientes, y determinadas características opcionales de las mismas se definen en las reivindicaciones dependientes.

En algunas realizaciones, la presente divulgación está dirigida a moléculas de ácido nucleico optimizadas por codones que codifican un polipéptido con actividad de FVIII. En algunas realizaciones, el polinucleótido codifica un polipéptido de FVIII de longitud completa. En otras realizaciones, la molécula de ácido nucleico codifica un polipéptido de FVIII con dominio B eliminado (BDD), en donde se elimina todo o una parte del dominio B de FVIII. En una realización particular, la molécula de ácido nucleico codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 86 %, al menos aproximadamente el 87 %, al menos aproximadamente el 88 %, al menos aproximadamente el 89 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 %, al menos aproximadamente el 96 %, al menos aproximadamente el 97 %, al menos aproximadamente el 98 % o al menos aproximadamente el 99 % de identidad de secuencia con SEQ ID NO: 17 (FIG. 1J) o un fragmento de la misma. En una realización, la molécula de ácido nucleico codifica un polipéptido que tiene la secuencia de aminoácidos de SEQ ID NO: 17 o un fragmento de la misma.

En algunas realizaciones, la molécula de ácido nucleico de la divulgación codifica un polipéptido de FVIII que comprende un péptido señal o un fragmento del mismo. En otras realizaciones, la molécula de ácido nucleico codifica un polipéptido de FVIII que carece de un péptido señal. En algunas realizaciones, el péptido señal comprende los aminoácidos 1-19 de SEQ ID NO: 17.

En otras realizaciones, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 % o al menos un 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, los nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B). En otras realizaciones, la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, los nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o el fragmento de dominio B) o los nucleótidos 58 a 4374 de SEQ ID NO: 71. En aún otras realizaciones, la secuencia de nucleótidos comprende los nucleótidos 1-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, los nucleótidos 1-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B) o los nucleótidos 1 a 4374 de SEQ ID NO: 71.

En algunas realizaciones, la secuencia de nucleótidos comprende una secuencia de ácido nucleico que codifica un péptido señal. En determinadas realizaciones, el péptido señal es un péptido señal FVIII. En algunas realizaciones, la secuencia de ácido nucleico que codifica un péptido señal está optimizada por codones. En una realización particular, la secuencia de ácido nucleico que codifica un péptido señal tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos al menos 98 %, al menos 99 % o 100 % de identidad de secuencia con (i) los nucleótidos 1 a 57 de SEQ ID NO: 1; (ii) los nucleótidos 1 a 57 de SEQ ID NO: 2; (iii) los nucleótidos 1 a 57 de SEQ ID NO: 3; (iv) nucleótidos 1 a 57 de SEQ ID NO: 4; (v) los nucleótidos 1 a 57 de SEQ ID NO: 5; (vi) los nucleótidos 1 a 57 de SEQ ID NO: 6; (vii) los nucleótidos 1 a 57 de SEQ ID NO: 70; (viii) los nucleótidos 1 a 57 de SEQ ID NO: 71; o (ix) los nucleótidos 1 a 57 de SEQ ID NO: 68.

Las SEQ ID NOs: 1-6, 70 y 71 son versiones optimizadas de SEQ ID NO: 16, la secuencia de nucleótidos de FVIII inicial o "parental" o "de tipo salvaje". La SEQ ID NO: 16 codifica un FVIII humano con el dominio B eliminado. Si bien las SEQ ID NOs: 1-6, 70 y 71 se derivan de una forma específica de FVIII con el dominio B eliminado (SEQ ID NO: 16), debe entenderse que la presente divulgación también está dirigida a versiones optimizadas de ácidos nucleico que codifican otras versiones de FVIII. Por ejemplo, otras versiones de FVIII pueden incluir FVIII de longitud completa, otras delecciones del dominio B de FVIII (descritas más adelante) u otros fragmentos de FVIII que conservan la actividad de FVIII.

"Un polipéptido con actividad de FVIII", tal como se utiliza en esta memoria, significa un polipéptido de FVIII funcional en su papel normal en la coagulación, a menos que se especifique lo contrario. La expresión polipéptido con actividad de FVIII incluye un fragmento, variante, análogo o derivado funcional del mismo que conserva la función del Factor VIII de tipo salvaje de longitud completa en la vía de la coagulación. "Un polipéptido con actividad de FVIII" se utiliza indistintamente con proteína FVIII, polipéptido FVIII o FVIII. Ejemplos de funciones del FVIII incluyen, pero no se limitan a la capacidad de activar la coagulación, la capacidad de actuar como cofactor del factor IX o la capacidad de formar un complejo de tenasa con el factor IX en presencia de  $\text{Ca}^{2+}$  y fosfolípidos, que luego convierte el Factor X en la forma activada Xa. En una realización, un polipéptido que tiene actividad de FVIII comprende dos cadenas polipeptídicas, teniendo la primera cadena la cadena pesada de FVIII y teniendo la segunda cadena la cadena ligera de FVIII. En otra realización, el polipéptido que tiene actividad de FVIII es FVIII de cadena sencilla. El FVIII de cadena sencilla puede contener una o más mutaciones o sustituciones en el residuo de aminoácido 1645 y/o 1648 correspondiente a la secuencia de FVIII maduro. Véase la solicitud internacional N° PCT/US2012/045784. La proteína FVIII puede ser la proteína FVIII humana, porcina, canina, de rata o murina. Además, las comparaciones entre el FVIII de seres humanos y otras especies han identificado residuos conservados que probablemente sean necesarios para la función (Cameron et al., *Thromb. Haemost.* 79:317-22 (1998); documento US 6.251.632).

El "dominio B" de FVIII, tal como se utiliza en esta memoria, es el mismo que el dominio B conocido en la técnica que se define por la identidad de la secuencia de aminoácidos interna y los sitios de escisión proteolítica por la trombina, p. ej., los residuos Ser741-Arg1648 de FVIII humano de longitud completa. Los otros dominios del FVIII humano están definidos por los siguientes residuos de aminoácidos: A1, residuos Ala1-Arg372; A2, residuos Ser373-Arg740; A3, residuos Ser1690-Ile2032; C1, residuos Arg2033-Asn2172; C2, residuos Ser2173-Tyr2332. La secuencia A3-C1-C2 incluye los residuos Ser1690-Tyr2332. A la secuencia restante, los residuos Glu1649-Arg1689, se la alude habitualmente como péptido de activación de cadena ligera de FVIII. Las ubicaciones de los límites para todos los dominios, incluyendo los dominios B, para el FVIII porcino, de ratón y canino también se conocen en la técnica. Un ejemplo de BDD FVIII es BDD FVIII recombinante REFACTO® (Wyeth Pharmaceuticals, Inc.).

Un "FVIII suprimido en el dominio B" puede tener las delecciones completas o parciales descritas en las Patentes de EE.UU. N°s 6.316.226, 6.346.513, 7.041.635, 5.789.203, 6.060.447, 5.595.886, 6.228.620, 5.972.885, 6.048.720, 5.543.502, 5.610.278, 5.171.844, 5.112.950, 4.868.112 y 6.458.563. En algunas realizaciones, una secuencia de FVIII con el dominio B eliminado de la presente divulgación comprende cualquiera de las delecciones descritas en la col. 4, línea 4 a la col. 5, línea 28 y los ejemplos 1-5 de la Patente de EE.UU. N° 6.316.226 (también en el documento de EE.UU. 6.346.513). En algunas realizaciones, un FVIII con el dominio B eliminado de la presente divulgación tiene una delección descrita en la col. 2, líneas 26-51 y los ejemplos 5-8 de la Patente de EE.UU. N° 5.789.203 (también documentos US 6.060.447, US 5.595.886 y US 6.228.620). En algunas realizaciones, un FVIII con el dominio B eliminado tiene una delección descrita en la col. 1, línea 25 a la col. 2, línea 40 de la Patente de EE.UU. N° 5.972.885; col. 6, líneas 1-22 y el ejemplo 1 de la Patente de EE.UU. 6.048.720; col. 2, líneas 17-46 de la Patente de EE.UU. N° 5.543.502; col. 4, línea 22 a la col. 5, línea 36 de la Patente de EE.UU. N° 5.171.844; col. 2, líneas 55-68, figura 2 y el ejemplo 1 de la Patente de EE.UU. N° 5.112.950; col. 2, línea 2 a la col. 19, línea 21 y tabla 2 de la Patente de EE.UU. N° 4.868.112; col. 2, línea 1 a la col. 3, línea 19, col. 3, línea 40 a la col. 4, línea 67, col. 7, línea 43 a la col. 8, línea 26 y col. 11, línea 5 a la col. 13, línea 39 de la Patente de EE.UU. N° 7.041.635; o col. 4, líneas 25-53, de la Patente de EE.UU. N° 6.458.563. En algunas realizaciones, un FVIII con el dominio B eliminado tiene una delección de la mayor parte del dominio B, pero aún contiene secuencias amino-terminales del dominio B que son esenciales para el procesamiento proteolítico *in vivo* del producto de traducción primario en dos cadenas polipeptídicas, como se describe en el documento WO 91/09122. En algunas realizaciones, un FVIII con dominio B eliminado se construye con una delección de los aminoácidos 747-1638, es decir, virtualmente una delección completa del dominio B. Hoeben R.C., et al. *J. Biol. Chem.* 265 (13): 7318-7323 (1990). Un FVIII con dominio B eliminado también puede contener una delección de los aminoácidos 771-1666 o los aminoácidos 868-1562 de FVIII. Meulien P., et al. *Protein Eng.* 2(4): 301-6 (1988). Delecciones del dominio B adicionales que forman parte de la divulgación incluyen, p. ej., delección de los aminoácidos 982 a 1562 o 760 a 1639 (Toole et al., *Proc. Natl. Acad. Sci. USA* (1986) 83, 5939-5942), 797 a 1562

(Eaton, et al. Biochemistry (1986) 25:8343-8347)), 741 a 1646 (Kaufman (solicitud PCT publicada N° WO 87/04187)), 747-1560 (Sarver, et al., DNA (1987) 6:553-564)), 741 a 1648 (Pasek (solicitud PCT N° 88/00831)), 816 a 1598 o 741 a 1689 (Lagner (Behring Inst. Mitt. (1988) N° 82:16-25, documento EP 295597). Cada una de las delecciones anteriores se pueden realizar en cualquier secuencia de FVIII.

- 5 Un cierto número de moléculas FVIII funcionales, incluyendo delecciones del dominio B, se describen en las siguientes patentes US 6.316.226 y US 6.346.513, ambas asignadas a Baxter; US 7.041.635 asignada a In2Gen; US 5.789.203, US 6.060.447, US 5.595.886 y US 6.228.620 asignadas a Chiron; US 5.972.885 y US 6.048.720 asignadas a Biovitrum, US 5.543.502 y US 5.610.278 asignadas a Novo Nordisk; US 5.171.844 asignada a Immuno Ag; US 5.112.950 asignada a Transgene S.A.; US 4.868.112 asignada a Genetics Institute.

## 10 Optimización por codones

En una realización, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de ácido nucleico ha sido optimizada por codones. En otra realización, la secuencia de ácido nucleico de partida que codifica un polipéptido con actividad FVIII y que está sujeta a optimización por codones es SEQ ID NO: 16. En algunas realizaciones, la secuencia que codifica un polipéptido con actividad FVIII está optimizada por codones para la expresión humana. En otras realizaciones, la secuencia que codifica un polipéptido con actividad de FVIII está optimizada por codones para la expresión murina. SEQ ID NOs: 1-6, 70 y 71 son versiones optimizadas por codones de SEQ ID NO: 16, optimizadas para la expresión humana.

La expresión "optimizado por codones" cuando se refiere a genes o regiones codificantes de moléculas de ácido nucleico para la transformación de diversos huéspedes, se refiere a la alteración de codones en el gen o las regiones codificantes de las moléculas de ácido nucleico para reflejar el uso del codón típico del organismo huésped sin alterar el polipéptido codificado por el ADN. Una optimización de este tipo incluye reemplazar al menos uno, o más de uno, o un número significativo de codones con uno o más codones que se utilizan con mayor frecuencia en los genes de ese organismo.

Desviaciones en la secuencia de nucleótidos que comprende los codones que codifican los aminoácidos de cualquier cadena polipeptídica permiten variaciones en la secuencia que codifica el gen. Dado que cada uno de los codones consiste en tres nucleótidos, y los nucleótidos que comprenden el ADN están restringidos a cuatro bases específicas, existen 64 combinaciones posibles de nucleótidos, 61 de las cuales codifican aminoácidos (los tres codones restantes codifican señales que terminan la traducción). El "código genético" que muestra qué codones codifican qué aminoácidos se reproduce en esta memoria como la Tabla 1. Como resultado, muchos aminoácidos están designados por más de un codón. Por ejemplo, los aminoácidos alanina y prolina están codificados por cuatro tripletes, la serina y la arginina por seis, mientras que el triptófano y la metionina están codificados por un solo triplete. Esta degeneración permite que la composición de las bases del ADN varíe en un amplio intervalo sin alterar la secuencia de aminoácidos de las proteínas codificadas por el ADN.

35 **Tabla 1:** El Código Genético Estándar

	<b>T</b>	<b>C</b>	<b>A</b>	<b>G</b>
<b>T</b>	TTT Phe (F) TTC " TTA Leu (L) TTG"	TCT Ser (S) TCC " TCA" TCG"	TAT Tyr (Y) TAC " TAA <b>Parada</b> TAG <b>Parada</b>	TGT Cys (C) TGC TGA <b>Parada</b> TGG Trp (W)
<b>C</b>	CTT Leu (L) CTC " CTA" CTG"	CCT Pro (P) CCC " CAA " CCG"	CAT His (H) CAC" CAA Gln (Q) CAG"	CGT Arg (R) CGC" CGA" CGG"
<b>A</b>	ATT Ile (I) ATC " ATA" ATG Met (M)	ACT Thr (T) ACC" ACA" ACG "	Asn AAT (N) AAC" AAA Lys (K) AAG "	AGT Ser (S) AGC" AGA Arg (R) AGG"
<b>G</b>	GTT Val (V) GTC " GTA" GTG"	GCT Ala (A) GCC " GCA " GCG "	GAT Asp (D) GAC " GAA Glu (E) GAG "	GGT Gly (G) GGC " GGA" GGG"

Muchos organismos muestran un sesgo por el uso de codones particulares para codificar la inserción de un aminoácido particular en una cadena peptídica en crecimiento. La preferencia de codones, o sesgo de codones, las diferencias en el uso de codones entre organismos, se debe a la degeneración del código genético y está bien documentada entre muchos organismos. El sesgo de codones a menudo se correlaciona con la eficiencia de la traducción del ARN mensajero (ARNm),

que a su vez se cree que depende, *entre otras cosas*, de las propiedades de los codones que se traducen y de la disponibilidad de moléculas de ARN de transferencia (ARNt) particulares. El predominio de ARNts seleccionados en una célula es generalmente un reflejo de los codones que se usan con mayor frecuencia en la síntesis de péptidos. Por consiguiente, los genes se pueden adaptar para una expresión génica óptima en un organismo dado en función de la optimización por codones.

Dado el gran número de secuencias de genes disponibles para una amplia diversidad de especies animales, vegetales y microbianas, se han calculado las frecuencias relativas del uso de codones. Las tablas de uso de codones están disponibles, por ejemplo, en "Codon Usage Database" disponible en [www.kazusa.or.jp/codon/](http://www.kazusa.or.jp/codon/) (visitada el 18 de junio de 2012). Véase Nakamura, Y., et al. Nucl. Acids Res. 28:292 (2000).

La asignación aleatoria de codones a una frecuencia optimizada para codificar una secuencia polipeptídica dada puede realizarse manualmente calculando las frecuencias de los codones para cada uno de los aminoácidos y luego asignando aleatoriamente los codones a la secuencia polipeptídica. Adicionalmente, se pueden utilizar diversos algoritmos y programas de software informáticos para calcular una secuencia óptima.

En algunas realizaciones, la molécula de ácido nucleico comprende una o más propiedades: (a) la molécula de ácido nucleico o una porción de la misma tiene un índice de adaptación de codones humanos aumentado en relación con SEQ ID NO: 16; (b) la secuencia de nucleótidos o una parte de la misma tiene una frecuencia incrementada de codones óptimos en relación con la SEQ ID NO: 16; (c) la secuencia de nucleótidos o una porción de la misma contiene un mayor porcentaje de nucleótidos G/C en comparación con el porcentaje de nucleótidos G/C en SEQ ID NO: 16; (d) la secuencia de nucleótidos o una porción de la misma tiene un uso de codones sinónimo relativo incrementado en relación con SEQ ID NO: 16; (e) la secuencia de nucleótidos o una porción de la misma es un número efectivo reducido de codones relativos a SEQ ID NO: 16; (f) la secuencia de nucleótidos contiene menos secuencias MARS/ARS (SEQ ID NOs: 21 y 22) en relación con la SEQ ID NO: 16; (g) la secuencia de nucleótidos contiene menos elementos desestabilizadores (SEQ ID NOs: 23 y 24) en relación con SEQ ID NO: 16; (i) la secuencia de nucleótidos no contiene una secuencia poli-T, (j) la secuencia de nucleótidos no contiene una secuencia poli-A; o (k) cualquier combinación de las mismas. En algunas realizaciones, las moléculas de ácido nucleico contienen al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve o diez características de (a) a (j).

### Índice de Adaptación de Codones

En una realización, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en esta memoria que codifica un polipéptido con actividad de FVIII, en donde el índice de adaptación de codones humanos aumenta en relación con SEQ ID NO: 16. Por ejemplo, la secuencia de nucleótidos puede tener un índice de adaptación de codones humanos que es al menos aproximadamente de 0,75 (75 %), al menos aproximadamente de 0,76 (76 %), al menos aproximadamente de 0,77 (77 %), al menos aproximadamente de 0,78 (78 %), al menos aproximadamente 0,79 (79 %), al menos aproximadamente 0,80 (80 %), al menos aproximadamente 0,81 (81 %), al menos aproximadamente 0,82 (82 %), al menos aproximadamente 0,83 (83 %), al menos aproximadamente 0,84 (84 %), al menos aproximadamente 0,85 (85 %), al menos aproximadamente 0,86 (86 %), al menos aproximadamente 0,87 (87 %), al menos aproximadamente 0,88 (88 %), al menos aproximadamente 0,89 (89 %), al menos aproximadamente 0,90 (90 %), al menos aproximadamente 0,91 (91 %), al menos aproximadamente 0,92 (92 %), al menos aproximadamente 0,93 (93 %), al menos aproximadamente 0,94 (94 %), al menos aproximadamente 0,95 (95 %), al menos aproximadamente 0,96 (96 %), al menos aproximadamente 0,97 (97 %), al menos aproximadamente 0,98 (98 %) o al menos aproximadamente 0,99 (99 %). En algunas realizaciones, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,88 (88 %). En otras realizaciones, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,91 (91 %). En otras realizaciones, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,91 (97 %).

En otras realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde el índice de adaptación de codones humanos de la secuencia de nucleótidos aumenta con respecto a SEQ ID NO: 16. En algunas realizaciones, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,75 (75 %), al menos aproximadamente 0,76 (76 %), al menos aproximadamente 0,77 (77 %), al menos aproximadamente 0,78 (78 %), al menos aproximadamente 0,79 (79 %), al menos aproximadamente 0,80 (80 %), al menos aproximadamente 0,81 (81 %), al menos aproximadamente 0,82 (82 %), al menos aproximadamente 0,83 (83 %), al menos aproximadamente 0,84 (84 %), al menos aproximadamente 0,85 (85 %), al menos aproximadamente 0,86 (86 %), al menos aproximadamente 0,87 (87 %) o al menos aproximadamente 0,88 (88 %). En una realización particular, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,75 (75 %). En otra realización, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,83 (83 %). En otra realización, la secuencia de nucleótidos tiene un índice de adaptación de codones

humanos que es al menos aproximadamente 0,88 (88 %). En otra realización, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,91 (91 %). En otra realización, la secuencia de nucleótidos tiene un índice de adaptación de codones humanos que es al menos aproximadamente 0,97 (97 %).

En algunas realizaciones, la molécula de ácido nucleico aislada de la presente divulgación tiene una frecuencia incrementada de codones óptimos (FOP, por sus siglas en inglés) en relación con SEQ ID NO: 16. En determinadas realizaciones, la FOP de la molécula de ácido nucleico aislada es al menos aproximadamente 40, al menos aproximadamente 45, al menos aproximadamente 50, al menos aproximadamente 55, al menos aproximadamente 60, al menos aproximadamente 64, al menos aproximadamente 65, al menos aproximadamente 70, al menos aproximadamente 75, al menos aproximadamente 79, al menos aproximadamente 80, al menos alrededor de 85 o al menos aproximadamente 90.

En otras realizaciones, la molécula de ácido nucleico aislada de la presente divulgación tiene un uso relativo de codones sinónimos (RCSU, por sus siglas en inglés) incrementado en relación con SEQ ID NO: 16. En algunas realizaciones, el RCSU de la molécula de ácido nucleico aislada es mayor que 1,5. En otras realizaciones, el RCSU de la molécula de ácido nucleico aislada es superior a 2,0. En determinadas realizaciones, el RCSU de la molécula de ácido nucleico aislada es al menos aproximadamente 1,5, al menos aproximadamente 1,6, al menos aproximadamente 1,7, al menos aproximadamente 1,8, al menos aproximadamente 1,9, al menos aproximadamente 2,0, al menos aproximadamente 2,1, al menos aproximadamente 2,2, al menos aproximadamente 2,3, al menos aproximadamente 2,4, al menos aproximadamente 2,5, al menos aproximadamente 2,6 o al menos aproximadamente 2,7.

Aún en otras realizaciones, la molécula de ácido nucleico aislada de la presente divulgación tiene un número efectivo reducido de codones con respecto a SEQ ID NO: 16. En algunas realizaciones, la molécula de ácido nucleico aislada tiene un número efectivo de codones de menos de aproximadamente 50, menos de aproximadamente 45, menos de aproximadamente 40, menos de aproximadamente 35, menos de aproximadamente 30 o menos de aproximadamente 25. En una realización particular, la molécula de ácido nucleico aislada tiene un número efectivo de codones de aproximadamente 40, aproximadamente 35, aproximadamente 30, aproximadamente 25 o aproximadamente 20.

## Optimización del Contenido de G/C

En algunas realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en esta memoria que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos contiene un porcentaje mayor de nucleótidos G/C en comparación con el porcentaje de nucleótidos G/C en SEQ ID NO: 16. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad FVIII tiene un contenido de G/C que es al menos aproximadamente 45 %, al menos aproximadamente 46 %, al menos aproximadamente 47 %, al menos aproximadamente 48 %, al menos aproximadamente 49 %, al menos aproximadamente 50 %, al menos aproximadamente 51 %, al menos aproximadamente 52 %, al menos aproximadamente 53 %, al menos aproximadamente 54 %, al menos aproximadamente 55 %, al menos aproximadamente 56 %, al menos aproximadamente 57 %, al menos aproximadamente 58 %, al menos aproximadamente 59 % o al menos aproximadamente 60%.

En otras realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos contiene un mayor porcentaje de nucleótidos G/C en comparación con el porcentaje de nucleótidos G/C en SEQ ID NO: 16. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad FVIII tiene un contenido de G/C que es al menos aproximadamente 45 %. En una realización particular, la secuencia de nucleótidos que codifica un polipéptido con actividad FVIII tiene un contenido de G/C de al menos aproximadamente 52 %. En otra realización, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 55 %. En otra realización, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 57 %. En otra realización, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C de al menos aproximadamente 58 %. En aún otra realización, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII tiene un contenido de G/C que es al menos aproximadamente 60 %.

El "contenido de G/C" (o contenido de guanina-citosina) o "porcentaje de nucleótidos de G/C" se refiere al porcentaje de bases nitrogenadas en una molécula de ADN que son guanina o citosina. El contenido de G/C se puede calcular utilizando la siguiente fórmula:

$$\frac{G+C}{A+T+G+C} \times 100 \quad (III)$$

Los genes humanos son muy heterogéneos en su contenido de G/C, teniendo algunos genes un contenido de G/C tan bajo como el 20 % y teniendo otros genes un contenido de G/C tan alto como el 95 %. En general, los genes ricos en G/C se expresan más. De hecho, se ha demostrado que aumentar el contenido de G/C de un gen puede conducir a una expresión incrementada del gen, debido principalmente a un aumento en la transcripción y a niveles más altos de ARNm en estado estacionario. Véase Kudla et al., PLoS Biol., 4(6): e180 (2006).

#### **Secuencias Similares a Regiones de Unión a Matriz**

En algunas realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en esta memoria que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos contiene menos secuencias MARS/ARS en relación con SEQ ID NO: 16. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad FVIII contiene a lo sumo 6, a lo sumo 5, a lo sumo 4, a lo sumo 3 o a lo sumo 2 secuencias MARS/ARS. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad FVIII contiene a lo sumo 1 secuencia MARS/ARS. En aún otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene una secuencia MARS/ARS.

En otras realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos contiene menos secuencias MARS/ARS en relación con SEQ ID NO: 16. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad FVIII contiene a lo sumo 6, a lo sumo 5, a lo sumo 4, a lo sumo 3 o a lo sumo 2 secuencias MARS/ARS. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad FVIII contiene a lo sumo 1 secuencia MARS/ARS. En aún otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene una secuencia MARS/ARS.

Se han identificado elementos ricos en AT en la secuencia de nucleótidos del FVIII humano que comparten similitud de secuencia con las secuencias de replicación autónoma (ARS, por sus siglas en inglés) y las regiones de unión a la matriz nuclear (MARs, por sus siglas en inglés) de *Saccharomyces cerevisiae*. (Falux et al., Mol. Cell. Biol. 16:4264-4272 (1996). Se ha demostrado que uno de estos elementos se une a factores nucleares *in vitro* y reprime la expresión de un gen indicador de cloranfenicol acetiltransferasa (CAT). *Id.* Se ha planteado la hipótesis de que estas secuencias pueden contribuir a la represión transcripcional del gen FVIII humano. Por lo tanto, en una realización, todas las secuencias MAR/ARS están abolidas en el gen FVIII de la presente divulgación. Hay cuatro secuencias MAR/ARS ATATTT (SEQ ID NO: 21) y tres secuencias MAR/ARS AAATAT (SEQ ID NO: 22) en la secuencia parental de FVIII (SEQ ID NO: 16). Todos estos sitios fueron mutados para destruir las secuencias MAR/ARS en las secuencias optimizada de FVIII (SEQ ID NOs: 1-6) La ubicación de cada uno de estos elementos y la secuencia de los nucleótidos correspondientes en las secuencias optimizadas se muestran en la Tabla 2, que figura a continuación.

Tabla 2: Resumen de Cambios a Elementos Represivos

Ubicación del Elemento	Inicio de la secuencia BDD FVIII (SEQ ID NO: 16)	Secuencia BDD FVIII optimizada								
		SEQ ID NO: 1	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 4	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 70	SEQ ID NO: 71	
Secuencias Destabilizadoras										
639	ATTTA	GTTTA	GTTCa	GTTCa	GTTCa	GTTCa	GTTCa	GTTCa	GTTCa	GTTCA
1338	ATTTA	GTTTA	GTTCa	GTTCa	CTTCA	GTTCa	GTTCa	GTTCa	GTTCa	GTTCA
1449	ATTTA	CTTTA	CTTCA	CTTCA	CTTCA	CTTCA	CTTCA	CTTCA	CTTCA	CTTCA
1590	TAAAT	TAAAT	CAAGT	CAAGT	TAAAGT	CAAGT	CAAGT	CAAGT	CAAGT	TAAAGT
1623	TAAAT	AAAA	GAGÁ	CTAAG	CAAGA	CAAGA	CAAGA	TAAAGT	CAAGA	CAAGA
2410	ATTTA	ATCTA	ATCTA	ATCTA	ATCTA	ATCTA	ATCTA	ATCTA	ATCTA	ATCTA
2586	ATTTA	GTTTA	GTTCa	GTTCa	GTTCa	GTTCa	GTTCa	GTTCa	GTTCa	GTTCA
2630	TAAT	TGAAT	TGAAC	TGAAC	TGAAC	TCAAT	TGAAC	TCAAT	TGAAC	TGAAC
3884	ATTTA	ATCTG	ACCTG	ACCTG	ACCTG	ATCTG	ACCTG	ATCTG	ACCTG	ACCTG
3887	TAAAT	TGAAC	TGAAC	TGAAC	TGAAC	TGAAC	TGAAC	TGAAC	TGAAC	TGAAC
Sitios de Unión al Promotor Potenciales										
641	TTATA	TTATC	TCATC	TCATT	TCATC	TCATC	TCATC	TCATT	TCATC	
1275	TATAA	CTATA	TTACA	CTACA	GTACA	CTACA	CTACA	GTACA	GTACA	
1276	TTATA	TATAA	TACAA	TACAA	TACAA	TACAA	TACAA	TACAA	TACAA	
1445	TTATA	TCATC	TCATC	TTATC	TCATC	TCATC	TCATC	TTATC	TCATC	
1474	TATAA	TATAA	TACAA	TACAA	TACAA	TACAA	TACAA	TACAA	TACAA	
1588	TATAA	TATAA	TACAA	TACAA	TATAA	TACAA	TACAA	TACAA	TATAA	
2614	TTATA	CTGTA	CTGTA	CTGTA	CTGTA	TTGTA	CTGTA	CTGTA	CTGTA	
2661	TATAA	CATCA	CATCA	CATCA	CATCA	CATCA	CATCC	CATCA	CATCC	
3286	TATAA	TATAA	TACAA	TACAA	TACAA	TACAA	TACAA	TACAA	TACAA	
3840	TTATA	TTACT	CTACA	CTACA	CTACA	CTACT	CTACT	CTACA	CTACT	



(continúa)

Secuencias Similares de Unión a Matriz (MARS/ARS)											
1287	ATATTT	GTATCT	GTACCT	GTATCT	GTACCT	GTACCT	GTACCT	GTACCT	GTACCT	GTATCT	GTATCT
1447	ATATTT	ATCTTC	ATCTTC	ATCTTC	ATCTTC	ATCTTC	ATCTTC	ATCTTC	ATCTTC	ATCTTC	ATCTTC
1577	AAATAT	AAATCT	AAATCT	AAATCT	AAATCT	AAATCT	AAATCT	AAATCT	AAATCT	AAATCT	AAATCT
1585	AAATAT	AAGTAT	AAGTAC	AAGTAT	AAGTAC	AAGTAC	AAGTAC	AAGTAC	AAGTAC	AAGTAT	AAGTAT
2231	ATATTT	ACATCA	ACATCA	ACATCA	ACATCT	ACATCT	ACATCT	ACATCT	ACATCT	ATAQUE	ATAQUE
3054	AAATAT	AAACAT	GAACAT	GAACAT	GAACAT	GAACAT	GAACAT	GAACAT	GAACAT	GAATAT	GAATAT
3788	ATATTT	ATAQUE	ACATCT	ACATCT	ACATCT	ACATCT	ACATCT	ACATCT	ACATCT	ACATCT	ACATCT
Elementos de Secuencia Ricos en AU (AREs)											
2468	1. ATT ATT	ACTCATC	ACTTCAT	ACTTCAT	ACTTCAT	ACTTCAT	ACTTCAT	ACTTCAT	ACTTCAT	ACTTCAT	ACTTCAT
3790	2. ATTTT TAA	ATCTTTAA	A TCTTCAA	ATCTTCAA	ATCTTCAA	ATCTTCAA	ATCTTCAA	ATCTTCAA	A TCTTCAA	A TCTTCAA	A TCTTCAA
Secuencias de PoliA/Poli T											
3273	AAAAAA	GAAGAA	GAAGAA	GAAGAA	GAAGAA	GAAGAA	GAAGAA	GAAGAA	GAAGAA	GAAGAA	GAAGAA
4195	TTTTTT	TTCTTT	TTCTTC	TTCTTC	TTCTTC	TTCTTC	TTCTTC	TTCTTC	TTCTTC	TTCTTC	TTCTTC
Sitios de Corte y Empalme											
2203	GGTGAT	GGGGAC	GGCGAC	GGGGAC	GGGGAC	GGGGAC	GGGGAC	GGGGAC	GGGGAC	GGGGAC	GGGGAC

### Secuencias Desestabilizadoras

En algunas realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en esta memoria que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos contiene menos elementos desestabilizadores en relación con SEQ ID NO: 16. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene a lo sumo 9, a lo sumo 8, a lo sumo 7, a lo sumo 6 o a lo sumo 5 elementos desestabilizadores. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene a lo sumo 4, a lo sumo 3, a lo sumo 2 o a lo sumo 1 elemento desestabilizador. En aún otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un elemento desestabilizador.

En otras realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con (i) los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, los nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos contiene menos elementos desestabilizadores en relación con SEQ ID NO: 16. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene a lo sumo 9, a lo sumo 8, a lo sumo 7, a lo sumo 6 o a lo sumo 5 elementos desestabilizadores. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene a lo sumo 4, a lo sumo 3, a lo sumo 2 o a lo sumo 1 elemento desestabilizador. En aún otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un elemento desestabilizador.

Existen diez elementos desestabilizadores en la secuencia parental de FVIII (SEQ ID NO: 16): seis secuencias ATTTA (SEQ ID NO: 23) y cuatro secuencias TAAAT (SEQ ID NO: 24). En una realización, las secuencias de estos sitios se mutaron para destruir los elementos desestabilizadores en FVIII optimizado SEQ ID NO: 71. La ubicación de cada uno de estos elementos y la secuencia de los nucleótidos correspondientes en las secuencias optimizadas se muestran en la Tabla 2.

### Sitios de Unión de Promotores Potenciales

En algunas realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos descrita en esta memoria que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos contiene menos sitios potenciales de unión al promotor en relación con SEQ ID NO: 16. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene a lo sumo 9, a lo sumo 8, a lo sumo 7, a lo sumo 6 o a lo sumo 5 sitios potenciales de unión al promotor. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene a lo sumo 4, a lo sumo 3, a lo sumo 2 o a lo sumo 1 sitio de unión al promotor potencial. En aún otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un sitio potencial de unión al promotor.

En otras realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos contiene menos sitios de unión a promotores potenciales en relación con SEQ ID NO: 16. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene a lo sumo 9, a lo sumo 8, a lo sumo 7, a lo sumo 6 o a lo sumo 5 sitios potenciales de unión al promotor. En otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII contiene a lo sumo 4, a lo sumo 3, a lo sumo 2 o a lo sumo 1 sitio de unión al promotor potencial. En aún otras realizaciones, la secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII no contiene un sitio potencial de unión al promotor.

Cajas TATA son secuencias reguladoras que se encuentran a menudo en las regiones del promotor de eucariotas. Sirven como el sitio de unión de la proteína de unión TATA (TBP, por sus siglas en inglés), un factor de transcripción general. Las cajas TATA comprenden habitualmente la secuencia TATAA (SEQ ID NO: 28) o una variante cercana. Sin embargo, las cajas TATA dentro de una secuencia codificante pueden inhibir la traducción de la proteína de longitud completa. Existen diez secuencias de unión a promotores potenciales en la secuencia de BDD FVIII de tipo salvaje (SEQ ID NO: 16): cinco secuencias TATAA (SEQ ID NO: 28) y cinco secuencias TTATA (SEQ ID NO: 29). En algunas realizaciones, al menos 1, al menos 2, al menos 3 o al menos 4 de los sitios de unión al promotor están abolidos en los genes FVIII de la presente divulgación. En algunas realizaciones, al menos 5 de los sitios de unión al promotor están abolidos en los genes FVIII de la presente divulgación. En otras realizaciones, al menos 6, al menos 7 o al menos 8 de los sitios de unión al promotor están abolidos en los genes FVIII de la presente divulgación. En una realización, al menos 9 de los sitios de unión al

promotor están abolidos en los genes FVIII de la presente divulgación. En una realización particular, todos los sitios de unión al promotor están abolidos en los genes FVIII de la presente divulgación. La ubicación de cada uno de los sitios de unión potencial al promotor y la secuencia de los nucleótidos correspondientes en las secuencias optimizadas se muestran en la Tabla 2.

## 5 Otros Elementos Reguladores Negativos que Actúan en Cis

Además de las secuencias MAR/ARS, elementos desestabilizadores y sitios de promotor potenciales arriba descritos, pueden identificarse varias secuencias potencialmente inhibitoras adicionales en la secuencia de BDD FVIII de tipo salvaje (SEQ ID NO: 16). Se pueden identificar dos elementos de secuencia ricos en AU (AREs, por sus siglas en inglés) (ATTTTATT (SEQ ID NO: 30); y ATTTTTAA (SEQ ID NO: 31), junto con un sitio poli-A (AAAAAAA; SEQ ID NO: 26), un sitio poli-T (TTTTTT; SEQ ID NO: 25) y un sitio de corte y empalme (GGTGAT; SEQ ID NO: 27) en la secuencia de BDD FVIII no optimizada. Uno o más de estos elementos se puede eliminar de las secuencias de FVIII optimizadas. La ubicación de cada uno de estos sitios y la secuencia de los nucleótidos correspondientes en las secuencias optimizadas se muestran en la Tabla 2.

En otras realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos no contiene uno o más elementos reguladores negativos que actúan en cis, por ejemplo, un sitio de corte y empalme, una secuencia poli-T, una secuencia poli-A, una secuencia ARE o cualquier combinación de los mismos.

En algunas realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos no contiene el sitio de corte y empalme GGTGAT (SEQ ID NO: 27). En algunas realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos no contiene una secuencia poli-T (SEQ ID NO: 25). En algunas realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos no contiene una secuencia poli-A (SEQ ID NO: 26). En algunas realizaciones, la presente divulgación proporciona una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 9 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 (es decir, nucleótidos 58-4374 de SEQ ID NO: 71 sin los nucleótidos que codifican el dominio B o fragmento de dominio B); y en donde la secuencia de nucleótidos no contiene un elemento ARE (SEQ ID NO: 30 o SEQ ID NO: 31).

En otras realizaciones, una secuencia de FVIII optimizada de la divulgación no comprende uno o más motivos antivirales, estructuras de tallo-bucle y secuencias repetitivas.

En aún otras realizaciones, los nucleótidos que rodean el sitio de inicio de la transcripción se cambian a una secuencia de consenso kozak (GCCGCCACC ATGC (SEQ ID NO: 32), en donde los nucleótidos subrayados son el codón de inicio). En otras realizaciones, los sitios de restricción pueden añadirse o eliminarse para facilitar el proceso de clonación.

## 50 Secuencias de Nucleótidos Heterólogos

En algunas realizaciones, las moléculas de ácido nucleico aisladas de la divulgación comprenden, además, una secuencia de nucleótidos heteróloga. En algunas realizaciones, las moléculas de ácido nucleico aisladas de la divulgación comprenden, además, al menos una secuencia de nucleótidos heteróloga. La secuencia de nucleótidos heteróloga se puede enlazar con las secuencias de nucleótidos de BDD-FVIII optimizadas de la divulgación en el extremo 5', en el extremo 3' o insertarse en el medio de la secuencia de nucleótidos de BDD-FVIII optimizada. Por lo tanto, en algunas

realizaciones, la secuencia de aminoácidos heteróloga codificada por la secuencia de nucleótidos heteróloga está enlazada al extremo N o al extremo C de la secuencia de aminoácidos de FVIII codificada por la secuencia de nucleótidos o está insertada entre dos aminoácidos en la secuencia de aminoácidos de FVIII. En algunas realizaciones, la secuencia de aminoácidos heteróloga se puede insertar entre dos aminoácidos en uno o más sitios de inserción seleccionados de la Tabla 3. En algunas realizaciones, la secuencia de aminoácidos heteróloga se puede insertar dentro del polipéptido FVIII codificado por la molécula de ácido nucleico de la divulgación en cualquier sitio descrito en la Publicación Internacional N° WO 2013/123457 A1, WO 2015/106052 A1 o la Publicación de EE.UU. N° 2015/0158929 A1.

En algunas realizaciones, la secuencia de aminoácidos heteróloga codificada por la secuencia de nucleótidos heteróloga se inserta dentro del dominio B o un fragmento del mismo. En algunas realizaciones, la secuencia de aminoácidos heteróloga se inserta dentro del FVIII inmediatamente aguas abajo de un aminoácido correspondiente al aminoácido 745 del FVIII humano maduro (SEQ ID NO: 15). En una realización particular, el FVIII comprende una delección de los aminoácidos 746-1646, correspondientes al FVIII humano maduro (SEQ ID NO : 15), y la secuencia de aminoácidos heteróloga codificada por la secuencia de nucleótidos heteróloga se inserta inmediatamente aguas abajo del aminoácido 745, correspondiente al FVIII humano maduro (SEQ ID NO:15).

**TABLA 3:** Sitios de Inserción de Restos Heterólogos

Sitio de inserción	de Dominio	Sitio de inserción	de Dominio	Sitio de inserción	de Dominio
3	A1	375	A2	1749	A3
18	A1	378	A2	1796	A3
22	A1	399	A2	1802	A3
26	A1	403	A2	1827	A3
40	A1	409	A2	1861	A3
60	A1	416	A2	1896	A3
65	A1	442	A2	1900	A3
81	A1	487	A2	1904	A3
116	A1	490	A2	1905	A3
119	A1	494	A2	1910	A3
130	A1	500	A2	1937	A3
188	A1	518	A2	2019	A3
211	A1	599	A2	2068	C1
216	A1	603	A2	2111	C1
220	A1	713	A2	2120	C1
224	A1	745	B	2171	C2
230	A1	1656	región a3	2188	C2
333	A1	1711	A3	2227	C2
336	A1	1720	A3	2332	CT
339	A1	1725	A3		
<b>Nota:</b> Los sitios de inserción indican la posición de aminoácido correspondiente a una posición de aminoácido de FVIII humano maduro (SEQ ID NO: 15).					

En otras realizaciones, las moléculas de ácido nucleico aisladas de la divulgación comprenden, además, dos, tres, cuatro, cinco, seis, siete u ocho secuencias de nucleótidos heterólogas. En algunas realizaciones, todas las secuencias de nucleótidos heterólogas son idénticas. En algunas realizaciones, al menos una secuencia de nucleótidos heteróloga es diferente de las otras secuencias de nucleótidos heterólogas. En algunas realizaciones, la divulgación puede comprender dos, tres, cuatro, cinco, seis o más de siete secuencias de nucleótidos heterólogas en tándem.

En algunas realizaciones, la secuencia de nucleótidos heteróloga codifica una secuencia de aminoácidos. En algunas realizaciones, la secuencia de aminoácidos codificada por la secuencia de nucleótidos heteróloga es un resto heterólogo que puede aumentar la semivida (un "extensor de la semivida") de una molécula de FVIII.

- En algunas realizaciones, el resto heterólogo es un péptido o un polipéptido con características estructuradas o no estructuradas que están asociadas con la prolongación de la semivida *in vivo* cuando se incorporan en una proteína de la divulgación. Ejemplos no limitantes incluyen albúmina, fragmentos de albúmina, fragmentos Fc de inmunoglobulinas, el péptido C-terminal (CTP, por sus siglas en inglés) de la subunidad  $\beta$  de la gonadotropina coriónica humana, una secuencia HAP, una secuencia XTEN, una transferrina o un fragmento de la misma, un polipéptido PAS, enlazadores de poliglicina, enlazadores de poliserina, restos de unión a albúmina o cualesquiera fragmentos, derivados, variantes o combinaciones de estos polipéptidos. En una realización particular, la secuencia de aminoácidos heteróloga es una región constante de inmunoglobulina o una porción de la misma, transferrina, albúmina o una secuencia PAS. En algunos aspectos, un resto heterólogo incluye factor de von Willebrand o un fragmento del mismo. En otros aspectos relacionados, un resto heterólogo puede incluir un sitio de unión (p. ej., un aminoácido de cisteína) para un resto no polipeptídico tal como polietilenglicol (PEG), hidroxietil almidón (HES, por sus siglas en inglés), ácido polisialico o cualesquiera derivados, variantes o combinaciones de estos elementos. En algunos aspectos, un resto heterólogo comprende un aminoácido de cisteína que funciona como un sitio de fijación para un resto no polipeptídico tal como polietilenglicol (PEG), hidroxietilalmidón (HES), ácido polisialico o cualesquiera derivados, variantes o combinaciones de estos elementos
- En una realización específica, una primera secuencia de nucleótidos heteróloga codifica un primer resto heterólogo que es una molécula que prolonga la semivida que se conoce en la técnica, y una segunda secuencia de nucleótidos heteróloga codifica un segundo resto heterólogo que también puede ser una molécula que prolonga la semivida que se conoce en la técnica. En determinadas realizaciones, el primer resto heterólogo (p. ej., un primer resto Fc) y el segundo resto heterólogo (p. ej., un segundo resto Fc) están asociados entre sí para formar un dímero. En una realización, el segundo resto heterólogo es un segundo resto Fc, en donde el segundo resto Fc está enlazado a o asociado con el primer resto heterólogo, p. ej., el primer resto Fc. Por ejemplo, el segundo resto heterólogo (p. ej., el segundo resto Fc) puede enlazarse al primer resto heterólogo (p. ej., el primer resto Fc) mediante un enlazador o puede asociarse con el primer resto heterólogo mediante un enlace covalente o no covalente.
- En algunas realizaciones, el resto heterólogo es un polipéptido que comprende, consiste esencialmente en o consiste en al menos aproximadamente 10, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 300, al menos aproximadamente 400, al menos aproximadamente 500, al menos aproximadamente 600, al menos aproximadamente 700, al menos aproximadamente 800, al menos aproximadamente 900, al menos aproximadamente 1000, al menos aproximadamente 1100, al menos aproximadamente 1200, al menos aproximadamente 1300, al menos aproximadamente 1400, al menos aproximadamente 1500, al menos aproximadamente 1600, al menos aproximadamente 1700, al menos aproximadamente 1800, al menos aproximadamente 1900, al menos aproximadamente 2000, al menos aproximadamente 2500, al menos aproximadamente 3000 o al menos aproximadamente 4000 aminoácidos. En otras realizaciones, el resto heterólogo es un polipéptido que comprende, consiste esencialmente en o consiste en aproximadamente 100 a aproximadamente 200 aminoácidos, aproximadamente 200 a aproximadamente 300 aminoácidos, aproximadamente 300 a aproximadamente 400 aminoácidos, aproximadamente 400 a aproximadamente 500 aminoácidos, aproximadamente 500 a aproximadamente 600 aminoácidos, aproximadamente 600 a aproximadamente 700 aminoácidos, aproximadamente 700 a aproximadamente 800 aminoácidos, aproximadamente 800 a aproximadamente 900 aminoácidos o aproximadamente 900 a aproximadamente 1000 aminoácidos.
- En determinadas realizaciones, un resto heterólogo mejora una o más propiedades farmacocinéticas de la proteína FVIII sin afectar significativamente a su actividad o función biológica.
- En determinadas realizaciones, un resto heterólogo aumenta la semivida *in vivo* y/o *in vitro* de la proteína FVIII de la divulgación. En otras realizaciones, un resto heterólogo facilita la visualización o localización de la proteína FVIII de la divulgación o un fragmento de la misma (p. ej., un fragmento que comprende un resto heterólogo después de la escisión proteolítica de la proteína FVIII). La visualización y/o ubicación de la proteína FVIII de la divulgación o un fragmento de la misma puede ser *in vivo*, *in vitro*, *ex vivo* o combinaciones de los mismos.
- En otras realizaciones, un resto heterólogo aumenta la estabilidad de la proteína FVIII de la divulgación o un fragmento de la misma (p. ej., un fragmento que comprende un resto heterólogo después de la escisión proteolítica de la proteína FVIII). Tal como se utiliza en esta memoria, el término "estabilidad" se refiere a una medida reconocida en la técnica del mantenimiento de una o más propiedades físicas de la proteína FVIII en respuesta a una condición del entorno (p. ej., una temperatura elevada o reducida). En determinados aspectos, la propiedad física puede ser el mantenimiento de la estructura covalente de la proteína FVIII (p. ej., la ausencia de escisión proteolítica, oxidación o desamidación no deseada). En otros aspectos, la propiedad física también puede ser la presencia de la proteína FVIII en un estado correctamente plegado (p. ej., la ausencia de agregados o precipitados solubles o insolubles). En un aspecto, la estabilidad de la proteína FVIII se mide ensayando una propiedad biofísica de la proteína FVIII, por ejemplo, estabilidad térmica, perfil de despliegue del pH, eliminación estable de glicosilación, solubilidad, función bioquímica (p. ej., capacidad para unirse a una proteína, receptor o ligando), etc., y/o combinaciones de los mismos. En otro aspecto, la función bioquímica se demuestra por la afinidad de unión de la interacción. En un aspecto, una medida de la estabilidad de la proteína es la estabilidad térmica, es decir, la resistencia al desafío térmico. La estabilidad se puede medir utilizando métodos conocidos en la técnica, tales como HPLC (cromatografía líquida de alta resolución), SEC (cromatografía de exclusión por tamaño), DLS (dispersión dinámica de la luz), etc. Métodos para medir la estabilidad térmica incluyen, pero no se limitan a calorimetría diferencial de barrido (DSC), fluorimetría diferencial de barrido (DSF), dicroísmo circular (CD) y ensayo de desafío térmico.

En determinados aspectos, una proteína FVIII codificada por la molécula de ácido nucleico de la divulgación comprende al menos un prolongador de la semivida, *es decir*, un resto heterólogo que aumenta la semivida *in vivo* de la proteína FVIII con respecto a la semivida *in vivo* de la proteína FVIII correspondiente que carece de un resto heterólogo de este tipo. La semivida *in vivo* de una proteína FVIII puede determinarse mediante cualquier método conocido por los expertos en la técnica, *p. ej.*, ensayos de actividad (ensayo cromogénico o ensayo de aPTT de coagulación en una fase), ELISA, ROTEM™, etc.

En algunas realizaciones, la presencia de uno o más prolongadores de la semivida da como resultado que la semivida de la proteína FVIII aumente en comparación con la semivida de la proteína correspondiente que carece de uno o más prolongadores de la semivida de este tipo. La semivida de la proteína FVIII que comprende un prolongador de la semivida es al menos aproximadamente 1,5 veces, al menos aproximadamente 2 veces, al menos aproximadamente 2,5 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 6 veces, al menos aproximadamente 7 veces, al menos aproximadamente 8 veces, al menos aproximadamente 9 veces, al menos aproximadamente 10 veces, al menos aproximadamente 11 veces o al menos aproximadamente 12 veces más que la semivida *in vivo* de la proteína FVIII correspondiente que carece de un prolongador de la semivida de este tipo.

En una realización, la semivida de la proteína FVIII que comprende un prolongador de la semivida es de aproximadamente 1,5 veces a aproximadamente 20 veces, de aproximadamente 1,5 veces a aproximadamente 15 veces o de aproximadamente 1,5 veces a aproximadamente 10 veces mayor que la semivida *in vivo* de la proteína correspondiente que carece de dicho prolongador de la semivida de este tipo. En otra realización, la semivida de la proteína FVIII que comprende un prolongador de la semivida se prolonga de aproximadamente 2 veces a aproximadamente 10 veces, de aproximadamente 2 veces a aproximadamente 9 veces, de aproximadamente 2 veces a aproximadamente 8 veces, de aproximadamente 2 veces a aproximadamente 7 veces, de aproximadamente 2 veces a aproximadamente 6 veces, de aproximadamente 2 veces a aproximadamente 5 veces, de aproximadamente 2 veces a aproximadamente 4 veces, de aproximadamente 2 veces a aproximadamente 3 veces, de aproximadamente 2,5 veces a aproximadamente 10 veces, de aproximadamente 2,5 veces a aproximadamente 9 veces, de aproximadamente 2,5 veces a aproximadamente 8 veces, de aproximadamente 2,5 veces a aproximadamente 7 veces, de aproximadamente 2,5 veces a aproximadamente 6 veces, de aproximadamente 2,5 veces a aproximadamente 5 veces, de aproximadamente 2,5 veces a aproximadamente 4 veces, de aproximadamente 2,5 veces a aproximadamente 3 veces, de aproximadamente 3 veces a aproximadamente 10 veces, de aproximadamente 3 veces a aproximadamente 9 veces, de aproximadamente 3 veces a aproximadamente 8 veces, de aproximadamente 3 veces a aproximadamente 7 veces, de aproximadamente 3 veces a aproximadamente 6 veces, de aproximadamente 3 veces a aproximadamente 5 veces, de aproximadamente 3 veces a aproximadamente 4 veces, de aproximadamente 4 veces a aproximadamente 6 veces, de aproximadamente 5 veces a aproximadamente 7 veces o de aproximadamente 6 veces a aproximadamente 8 veces en comparación con la semivida *in vivo* de la proteína correspondiente que carece de un prolongador de la semivida de este tipo.

En otras realizaciones, la semivida de la proteína FVIII que comprende un prolongador de la semivida es de al menos aproximadamente 17 horas, al menos aproximadamente 18 horas, al menos aproximadamente 19 horas, al menos aproximadamente 20 horas, al menos aproximadamente 21 horas, al menos aproximadamente 22 horas, al menos aproximadamente 23 horas, al menos aproximadamente 24 horas, al menos aproximadamente 25 horas, al menos aproximadamente 26 horas, al menos aproximadamente 27 horas, al menos aproximadamente 28 horas, al menos aproximadamente 29 horas, al menos aproximadamente 30 horas, al menos aproximadamente 31 horas, al menos aproximadamente 32 horas, al menos aproximadamente 33 horas, al menos aproximadamente 34 horas, al menos aproximadamente 35 horas, al menos aproximadamente 36 horas, al menos aproximadamente 48 horas, al menos aproximadamente 60 horas, al menos aproximadamente 72 horas, al menos aproximadamente 84 horas, al menos aproximadamente 96 horas o al menos aproximadamente 108 horas.

En aún otras realizaciones, la semivida de la proteína FVIII que comprende un prolongador de la semivida es de aproximadamente 15 horas a aproximadamente dos semanas, de aproximadamente 16 horas a aproximadamente una semana, de aproximadamente 17 horas a aproximadamente una semana, de aproximadamente 18 horas a aproximadamente una semana, de aproximadamente 19 horas a aproximadamente una semana, de aproximadamente 20 horas a aproximadamente una semana, de aproximadamente 21 horas a aproximadamente una semana, de aproximadamente 22 horas a aproximadamente una semana, de aproximadamente 23 horas a aproximadamente una semana, de aproximadamente 24 horas a aproximadamente una semana, de aproximadamente 36 horas a aproximadamente una semana, de aproximadamente 48 horas a aproximadamente una semana, de aproximadamente 60 horas a aproximadamente una semana, de aproximadamente 24 horas a aproximadamente seis días, de aproximadamente 24 horas a aproximadamente cinco días, de aproximadamente 24 horas a aproximadamente cuatro días, de aproximadamente 24 horas a aproximadamente tres días o de aproximadamente 24 horas a aproximadamente dos días.

En algunas realizaciones, la semivida media por sujeto de la proteína FVIII que comprende un prolongador de semivida es de aproximadamente 15 horas, aproximadamente 16 horas, aproximadamente 17 horas, aproximadamente 18 horas, aproximadamente 19 horas, aproximadamente 20 horas, aproximadamente 21 horas, aproximadamente 22 horas, aproximadamente 23 horas, aproximadamente 24 horas (1 día), aproximadamente 25 horas, aproximadamente 26 horas,

aproximadamente 27 horas, aproximadamente 28 horas, aproximadamente 29 horas, aproximadamente 30 horas, aproximadamente 31 horas, aproximadamente 32 horas, aproximadamente 33 horas, aproximadamente 34 horas, aproximadamente 35 horas, aproximadamente 36 horas, aproximadamente 40 horas, aproximadamente 44 horas, aproximadamente 48 horas (2 días), aproximadamente 54 horas, aproximadamente 60 horas, aproximadamente 72 horas (3 días), aproximadamente 84 horas, aproximadamente 96 horas (4 días), aproximadamente 108 horas, aproximadamente 120 horas (5 días), aproximadamente seis días, aproximadamente siete días (una semana), aproximadamente ocho días, aproximadamente nueve días, aproximadamente 10 días, aproximadamente 11 días, aproximadamente 12 días, aproximadamente 13 días o aproximadamente 14 días.

Uno o más prolongadores de la semivida se pueden condensar con el extremo C o el extremo N de FVIII o insertarse dentro de FVIII.

### 1. Una Región Constante de Inmunoglobulina o una Porción de la Misma

En otro aspecto, un resto heterólogo comprende una o más regiones constantes de inmunoglobulina o porciones de las mismas (p. *ej.*, una región Fc). En una realización, una molécula de ácido nucleico aislada de la divulgación comprende, además, una secuencia de ácido nucleico heteróloga que codifica una región constante de inmunoglobulina o una porción de la misma. En algunas realizaciones, la región constante de inmunoglobulina o una porción de la misma es una región Fc.

Una región constante de inmunoglobulina está compuesta por dominios denominados CH (constante pesado) (CH1, CH2, etc.). Dependiendo del isotipo (*es decir*, IgG, IgM, IgA IgD o IgE), la región constante puede estar compuesta por tres o cuatro dominios CH. Algunas regiones constantes de isotipos (p. *ej.*, IgG) también contienen una región bisagra. Véase Janeway et al. 2001, Immunobiology, Garland Publishing, N.Y., N.Y.

Una región constante de inmunoglobulina o una porción de la misma para producir la proteína FVIII de la presente divulgación puede obtenerse de un cierto número de fuentes diferentes. En una realización, una región constante de inmunoglobulina o una porción de la misma se deriva de una inmunoglobulina humana. Sin embargo, se entiende que la región constante de inmunoglobulina o una porción de la misma se puede derivar de una inmunoglobulina de otra especie de mamífero, incluyendo, por ejemplo, un roedor (p. *ej.*, un ratón, rata, conejo, cobaya) o una especie de primate no humano (p. *ej.*, chimpancé, macaco). Además, la región constante de inmunoglobulina o una parte de la misma se puede derivar de cualquier clase de inmunoglobulina, incluyendo IgM, IgG, IgD, IgA e IgE, y cualquier isotipo de inmunoglobulina, incluyendo IgG1, IgG2, IgG3 e IgG4. En una realización, se utiliza el isotipo humano IgG1.

Una diversidad de secuencias génicas de la región constante de inmunoglobulina (p. *ej.*, secuencias génicas de la región constante humana) están disponibles en forma de depósitos de acceso público. La secuencia de dominios de la región constante se puede seleccionar con una función de efector particular (o sin una función de efector particular) o con una modificación particular para reducir la inmunogenicidad. Se han publicado muchas secuencias de anticuerpos y genes que codifican anticuerpos y pueden obtenerse secuencias de regiones constantes de Ig adecuadas (p. *ej.*, secuencias bisagra, CH2 y/o CH3, o porciones de las mismas) a partir de estas secuencias utilizando técnicas reconocidas en la técnica. El material genético obtenido utilizando cualquiera de los métodos anteriores puede entonces alterarse o sintetizarse para obtener los polipéptidos de la presente divulgación. Se apreciará además que el alcance de esta divulgación abarca alelos, variantes y mutaciones de secuencias de ADN de región constante.

Las secuencias de la región constante de inmunoglobulina o una porción de la misma pueden clonarse, p. *ej.*, utilizando la reacción en cadena de la polimerasa y cebadores que se seleccionan para amplificar el dominio de interés. Para clonar una secuencia de la región constante de inmunoglobulina o una porción de la misma a partir de un anticuerpo, se puede aislar el ARNm de células de hibridoma, bazo o linfa, transcribirlo inversamente en ADN y amplificar los genes del anticuerpo mediante PCR. Los métodos de amplificación por PCR se describen en detalle en las Pat. de EE.UU. N°s 4.683.195; 4.683.202; 4.800.159; 4.965.188; y, p. *ej.*, en "PCR Protocols: A Guide to Methods and Applications", Innis et al. eds., Academic Press, San Diego, CA (1990); Ho et al. 1989. Gen 77:51; Horton et al. 1993. Methods Enzymol. 217:270). La PCR puede iniciarse con cebadores de región constante de consenso o con cebadores más específicos basados en las secuencias de aminoácidos y ADN de cadena ligera y pesada publicadas. La PCR también se puede utilizar para aislar clones de ADN que codifican las cadenas ligera y pesada del anticuerpo. En este caso, las colecciones pueden seleccionarse mediante cebadores de consenso o sondas homólogas más grandes, tales como sondas de región constante de ratón. Se conocen en la técnica numerosos conjuntos de cebadores adecuados para la amplificación de genes de anticuerpos (p. *ej.*, cebadores 5' basados en la secuencia N-terminal de anticuerpos purificados (Benhar y Pastan. 1994. Protein Engineering 7:1509); amplificación rápida de extremos de ADNc (Ruberti, F. et al. 1994. J. Immunol. Methods 173:33), secuencias conductoras de anticuerpos (Larrick et al. 1989 Biochem. Biophys. Res. Commun. 160:1250). La clonación de secuencias de anticuerpos se describe adicionalmente en Newman et al., Pat. de EE.UU. N° 5.658.570, presentada el 25 de enero de 1995.

Una región constante de inmunoglobulina utilizada en esta memoria puede incluir todos los dominios y la región bisagra o porciones de la misma. En una realización, la región constante de inmunoglobulina o una porción de la misma comprende un dominio CH2, un dominio CH3 y una región bisagra, *es decir*, una región Fc o un participante en la unión a FcRn.

Tal como se utiliza en esta memoria, la expresión "región Fc" se define como la porción de un polipéptido que corresponde a la región Fc de la Ig nativa, *es decir*, como se forma por la asociación dimérica de los respectivos dominios Fc de sus dos cadenas pesadas. Una región Fc nativa forma un homodímero con otra región Fc. Por el contrario, la expresión "región Fc condensada genéticamente" o "región Fc de cadena sencilla" (región scFc), tal como se utiliza en esta memoria, se refiere a una región Fc dimérica sintética compuesta por dominios Fc enlazados genéticamente dentro de una cadena polipeptídica única (*es decir*, codificada en una única secuencia genética contigua). Véase la Publicación Internacional N° WO 2012/006635.

En una realización, la "región Fc" se refiere a la porción de una sola cadena pesada de Ig que comienza en la región bisagra justo aguas arriba del sitio de escisión de papaína (*es decir*, el residuo 216 en IgG, tomando el primer residuo de la región constante de la cadena pesada como 114) y terminando en el extremo C del anticuerpo. Por consiguiente, una región Fc completa comprende al menos un dominio bisagra, un dominio CH2 y un dominio CH3.

Una región constante de inmunoglobulina o una porción de la misma puede ser una pareja de unión a FcRn. FcRn es activo en tejidos epiteliales adultos y se expresa en la luz de los intestinos, las vías respiratorias pulmonares, las superficies nasales, las superficies vaginales, el colon y las superficies rectales (Pat. de EE.UU. N° 6.485.726). Un participante en la unión a FcRn es una porción de una inmunoglobulina que se une a FcRn.

El receptor de FcRn ha sido aislado de varias especies de mamíferos, incluyendo los seres humanos. Las secuencias del FcRn humano, FcRn de mono, FcRn de rata y FcRn de ratón son conocidas (Story et al. 1994, J. Exp. Med. 180:2377). El receptor FcRn se une a IgG (pero no a otras clases de inmunoglobulinas tales como IgA, IgM, IgD e IgE) a un pH relativamente bajo, transporta activamente la IgG transcelularmente en una dirección luminal a serosa y luego libera la IgG a un pH relativamente más alto que se encuentra en los fluidos intersticiales. Se expresa en tejido epitelial adulto (Pat. de EE.UU. N°s 6.485.726, 6.030.613, 6.086.875; WO 03/077834; US2003-0235536A1) incluyendo epitelio pulmonar e intestinal (Israel et al. 1997, Immunology 92:69), epitelio tubular proximal renal (Kobayashi et al., 2002, Am. J. Physiol. Renal Physiol. 282:F358), así como epitelio nasal, superficies vaginales y superficies de árboles biliares.

Participantes en la unión a FcRn útiles en la presente divulgación abarcan moléculas que pueden unirse específicamente al receptor de FcRn que incluyen IgG completa, el fragmento Fc de IgG y otros fragmentos que incluyen la región de unión completa del receptor de FcRn. La región de la porción Fc de IgG que se une al receptor de FcRn se ha descrito basándose en cristalografía de rayos X (Burmeister et al. 1994, Nature 372:379). La principal área de contacto de Fc con FcRn está cerca de la unión de los dominios CH2 y CH3. Los contactos Fc-FcRn están todos dentro de una única cadena pesada de Ig. Los participantes en la unión a FcRn incluyen IgG completa, el fragmento Fc de IgG y otros fragmentos de IgG que incluyen la región de unión completa de FcRn. Los principales sitios de contacto incluyen los residuos de aminoácidos 248, 250-257, 272, 285, 288, 290-291, 308-311 y 314 del dominio CH2 y los residuos de aminoácidos 385-387, 428 y 433-436 del dominio CH3. Las referencias hechas a la numeración de aminoácidos de inmunoglobulinas o fragmentos de inmunoglobulina, o regiones, se basan todas en Kabat et al. 1991, Sequences of Proteins of Immunological Interest, Departamento de Salud Pública de EE.UU., Bethesda, Md.

Las regiones Fc o los participantes en la unión a FcRn unidos a FcRn pueden transportarse eficazmente a través de las barreras epiteliales mediante FcRn, proporcionando así un medio no invasivo para administrar sistémicamente una molécula terapéutica deseada. Adicionalmente, las proteínas de fusión que comprenden una región Fc o un participante en la unión a FcRn son sometidas a endocitosis por parte de células que expresan FcRn. Pero en lugar de ser marcadas para la degradación, estas proteínas de fusión se reciclan de nuevo a la circulación, aumentando así la semivida *in vivo* de estas proteínas. En determinadas realizaciones, las porciones de regiones constantes de inmunoglobulina son una región Fc o un participante en la unión a FcRn que típicamente se asocia, a través de enlaces disulfuro y otras interacciones no específicas, con otra región Fc u otro participante en la unión a FcRn para formar dímeros y multímeros de orden superior.

Dos receptores FcRn pueden unirse a una única molécula de Fc. Los datos cristalográficos sugieren que cada molécula de FcRn se une a un solo polipéptido del homodímero Fc. En una realización, el participante en la unión a FcRn, *p. ej.*, un fragmento Fc de una IgG, a una molécula biológicamente activa proporciona un medio para suministrar la molécula biológicamente activa por vía oral, bucal, sublingual, rectal, vaginal, como un aerosol administrado por vía nasal o por vía pulmonar, o por vía ocular. En otra realización, la proteína FVIII se puede administrar de forma invasiva, *p. ej.*, por vía subcutánea, por vía intravenosa.

Una región del participante en la unión a FcRn es una molécula o porción de la misma que puede unirse específicamente al receptor FcRn con el consiguiente transporte activo por parte del receptor FcRn de la región Fc. Específicamente, la unión se refiere a dos moléculas que forman un complejo que es relativamente estable en condiciones fisiológicas. La unión específica se caracteriza por una alta afinidad y una capacidad de baja a moderada, a diferencia de la unión no específica que habitualmente tiene una baja afinidad con una capacidad de moderada a alta. Típicamente, la unión se considera específica cuando la constante de afinidad  $K_A$  es mayor que  $10^6 \text{ M}^{-1}$  o mayor que  $10^8 \text{ M}^{-1}$ . Si es necesario, la unión no específica puede reducirse sin afectar sustancialmente a la unión específica variando las condiciones de unión. Las condiciones de unión apropiadas, tales como la concentración de las moléculas, la fuerza iónica de la solución, la temperatura, el tiempo permitido para la unión, la concentración de un agente de bloqueo (*p. ej.*, albúmina sérica, caseína de la leche), etc., pueden ser optimizadas por un experto en la materia utilizando técnicas de rutina.



En determinadas realizaciones, una proteína FVIII codificada por la molécula de ácido nucleico de la divulgación comprende una o más regiones Fc truncadas que, no obstante, son suficientes para conferir propiedades de unión al receptor Fc (FcR) a la región Fc. Por ejemplo, la porción de una región Fc que se une a FcRn (*es decir*, la porción de unión a FcRn) comprende aproximadamente los aminoácidos 282-438 de IgG1, numeración de la UE (siendo los sitios de contacto principales los aminoácidos 248, 250-257, 272, 285, 288, 290-291, 308-311 y 314 del dominio CH2 y los residuos de aminoácidos 385-387, 428 y 433-436 del dominio CH3. Por lo tanto, una región Fc de la divulgación puede comprender o consistir en una porción de unión a FcRn. Porciones de unión a FcRn se pueden derivar de cadenas pesadas de cualquier isotipo, incluyendo IgG1, IgG2, IgG3 e IgG4. En una realización, se utiliza una porción de unión a FcRn de un anticuerpo del isotipo humano IgG1. En otra realización, se utiliza una porción de unión a FcRn de un anticuerpo del isotipo humano IgG4.

La región Fc se puede obtener de un cierto número de fuentes diferentes. En una realización, una región Fc del polipéptido se deriva de una inmunoglobulina humana. Sin embargo, se entiende que un resto Fc se puede derivar de una inmunoglobulina de otra especie de mamífero, incluyendo, por ejemplo, un roedor (p. ej., un ratón, rata, conejo, cobaya) o una especie de primate no humano (p. ej., chimpancé, macaco). Además, el polipéptido de los dominios Fc o sus porciones se pueden derivar de cualquier clase de inmunoglobulina, incluyendo IgM, IgG, IgD, IgA e IgE, y cualquier isotipo de inmunoglobulina, incluyendo IgG1, IgG2, IgG3 e IgG4. En otra realización, se utiliza el isotipo humano IgG1.

En determinadas realizaciones, la variante Fc confiere un cambio en al menos una función efectora impartida por un resto Fc que comprende dicho dominio Fc de tipo salvaje (p. ej., una mejora o reducción en la capacidad de la región Fc de unirse a receptores Fc (p. ej., FcγRI, FcγRII o FcγRIII) o proteínas del complemento (p. ej., C1q), o para desencadenar citotoxicidad dependiente de anticuerpos (ADCC, por sus siglas en inglés), fagocitosis o citotoxicidad dependiente del complemento (CDC, por sus siglas en inglés)). En otras realizaciones, la variante Fc proporciona un residuo de cisteína diseñado.

La región Fc de la divulgación puede emplear variantes de Fc reconocidas en la técnica que se sabe que imparten un cambio (p. ej., una mejora o reducción) en la función del efector y/o la unión de FcR o FcRn. Específicamente, una región Fc de la divulgación puede incluir, por ejemplo, un cambio (p. ej., una sustitución) en una o más de las posiciones de aminoácidos descritas en las Publicaciones Internacionales PCT WO88/07089A1, WO96/14339A1, WO98/05787A1, WO98/23289A1, WO99/51642A1, WO99/58572A1, WO00/09560A2, WO00/32767A1, WO00/42072A2, WO02/44215A2, WO02/060919A2, WO03/074569A2, WO04/016750A2, WO04/029207A2, WO04/035752A2, WO04/063351A2, WO04/074455A2, WO04/099249A2, WO05/040217A2, WO04/044859, WO05/070963A1, WO05/077981A2, WO05/092925A2, WO05/123780A2, WO06/019447A1, WO06/047350A2 y WO06/085967A2; Publicaciones de Patentes de EE. UU. N°s US2007/0231329, US2007/0231329, US2007/0237765, US2007/0237766, US2007/0237767, US2007/0243188, US2007/0248603, US2007/0286859, US2008/0057056; o las Patentes de EE.UU. 5.648.260; 5.739.277; 5.834.250; 5.869.046; 6.096.871; 6.121.022; 6.194.551; 6.242.195; 6.277.375; 6.528.624; 6.538.124; 6.737.056; 6.821.505; 6.998.253; 7.083.784; 7.404.956 y 7.317.091. En una realización, el cambio específico (p. ej., la sustitución específica de uno o más aminoácidos descritos en la técnica) se puede realizar en una o más de las posiciones de aminoácidos descritas. En otra realización, se puede realizar un cambio diferente en una o más de las posiciones de aminoácidos descritas (p. ej., la sustitución diferente de una o más posiciones de aminoácidos descritas en la técnica).

La región Fc o el participante en la unión a FcRn de IgG se puede modificar de acuerdo con procedimientos bien reconocidos tales como mutagénesis dirigida al sitio y similares para producir fragmentos de IgG o Fc modificados o porciones de los mismos que se unirán a FcRn. Modificaciones de este tipo incluyen modificaciones alejadas de los sitios de contacto de FcRn, así como modificaciones dentro de los sitios de contacto que conservan o incluso potencian la unión a FcRn. Por ejemplo, los siguientes residuos de aminoácidos individuales en IgG1 Fc humana (Fc γ1) pueden sustituirse sin una pérdida significativa de la afinidad de unión de Fc por FcRn: P238A, S239A, K246A, K248A, D249A, M252A, T256A, E258A, T260A, D265A, S267A, H268A, E269A, D270A, E272A, L274A, N276A, Y278A, D280A, V282A, E283A, H285A, N286A, T289A, K290A, R292A, E293A, E294A, Q295A, Y296F, N297A, S298A, Y300F, R301A, V303A, V305A, T307A, L309A, Q311A, D312A, N315A, K317A, E318A, K320A, K322A, S324A, K326A, A327Q, P329A, A330Q, P331A, E333A, K334A, T335A, S337A, K338A, K340A, Q342A, R344A, E345A, Q347A, R355A, E356A, M358A, T359A, K360A, N361A, Q362A, Y373A, S375A, D376A, A378Q, E380A, E382A, S383A, N384A, Q386A, E388A, N389A, N390A, Y391F, K392A, L398A, S400A, D401A, D413A, K414A, R416A, Q418A, Q419A, N421A, V422A, S424A, E430A, N434A, T437A, Q438A, K439A, S440A, S444A y K447A, en que, por ejemplo, P238A representa prolina de tipo salvaje sustituida con alanina en la posición número 238. Como un ejemplo, una realización específica incorpora la mutación N297A, eliminando un sitio de N-glicosilación altamente conservado. Además de la alanina, los aminoácidos de tipo salvaje pueden sustituirse con otros aminoácidos en las posiciones arriba especificadas. Se pueden introducir mutaciones individualmente en Fc, dando lugar a más de cien regiones Fc distintas de la Fc nativa. Adicionalmente, se pueden introducir juntas combinaciones de dos, tres o más de estas mutaciones individuales, dando lugar a cientos de regiones Fc más.

Determinadas mutaciones anteriores pueden conferir una nueva funcionalidad a la región Fc o al participante en la unión a FcRn. Por ejemplo, una realización incorpora N297A, eliminando un sitio de N-glicosilación altamente conservado. El efecto de esta mutación es reducir la inmunogenicidad, potenciando con ello la semivida circulante de la región Fc y de

hacer que la región Fc sea incapaz de unirse a FcγRI, FcγRIIA, FcγRIIB y FcγRIIIA, sin comprometer la afinidad por FcRn (Routledge et al. 1995, Transplantation 60:847, Friend et al., 1999, Transplantation 68:1632, Shields et al., 1995, J. Biol. Chem. 276:6591). Como un ejemplo adicional de la nueva funcionalidad que surge de las mutaciones arriba descritas, la afinidad por FcRn puede incrementarse más allá de la del tipo salvaje en algunos casos. Esta afinidad incrementada puede reflejar una mayor tasa de "encendido", una disminución de la tasa de "apagado" o tanto una mayor tasa de "encendido" como una disminución de la tasa de "apagado". Ejemplos de mutaciones que se cree que imparten una afinidad incrementada por FcRn incluyen, pero no se limitan a T256A, T307A, E380A y N434A (Shields et al. 2001, J. Biol. Chem. 276:6591).

Adicionalmente, al menos tres receptores Fc gamma humanos parecen reconocer un sitio de unión en IgG dentro de la región bisagra inferior, generalmente los aminoácidos 234-237. Por lo tanto, otro ejemplo de nueva funcionalidad y potencial inmunogenicidad disminuida puede surgir de mutaciones de esta región tales como, por ejemplo, reemplazando los aminoácidos 233-236 de IgG1 humana "ELLG" (SEQ ID NO: 45) a la secuencia correspondiente de IgG2 "PVA" (con una delección de un aminoácido). Se ha demostrado que FcγRI, FcγRII y FcγRIII, que median en diversas funciones del efector, no se unirán a IgG1 cuando se hayan introducido mutaciones de este tipo. Ward y Ghetie 1995, Therapeutic Immunology 2:77 y Armour et al. 1999, Eur. J. Immunol. 29:2613.

En otra realización, la región constante de inmunoglobulina o una porción de la misma comprende una secuencia de aminoácidos en la región bisagra o una porción de la misma que forma uno o más enlaces disulfuro con una segunda región constante de inmunoglobulina o una porción de la misma. La segunda región constante de inmunoglobulina o una porción de la misma se puede enlazar a un segundo polipéptido, uniendo la proteína FVIII y el segundo polipéptido. En algunas realizaciones, el segundo polipéptido es un resto potenciador. Tal como se utiliza en esta memoria, la expresión "resto potenciador" se refiere a una molécula, un fragmento de la misma o un componente de un polipéptido que es capaz de potenciar la actividad procoagulante de FVIII. El resto potenciador puede ser un cofactor, tal como el factor tisular soluble (sTF, por sus siglas en inglés), o un péptido procoagulante. Por lo tanto, tras la activación de FVIII, el resto potenciador está disponible para potenciar la actividad de FVIII.

En determinadas realizaciones, una proteína FVIII codificada por una molécula de ácido nucleico de la divulgación comprende una sustitución de un aminoácido en una región constante de inmunoglobulina o una porción de la misma (*p. ej.*, variantes Fc), que altera las funciones efectoras independientes del antígeno de la región constante de Ig, en particular la semivida circulante de la proteína.

## 2. Regiones scFc

En otro aspecto, un resto heterólogo comprende una región scFc (Fc de cadena sencilla). En una realización, una molécula de ácido nucleico aislada de la divulgación comprende, además, una secuencia de ácido nucleico heteróloga que codifica una región scFc. La región scFc comprende al menos dos regiones constantes de inmunoglobulina o porciones de las mismas (*p. ej.*, restos o dominios Fc (*p. ej.*, 2, 3, 4, 5, 6 o más restos o dominios Fc)) dentro de la misma cadena polipeptídica lineal que son capaces de plegarse (*p. ej.*, plegarse intramolecular o intermolecularmente) para formar una región scFc funcional que está enlazada por un enlazador peptídico Fc. Por ejemplo, en una realización, un polipéptido de la divulgación es capaz de unirse, a través de su región scFc, a al menos un receptor Fc (*p. ej.*, un FcRn, un receptor FcγR (*p. ej.*, FcγRIII) o una proteína del complemento (*p. ej.*, C1q)) con el fin de mejorar la semivida o desencadenar una función efectora inmunitaria (*p. ej.*, citotoxicidad dependiente de anticuerpos (ADCC), fagocitosis o citotoxicidad dependiente del complemento (CDCC) y/o para mejorar la capacidad de fabricación).

## 3. CTP

En otro aspecto, un resto heterólogo comprende un péptido C-terminal (CTP) de la subunidad β de la gonadotropina coriónica humana o un fragmento, variante o derivado del mismo. Se sabe que uno o más péptidos CTP insertados en una proteína recombinante aumentan la semivida *in vivo* de esa proteína. Véase, *p. ej.*, la Patente de EE.UU. N° 5.712.122.

Péptidos CTP ejemplares incluyen DPRFQDSSSSKAPPSRLPGPSDTPIL (SEQ ID NO: 33) o SSSSKAPPSRLPGPSDTPILPQ (SEQ ID NO: 34). Véase, *p. ej.*, la Publicación de Solicitud de Patente de EE.UU. N° US 2009/0087411 A1.

## 4. Secuencia XTEN

En algunas realizaciones, un resto heterólogo comprende una o más secuencias XTEN, fragmentos, variantes o derivados de las mismas. Tal como se utiliza aquí, "secuencia XTEN" se refiere a polipéptidos de longitud extendida con secuencias que se producen de forma no natural, sustancialmente no repetitivas, que están compuestas principalmente de pequeños aminoácidos hidrófilos, teniendo la secuencia un grado bajo o ninguna estructura secundaria o terciaria bajo condiciones fisiológicas. Como resto heterólogo, las XTENs pueden servir como resto de prolongación de la semivida. Además, XTEN puede proporcionar propiedades deseables que incluyen, pero no se limitan a parámetros farmacocinéticos potenciados y características de solubilidad.

La incorporación de un resto heterólogo que comprende una secuencia XTEN en una proteína de la divulgación puede conferir a la proteína una o más de las siguientes propiedades ventajosas: flexibilidad conformacional, solubilidad acuosa potenciada, alto grado de resistencia a la proteasa, baja inmunogenicidad, baja unión a receptores de mamíferos, o radios hidrodinámicos (o de Stokes) incrementados.

- 5 En determinados aspectos, una secuencia de XTEN puede aumentar las propiedades farmacocinéticas, tales como una semivida *in vivo* más prolongada o un área bajo la curva (AUC, por sus siglas en inglés) incrementada, de modo que una proteína de la divulgación permanece *in vivo* y tiene actividad procoagulante durante un período de tiempo prolongado en comparación con una proteína con el mismo pero sin el resto heterólogo XTEN.

- 10 En algunas realizaciones, la secuencia de XTEN útil para la divulgación es un péptido o un polipéptido que tiene más de aproximadamente 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1200, 1400, 1600, 1800 o 2000 residuos aminoácidos. En determinadas realizaciones, XTEN es un péptido o un polipéptido que tiene más de aproximadamente 20 a aproximadamente 3000 residuos aminoácidos, más de 30 a aproximadamente 2500 residuos, más de 40 a aproximadamente 2000 residuos, más de 50 a aproximadamente 1500 residuos, más de 60 a aproximadamente 1000 residuos, más de 70 a aproximadamente 900 residuos, más de 80 a aproximadamente 800 residuos, más de 90 a aproximadamente 700 residuos, más de 100 a aproximadamente 600 residuos, más de 110 a aproximadamente 500 residuos, o más de 120 a aproximadamente 400 residuos. En una realización particular, la XTEN comprende una secuencia de aminoácidos de más de 42 aminoácidos de longitud y menos de 144 aminoácidos de longitud.

- 20 La secuencia de XTEN de la divulgación puede comprender uno o más motivos de secuencia de 5 a 14 (p. ej., 9 a 14) residuos aminoácidos o una secuencia de aminoácidos de al menos 80 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % idéntico al motivo de secuencia, en donde el motivo comprende, consiste esencialmente en o consiste en 4 a 6 tipos de aminoácidos (p. ej., 5 aminoácidos) seleccionados del grupo que consiste en glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P). Véase el documento US 2010-0239554 A1.

- 25 En algunas realizaciones, XTEN comprende motivos de secuencia que no se solapan, en los que aproximadamente el 80 %, o al menos aproximadamente el 85 %, o al menos aproximadamente el 90 %, o aproximadamente el 91 %, o aproximadamente el 92 %, o aproximadamente el 93 %, o aproximadamente el 94 %, o aproximadamente el 95 %, o aproximadamente el 96 %, o aproximadamente el 97 %, o aproximadamente el 98 %, o aproximadamente el 99 % o aproximadamente el 100 % de la secuencia consiste en múltiples unidades de secuencias no superpuestas seleccionadas de una familia de un solo motivo seleccionada de la Tabla 4, dando como resultado una secuencia de la familia. Tal como se utiliza en esta memoria, "familia" significa que XTEN tiene motivos seleccionados solo de una sola categoría de motivos de la Tabla 4; es decir, AD, AE, AF, AG, AM, AQ, BC o BD XTEN, y que cualquier otro aminoácido en XTEN que no sea de un motivo de la familia se seleccione para lograr una propiedad necesaria, tal como para permitir la incorporación de un restricción sitio mediante la codificación de nucleótidos, incorporación de una secuencia de escisión, o para lograr un mejor enlace a FVIII. En algunas realizaciones de familias XTEN, una secuencia XTEN comprende múltiples unidades de motivos de secuencia no solapantes de la familia de motivos AD o de la familia de motivos AE o de la familia de motivos AF o de la familia de motivos AG o de la familia de motivos AM o de la familia de motivos AQ, o de la familia BC o de la familia BD, exhibiendo el XTEN resultante el intervalo de homología arriba descrito. En otras realizaciones, XTEN comprende múltiples unidades de secuencias de motivos de dos o más de las familias de motivos de la Tabla 2A. Estas secuencias se pueden seleccionar para lograr las características físicas/químicas deseadas, incluyendo propiedades tales como carga neta, hidrofilia, falta de estructura secundaria o falta de repetitividad que confiere la composición de aminoácidos de los motivos, que se describen con más detalle más adelante. En las realizaciones descritas anteriormente en esta memoria, los motivos incorporados en el XTEN se pueden seleccionar y ensamblar utilizando los métodos descritos en esta memoria para lograr un XTEN de aproximadamente 36 a aproximadamente 3000 residuos de aminoácidos.

**Tabla 4. Motivos de Secuencia XTEN de 12 Aminoácidos y Familias de Motivos**

Familia de Motivos *	SECUENCIA DE MOTIVOS	SEQ ID NO:
AD	GESPGGSSGSES	73
AD	GSEGSSGPGESS	74
AD	GSSESGSSEGGP	75
AD	GSGGEPSESGSS	76
AE, AM	GSPAGSPTSTEE	77
AE, AM, AQ	GSEPATSGSETP	78
AE, AM, AQ	GTSESATPESGP	79
AE, AM, AQ	GTSTEPSEGSAP	80
AF, AM	GSTSESPSGTAP	81
AF, AM	GTSTPESGSASP	82
AF, AM	GTSPSGESTAP	83
AF, AM	GSTSSTAESPGP	84
AG, AM	GTPGSGTASSSP	85
AG, AM	GSSTPSGATGSP	86
AG, AM	GSSPSASTGTGP	87
AG, AM	GASPGTSSTGSP	88
AQ	GEPAGSPTSTSE	89
AQ	GTGEPSSTPASO	90
AQ	GSGPSTESAPTE	91
AQ	GSETPSGPSETA	92
AQ	GPSETSTSEPGA	93
AQ	GSPSEPTGTSA	94
BC	GSGASEPTPASO	95
BC	GSEPATSGTEPS	96
BC	GTSEPSTSEPGA	97
BC	GTSTEPSEPGSA	98
BD	GSTAGSETSTEA	99
BD	GSETATGSETA	100
BD	GTSESATSESGA	101
BD	GTSTEASEGSAS	102
* Designa secuencias de motivos individuales que, cuando se utilizan juntas en diversas permutaciones, dan como resultado una "secuencia de familiar"		

- 5 Ejemplos de secuencias de XTEN que pueden utilizarse como restos heterólogos en proteínas quiméricas de la divulgación se describen, *p. ej.*, en las Publicaciones de Patentes de EE.UU. N<sup>os</sup> 2010/0239554 A1, 2010/0323956 A1, 2011/0046060 A1, 2011/0046061 A1, 2011/0077199 A1 o 2011/0172146 A1, o las Publicaciones de Patente Internacional N<sup>os</sup> WO 2010091122 A1, WO 2010144502 A2, WO 2010144508 A1, WO 2011028228 A1, WO 2011028229 A1 o WO 2011028344 A2.
- 10 XTEN puede tener longitudes variables para la inserción o el enlace a FVIII. En una realización, la longitud de la o las secuencias de XTEN se elige en base a la propiedad o función a lograr en la proteína de fusión. Dependiendo de la propiedad o función prevista, XTEN puede ser una secuencia de longitud corta o intermedia o una secuencia más larga que puede servir como soporte. En determinadas realizaciones, XTEN incluye segmentos cortos de aproximadamente 6 a aproximadamente 99 residuos de aminoácidos, longitudes intermedias de aproximadamente 100 a aproximadamente 399
- 15 residuos de aminoácidos y longitudes más largas de aproximadamente 400 a aproximadamente 1000 y hasta aproximadamente 3000 residuos de aminoácidos. Por lo tanto, XTEN insertado en o enlazado a FVIII puede tener longitudes de aproximadamente 6, aproximadamente 12, aproximadamente 36, aproximadamente 40, aproximadamente 42, aproximadamente 72, aproximadamente 96, aproximadamente 144, aproximadamente 288, aproximadamente 400, aproximadamente 500, aproximadamente 576, aproximadamente 600, aproximadamente 700, aproximadamente 800,
- 20 aproximadamente 864, aproximadamente 900, aproximadamente 1000, aproximadamente 1500, aproximadamente 2000, aproximadamente 2500 o hasta aproximadamente 3000 residuos aminoácidos de longitud. En otras realizaciones, las secuencias de XTEN son de aproximadamente 6 a aproximadamente 50, de aproximadamente 50 a aproximadamente 100, de aproximadamente 100 a 150, de aproximadamente 150 a 250, de aproximadamente 250 a 400, de aproximadamente 400 a 500, de aproximadamente 500 a aproximadamente 900, de aproximadamente 900 a 1500, de
- 25 aproximadamente 1500 a 2000 o de aproximadamente 2000 a aproximadamente 3000 residuos aminoácidos de longitud. La longitud precisa de un XTEN insertado en o enlazado a FVIII puede variar sin afectar negativamente a la actividad del FVIII. En una realización, uno o más de los XTEN utilizados en esta memoria tienen 42 aminoácidos, 72 aminoácidos, 144 aminoácidos, 288 aminoácidos, 576 aminoácidos u 864 aminoácidos de longitud y se pueden seleccionar de uno o más de las secuencias de familias XTEN; *es decir*, AD, AE, AF, AG, AM, AQ, BC o BD.

En algunas realizaciones, la secuencia de XTEN utilizada en la divulgación es al menos 60 %, 70 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % idéntico a una secuencia seleccionada del grupo que consiste en AE42, AG42, AE48, AM48, AE72, AG72, AE108, AG108, AE144, AF144, AG144, AE180, AG180, AE216, AG216, AE252, AG252, AE288, AG288, AE324, AG324, AE360, AG360, AE396, AG396, AE432, AG432, AE468, AG468, AE504, AG504, AF504, AE540, AG540, AF540, AD576, AE576, AF576, AG576, AE612, AG612, AE624, AE648, AG648, AG684, AE720, AG720, AE756, AG756, AE792, AG792, AE828, AG828, AD836, AE864, AF864, AG864, AM875, AE912, AM923, AM1318, BC864, BD864, AE948, AE1044, AE1140, AE1236, AE1332, AE1428, AE1524, AE1620, AE1716, AE1812, AE1908, AE2004A, AG948, AG1044, AG1140, AG1236, AG1332, AG1428, AG1524, AG1620, AG1716, AG1812, AG1908, AG2004, y cualquier combinación de los mismos. Véase el documento US 2010-0239554 A1. En una realización particular, el XTEN comprende AE42, AE72, AE144, AE288, AE576, AE864, AG 42, AG72, AG144, AG288, AG576, AG864 o cualquier combinación de los mismos.

Secuencias ejemplares de XTEN que se pueden utilizar como restos heterólogos en la proteína quimérica de la divulgación incluyen XTEN AE42-4 (SEQ ID NO: 46, codificada por SEQ ID NO: 47; FIGs. 11C y 11D, respectivamente), XTEN 144-2A (SEQ ID NO: 48, codificada por SEQ ID NO: 49; FIGs. 11E y 11F, respectivamente), XTEN A144-3B (SEQ ID NO: 50, codificada por SEQ ID NO: 51; FIGs. 11G y 11H, respectivamente), XTEN AE144-4A (SEQ ID NO: 52, codificada por SEQ ID NO: 53; FIGs. 11I y 11J, respectivamente), XTEN AE144-5A (SEQ ID NO: 54, codificada por SEQ ID NO: 55; FIGs. 11K y 11L, respectivamente), XTEN AE144-6B (SEQ ID NO: 56, codificada por SEQ ID NO: 57; FIGs. 11M y 11N, respectivamente), XTEN AG144-1 (SEQ ID NO: 58, codificada por SEQ ID NO: 59; FIGs. 11O y 11P, respectivamente), XTEN AG144-A (SEQ ID NO: 60, codificada por SEQ ID NO: 61; FIGs. 11Q y 11R, respectivamente), XTEN AG144-B (SEQ ID NO: 62, codificada por SEQ ID NO: 63; FIGs. 11S y 11T, respectivamente), XTEN AG144-C (SEQ ID NO: 64, codificada por SEQ ID NO: 65; FIGs. 11U y 11V, respectivamente) y XTEN AG144-F (SEQ ID NO: 66, codificada por SEQ ID NO: 67; FIGs. 11W y 11X, respectivamente). En una realización particular, el XTEN es codificado por SEQ ID NO :18.

En algunas realizaciones, menos del 100 % de los aminoácidos de un XTEN se seleccionan de glicina (G), alanina (A), serina (S), treonina (T), glutamato (E) y prolina (P), o menos del 100 % de la secuencia consiste en los motivos de secuencia de la Tabla 2A o una secuencia XTEN proporcionada en esta memoria. En tales realizaciones, los residuos aminoácidos restantes de XTEN se seleccionan de cualquiera de los otros 14 L-aminoácidos naturales, pero se pueden seleccionar preferentemente de aminoácidos hidrófilos de modo que la secuencia de XTEN contenga al menos aproximadamente 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o al menos aproximadamente 99 % de aminoácidos hidrófilos. El contenido de aminoácidos hidrófobos en el XTEN utilizado en las construcciones de conjugación puede ser menor que 5 %, menor que 2 % o menor que 1 % de contenido de aminoácidos hidrófobos. Los residuos hidrófobos menos favorecidos en la construcción de XTEN incluyen triptófano, fenilalanina, tirosina, leucina, isoleucina, valina y metionina. Además, las secuencias de XTEN pueden contener menos del 5 % o menos del 4 % o menos del 3 % o menos del 2 % o menos del 1 % o ninguno de los siguientes aminoácidos: metionina (por ejemplo, para evitar la oxidación) o asparagina y glutamina (para evitar la desamidación).

La una o más secuencias de XTEN pueden insertarse en el extremo C o en el extremo N de la secuencia de aminoácidos codificada por la secuencia de nucleótidos o pueden insertarse entre dos aminoácidos en la secuencia de aminoácidos codificada por la secuencia de nucleótidos. Por ejemplo, el XTEN se puede insertar entre dos aminoácidos en uno o más sitios de inserción seleccionados de la Tabla 3. Se pueden encontrar ejemplos de sitios dentro del FVIII que son permisibles para la inserción de XTEN, *p. ej.*, en la Publicación Internacional N° WO 2013/123457 A1 o la publicación de EE.UU. N° 2015/0158929 A1.

## 5. Albúmina o Fragmento, Derivado o Variante del Mismo

En algunas realizaciones, un resto heterólogo comprende albúmina o un fragmento funcional de la misma. La albúmina sérica humana (HSA o HA, por sus siglas en inglés), una proteína de 609 aminoácidos en su forma completa, es responsable de una proporción significativa de la presión osmótica del suero y también funciona como un soporte de ligandos endógenos y exógenos. El término "albúmina", tal como se utiliza en esta memoria, incluye albúmina de longitud completa o un fragmento funcional, variante, derivado o análogo de la misma. Ejemplos de albúmina o los fragmentos o variantes de la misma se describen en las Publ. de Pat. de EE.UU. N°s 2008/0194481A1, 2008/0004206 A1, 2008/0161243 A1, 2008/0261877 A1 o 2008/0153751 A1 o la Publ. de Solicitud PCT N° 2008/033413 A2, 2009/058322 A1 o 2007/021494 A2.

En una realización, la proteína FVIII codificada por una molécula de ácido nucleico de la divulgación comprende albúmina, un fragmento o una variante de la misma que está además enlazada a un segundo resto heterólogo seleccionado del grupo que consiste en una región o porción constante de inmunoglobulina del mismo (*p. ej.*, una región Fc), una secuencia PAS, HES y PEG.

## 6. Resto de unión a Albúmina

En determinadas realizaciones, el resto heterólogo es un resto de unión a albúmina, que comprende un péptido de unión a albúmina, un dominio de unión a albúmina bacteriano, un fragmento de anticuerpo de unión a albúmina o cualquier combinación de los mismos.

Por ejemplo, la proteína de unión a albúmina puede ser una proteína de unión a albúmina bacteriana, un anticuerpo o un fragmento de anticuerpo que incluye anticuerpos de dominio (véase la Pat. de EE.UU. N° 6.696.245). Una proteína de unión a albúmina, por ejemplo, puede ser un dominio de unión a albúmina bacteriano tal como el de la proteína G estreptocócica (Konig, T. y Skerra, A. (1998) J. Immunol. Methods 218, 73-83). Otros ejemplos de péptidos de unión a albúmina que pueden utilizarse como participante en la conjugación son, por ejemplo, los que tienen una secuencia de consenso Cys-Xaa<sub>1</sub> - Xaa<sub>2</sub> - Xaa<sub>3</sub> - Xaa<sub>4</sub>-Cys, en donde Xaa<sub>1</sub> es Asp, Asn, Ser, Thr o Trp; Xaa<sub>2</sub> es Asn, Gln, His, Ile, Leu o Lys; Xaa<sub>3</sub> es Ala, Asp, Phe, Trp o Tyr; y Xaa<sub>4</sub> es Asp, Gly, Leu, Phe, Ser o Thr tal como se describe en la solicitud de patente estadounidense 2003/0069395 o Dennis et al. (Dennis et al. (2002) J. Biol. Chem. 277, 35035-35043).

El dominio 3 de la proteína G estreptocócica, tal como se describe por Kraulis et al., FEBS Lett. 378:190-194 (1996) y Linhult et al., Protein Sci. 11:206-213 (2002) es un ejemplo de un dominio de unión a albúmina bacteriano. Ejemplos de péptidos de unión a albúmina incluyen una serie de péptidos que tienen la secuencia central DICLPRWGCLW (SEQ ID NO: 35). Véase, p. ej., Dennis et al., J. Biol. Chem. 2002, 277: 35035-35043 (2002). Se describen ejemplos de fragmentos de anticuerpos que se unen a albúmina en Muller y Kontermann, Curr. Opin. Mol. Ther. 9:319-326 (2007); Roovers et al., Cancer Immunol. Immunother. 56:303-317 (2007) y Holt et al., Prot. Eng. Design Sci., 21:283-288 (2008). Un ejemplo de dicho resto de unión a albúmina de este tipo es hexanoato de 2-(3-maleimidopropanamido)-6-(4-(4-yodofenil) butanamido) ("etiqueta Albu") tal como se describe por Trussel et al., Bioconjugate Chem. 20:2286-2292 (2009).

Ácidos grasos, en particular los ácidos grasos de cadena larga (LCFA, por sus siglas en inglés) y compuestos de unión a albúmina de tipo ácido graso de cadena larga pueden utilizarse para extender la semivida *in vivo* de las proteínas FVIII de la divulgación. Un ejemplo de un compuesto de unión a albúmina similar a LCFA es ácido 16-(1-(3-(9-(((2,5-dioxopirrolidin-1-iloxi)carboniloxi)-metili)-7-sulfo-9H-fluoren-2-ilamino)-3-oxopropil)-2,5-dioxopirrolidin-3-iltio)hexadecanoico (véase, p. ej., el documento WO 2010/140148).

## 7. Secuencia PAS

En otras realizaciones, el resto heterólogo es una secuencia PAS. Una secuencia PAS, tal como se utiliza en esta memoria, significa una secuencia de aminoácidos que comprende principalmente residuos alanina y serina o que comprende principalmente residuos alanina, serina y prolina, formando la secuencia de aminoácidos una conformación de espiral aleatoria en condiciones fisiológicas. Por consiguiente, la secuencia PAS es un bloque de construcción, un polímero de aminoácidos o un casete de secuencia que comprende, consiste esencialmente en o consiste en alanina, serina y prolina que puede utilizarse como una parte del resto heterólogo en la proteína quimérica. Sin embargo, la persona experta es consciente de que un polímero de aminoácidos también puede formar una conformación de espiral aleatoria cuando se añaden residuos distintos de alanina, serina y prolina como un constituyente menor en la secuencia PAS. La expresión "constituyente menor", tal como se utiliza en esta memoria, significa que se pueden añadir aminoácidos distintos de alanina, serina y prolina en la secuencia PAS hasta cierto punto, p. ej., hasta aproximadamente el 12 %, es decir, aproximadamente 12 de 100 aminoácidos de la secuencia PAS, hasta aproximadamente el 10 %, es decir, aproximadamente 10 de 100 aminoácidos de la secuencia PAS, hasta aproximadamente el 9 %, es decir, aproximadamente 9 de 100 aminoácidos, hasta aproximadamente el 8 %, es decir, aproximadamente 8 de 100 aminoácidos, aproximadamente el 6%, es decir, aproximadamente 6 de 100 aminoácidos, aproximadamente el 5%, es decir, aproximadamente 5 de 100 aminoácidos, aproximadamente el 4%, es decir, aproximadamente 4 de 100 aminoácidos, aproximadamente el 3%, es decir, aproximadamente 3 de 100 aminoácidos, aproximadamente el 2%, es decir, aproximadamente 2 de 100 aminoácidos, aproximadamente el 1%, es decir, aproximadamente 1 de 100 de los aminoácidos. Los aminoácidos diferentes de alanina, serina y prolina pueden seleccionarse del grupo formado por Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Thr, Trp, Tyr y Val.

Bajo condiciones fisiológicas, la extensión de la secuencia PAS forma una conformación de espiral aleatoria y, por lo tanto, puede mediar en una estabilidad incrementada *in vivo* y/o *in vitro* para la proteína FVIII. Dado que el dominio de espiral aleatorio no adopta una estructura o función estable por sí mismo, la actividad biológica mediada por la proteína FVIII se conserva esencialmente. En otras realizaciones, las secuencias PAS que forman el dominio de espiral aleatoria son biológicamente inertes, especialmente con respecto a la proteólisis en el plasma sanguíneo, la inmunogenicidad, el punto isoelectrónico/comportamiento electrostático, la unión a los receptores de la superficie celular o la internalización, pero siguen siendo biodegradables, lo que proporciona claras ventajas frente a polímeros sintéticos tales como PEG.

Ejemplos no limitantes de las secuencias PAS que forman una conformación de espiral aleatoria comprenden una secuencia de aminoácidos seleccionada del grupo que consiste en ASPAAPAPASPAAPAPSAPA (SEQ ID NO: 36), AAPASPAPAPASAPAPAAPS (SEQ ID NO: 37), APSSPSPAPSSPSPASPSS (SEQ ID NO: 38), APSSPSPAPSSPSPASPS (SEQ ID NO: 39), SSPSPAPSSPSPASPSSPA (SEQ ID NO: 40), AASPAAPSAPAAASPAAPSAPPA (SEQ ID NO: 41) y ASAAAPAAASAAASAPSAAA (SEQ ID NO: 42) o cualesquiera combinaciones de las mismas. Se conocen ejemplos adicionales de secuencias PAS de, p. ej., la Publ. de la Pat. de EE.UU. N° 2010/0292130 A1 y la Publ. de la solicitud PCT. N° WO 2008/155134 A1.

## 8. Secuencia HAP

En determinadas realizaciones, el resto heterólogo es un polímero de homo-aminoácido (HAP, por sus siglas en inglés) rico en glicina. La secuencia HAP puede comprender una secuencia repetitiva de glicina, que tiene al menos 50 aminoácidos, al menos 100 aminoácidos, 120 aminoácidos, 140 aminoácidos, 160 aminoácidos, 180 aminoácidos, 200 aminoácidos, 250 aminoácidos, 300 aminoácidos, 350 aminoácidos, 400 aminoácidos, 450 aminoácidos o 500 aminoácidos de longitud. En una realización, la secuencia HAP es capaz de prolongar la semivida de un resto condensado o enlazado a la secuencia HAP. Ejemplos no limitantes de la secuencia HAP incluyen, pero no se limitan a (Gly)<sub>n</sub>, (Gly<sub>4</sub>Ser)<sub>n</sub> o S(Gly<sub>4</sub>Ser)<sub>n</sub>, en donde n es 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20. En una realización, n es 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40. En otra realización, n es 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190 o 200.

## 9. Transferrina o Fragmento de la misma

En determinadas realizaciones, el resto heterólogo es transferrina o un fragmento de la misma. Puede utilizarse cualquier transferrina para producir las proteínas FVIII de la divulgación. Como un ejemplo, el TF humano de tipo salvaje (TF) es una proteína de 679 aminoácidos, de aproximadamente 75 kDa (sin tener en cuenta la glicosilación), con dos dominios principales, N (aproximadamente 330 aminoácidos) y C (aproximadamente 340 aminoácidos), que parecen originarse a partir de una duplicación de genes. Véanse los números de acceso de GenBank NM001063, XM002793, M12530, XM039845, XM039847 y S95936 ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). La transferrina comprende dos dominios, el dominio N y el dominio C. El dominio N comprende dos subdominios, el dominio N1 y el dominio N2, y el dominio C comprende dos subdominios, el dominio C1 y el dominio C2.

En una realización, el resto heterólogo de transferrina incluye una variante de corte y empalme de transferrina. En un ejemplo, una variante de corte y empalme de transferrina puede ser una variante de corte y empalme de transferrina humana, *p. ej.*, acceso a Genbank AAA61140. En otra realización, la porción de transferrina de la proteína quimérica incluye uno o más dominios de la secuencia de transferrina, *p. ej.*, dominio N, dominio C, dominio N1, dominio N2, dominio C1, dominio C2 o cualquier combinación de los mismos.

## 10. Receptores de Aclaramiento

En determinadas realizaciones, el resto heterólogo es un receptor de aclaramiento, fragmento, variante o derivado del mismo. LRP1 es una proteína de membrana integral de 600 kDa que está implicada en el aclaramiento mediado por el receptor de una diversidad de proteínas, tales como el Factor X. Véase, *p. ej.*, Narita et al., Blood 91:555-560 (1998).

## 11. Factor de von Willebrand o Fragmentos del Mismo

En determinadas realizaciones, el resto heterólogo es el factor de von Willebrand (VWF) o uno o más fragmentos del mismo.

El VWF (también conocido como F8VWF) es una gran glicoproteína multimérica presente en el plasma sanguíneo y producida constitutivamente en el endotelio (en los cuerpos de Weibel -Palade), megacariocitos (gránulos de plaquetas) y tejido conjuntivo subendoteliano. El monómero básico del VWF es una proteína de 2813 aminoácidos. Cada uno de los monómeros contiene un cierto número de dominios específicos con una función específica, los dominios D' y D3 (que se unen juntos al Factor VIII), el dominio A1 (que se une al receptor plaquetario GPIb, heparina y/o posiblemente colágeno), el dominio A3 (que se une al colágeno), el dominio C1 (en el que el dominio RGD se une a la integrina plaquetaria αIIbβ3 cuando ésta se activa) y el dominio "nudo de cisteína" en el extremo C de la proteína (que VWF comparte con el factor de crecimiento derivado de plaquetas (PDGF), el factor de crecimiento transformante-β (TGFβ) y la gonadotropina coriónica humana β (βHCG)).

La secuencia de 2813 aminoácidos del monómero para el VWF humano se reseña como número de acceso NP000543.2 en Genbank. La secuencia de nucleótidos que codifica el VWF humano se reseña como número de acceso NM000552.3 en Genbank. SEQ ID NO: 44 (FIG. 11B) es la secuencia de aminoácidos codificada por SEQ ID NO: 43. El dominio D' incluye los aminoácidos 764 a 866 de SEQ ID NO: 44. El dominio D3 incluye los aminoácidos 867 a 1240 de SEQ ID NO: 44.

En el plasma, el 95-98 % del FVIII circula en un complejo no covalente compacto con el VWF de longitud completa. La formación de este complejo es importante para el mantenimiento de niveles plasmáticos apropiados de FVIII *in vivo*. Lenting et al., Blood. 92(11): 3983-96 (1998); Lenting et al., J. Thromb. Hemost. 5(7): 1353-60 (2007). Cuando el FVIII se activa debido a la proteólisis en las posiciones 372 y 740 de la cadena pesada y en la posición 1689 de la cadena ligera, el VWF unido al FVIII se elimina del FVIII activado.

En determinadas realizaciones, el resto heterólogo es factor de von Willebrand de longitud completa. En otras realizaciones, el resto heterólogo es un fragmento del factor de von Willebrand. Tal como se utiliza en esta memoria, la

expresión "fragmento de VWF" o "fragmentos de VWF" utilizada en esta memoria significa cualesquiera fragmentos de VWF que interactúen con FVIII y conserven al menos una o más propiedades que normalmente se proporcionan a FVIII por VWF de longitud completa, *p. ej.*, prevenir la activación prematura a FVIIIa, prevenir la proteólisis prematura, prevenir la asociación con membranas de fosfolípidos que podrían conducir a un aclaramiento prematuro, prevenir la unión a receptores de aclaramiento de FVIII que pueden unirse a FVIII desnudo pero no a FVIII unido a VWF, y/o estabilizar las Interacciones entre la cadena pesada y la cadena ligera de FVIII. En una realización específica, el resto heterólogo es un fragmento (VWF) que comprende un dominio D' y un dominio D3 de VWF. El fragmento de VWF que comprende el dominio D' y el dominio D3 puede comprender, además, un dominio de VWF seleccionado del grupo que consiste en un dominio A1, un dominio A2, un dominio A3, un dominio D1, un dominio D2, un dominio D4, un dominio B1, un dominio B2, un dominio B3, un dominio C1, un dominio C2, un dominio CK, uno o más fragmentos de los mismos, y cualquier combinación de los mismos. En la solicitud de patente provisional de EE.UU. nº 61/667.901, presentada el 3 de julio de 2012, y la Publicación de EE.UU. Nº 2015/0023959 A1 se describen ejemplos adicionales del polipéptido que tiene la actividad de FVIII condensado al fragmento de VWF.

## 12. Restos enlazadores

En determinadas realizaciones, el resto heterólogo es un enlazador peptídico.

Tal como se utiliza en esta memoria, las expresiones "enlazadores peptídicos" o "restos enlazadores" se refieren a una secuencia peptídica o polipeptídica (*p. ej.*, una secuencia peptídica o polipeptídica sintética) que conecta dos dominios en una secuencia lineal de aminoácidos de una cadena polipeptídica.

En algunas realizaciones, las secuencias de nucleótidos heterólogas que codifican enlazadores peptídicos pueden insertarse entre las secuencias de polinucleótidos de FVIII optimizadas de la divulgación y una secuencia de nucleótidos heteróloga que codifica, por ejemplo, uno de los restos heterólogos arriba descritos tales como la albúmina. Los enlazadores peptídicos pueden proporcionar flexibilidad a la molécula polipeptídica quimérica. Típicamente, los enlazadores no se escinden, sin embargo, una escisión de este tipo puede ser deseable. En una realización, estos enlazadores no se eliminan durante el procesamiento.

Un tipo de enlazador que puede estar presente en una proteína quimérica de la divulgación es un enlazador escindible por proteasa que comprende un sitio de escisión (*es decir*, un sustrato del sitio de escisión de proteasa, *p. ej.*, un sitio de escisión de factor Xla, Xa o trombina) y que pueden incluir enlazadores adicionales en el extremo N o en el extremo C o en ambos lados del sitio de escisión. Estos enlazadores escindibles, cuando se incorporan en una construcción de la divulgación, dan como resultado una molécula quimérica que tiene un sitio de escisión heterólogo.

En una realización, un polipéptido de FVIII codificado por una molécula de ácido nucleico de la presente divulgación comprende dos o más dominios Fc o restos enlazados a través de un enlazador cscFc para formar una región Fc comprendida en una única cadena polipeptídica. El enlazador cscFc está flanqueado por al menos un sitio de procesamiento intracelular, *es decir*, un sitio escindido por una enzima intracelular. La escisión del polipéptido en al menos un sitio de procesamiento intracelular da como resultado un polipéptido que comprende al menos dos cadenas polipeptídicas.

Opcionalmente, se pueden utilizar otros enlazadores peptídicos en una construcción de la divulgación, *p. ej.*, para conectar una proteína FVIII a una región Fc. Algunos enlazadores ejemplares que se pueden utilizar en relación con la divulgación incluyen, *p. ej.*, polipéptidos que comprenden aminoácidos GlySer descritos con más detalle más adelante.

En una realización, el enlazador peptídico es sintético, *es decir*, no se produce de forma natural. En una realización, un enlazador peptídico incluye péptidos (o polipéptidos) (que pueden o no ser naturales) que comprenden una secuencia de aminoácidos que enlaza o condensa genéticamente una primera secuencia lineal de aminoácidos con una segunda secuencia lineal de aminoácidos a la que no está enlazado de forma natural ni condensado genéticamente en la naturaleza. Por ejemplo, en una realización, el enlazador peptídico puede comprender polipéptidos que no se producen de forma natural que son formas modificadas de polipéptidos que se producen de forma natural (*p. ej.*, que comprenden una mutación tal como una adición, sustitución o delección). En otra realización, el enlazador peptídico puede comprender aminoácidos que no se producen de forma natural. En otra realización, el enlazador peptídico puede comprender aminoácidos que se producen de forma natural que se presentan en una secuencia lineal que no se presenta en la naturaleza. En aún otra realización, el enlazador peptídico puede comprender una secuencia polipeptídica que se produce de forma natural.

Por ejemplo, en determinadas realizaciones, se puede utilizar un enlazador peptídico para condensar restos Fc idénticos, formando así una región scFc homodimérica. En otras realizaciones, se puede utilizar un enlazador peptídico para condensar diferentes restos Fc (*p. ej.*, un resto Fc de tipo salvaje y una variante del resto Fc), formando así una región scFc heterodimérica.



- En otra realización, un enlazador peptídico comprende o consiste en un enlazador gly-ser. En una realización, un enlazador scFc o cscFc comprende al menos una porción de una bisagra de inmunoglobulina y un enlazador gly-ser. Tal como se utiliza en esta memoria, la expresión "enlazador gly-ser" se refiere a un péptido que consiste en residuos de glicina y serina. En determinadas realizaciones, dicho enlazador gly-ser se puede insertar entre otras dos secuencias del enlazador peptídico. En otras realizaciones, un enlazador gly-ser se fija a uno o ambos extremos de otra secuencia del enlazador peptídico. En aún otras realizaciones, se incorporan en serie dos o más enlazadores gly-ser en un enlazador peptídico. En una realización, un enlazador peptídico de la divulgación comprende al menos una porción de una región bisagra superior (*p. ej.*, derivada de una molécula IgG1, IgG2, IgG3 o IgG4), al menos una porción de una región bisagra media (*p. ej.*, derivado de una molécula de IgG1, IgG2, IgG3 o IgG4) y una serie de residuos de aminoácidos gly/ser.
- Enlazadores peptídicos de la divulgación tienen una longitud de al menos un aminoácido y pueden tener longitudes variables. En una realización, un enlazador peptídico de la divulgación tiene una longitud de aproximadamente 1 a aproximadamente 50 aminoácidos. Tal como se utiliza en este contexto, el término "aproximadamente" indica +/- dos residuos aminoácidos. Dado que la longitud del enlazador debe ser un número entero positivo, la longitud de aproximadamente 1 a aproximadamente 50 aminoácidos de longitud significa una longitud de 1-3 a 48-52 aminoácidos de longitud. En otra realización, un enlazador peptídico de la divulgación tiene una longitud de aproximadamente 10 a aproximadamente 20 aminoácidos. En otra realización, un enlazador peptídico de la divulgación tiene una longitud de aproximadamente 15 a aproximadamente 50 aminoácidos. En otra realización, un enlazador peptídico de la divulgación tiene una longitud de aproximadamente 20 a aproximadamente 45 aminoácidos. En otra realización, un enlazador peptídico de la divulgación tiene una longitud de aproximadamente 15 a aproximadamente 35 o de aproximadamente 20 a aproximadamente 30 aminoácidos. En otra realización, un enlazador peptídico de la divulgación es de aproximadamente 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, 40, 50, 60, 70, 80, 90, 100, 500, 1000 o 2000 aminoácidos de longitud. En una realización, un enlazador peptídico de la divulgación tiene una longitud de 20 o 30 aminoácidos.
- En algunas realizaciones, el enlazador peptídico puede comprender al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos 10, al menos 20, al menos 30, al menos 40, al menos 50, al menos 60, al menos 70, al menos 80, al menos 90 o al menos 100 aminoácidos. En otras realizaciones, el enlazador peptídico puede comprender al menos 200, al menos 300, al menos 400, al menos 500, al menos 600, al menos 700, al menos 800, al menos 900 o al menos 1000 aminoácidos. En algunas realizaciones, el enlazador peptídico puede comprender al menos aproximadamente 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900 o 2000 aminoácidos. El enlazador peptídico puede comprender 1-5 aminoácidos, 1-10 aminoácidos, 1-20 aminoácidos, 10-50 aminoácidos, 50-100 aminoácidos, 100-200 aminoácidos, 200-300 aminoácidos, 300-400 aminoácidos, 400-500 aminoácidos, 500-600 aminoácidos, 600-700 aminoácidos, 700-800 aminoácidos, 800-900 aminoácidos o 900-1000 aminoácidos.
- Enlazadores peptídicos pueden introducirse en secuencias polipeptídicas utilizando técnicas conocidas en la técnica. Modificaciones pueden confirmarse mediante análisis de secuencia de ADN. El ADN de plásmido puede utilizarse para transformar células huésped para la producción estable de los polipéptidos producidos.

### Híbridos de Monómero-Dímero

- En algunas realizaciones, las moléculas de ácido nucleico aisladas de la divulgación que, además, comprenden una secuencia de nucleótidos heteróloga codifican una molécula híbrida de monómero-dímero que comprende FVIII.
- La expresión "híbrido de monómero-dímero" utilizada en esta memoria se refiere a una proteína quimérica que comprende una primera cadena polipeptídica y una segunda cadena polipeptídica, que están asociadas entre sí mediante un enlace disulfuro, en donde la primera cadena comprende Factor VIII y una primera región Fc y la segunda cadena comprende, consiste esencialmente en o consiste en una segunda región Fc sin el FVIII. La construcción híbrida de monómero-dímero es, por lo tanto, un híbrido que comprende un aspecto de monómero que tiene solo un factor de coagulación y un aspecto de dímero que tiene dos regiones Fc.

### Elemento de Control de la Expresión

- En algunas realizaciones, la molécula de ácido nucleico o el vector de la divulgación comprenden, además, al menos una secuencia de control de la expresión. Una secuencia de control de la expresión, tal como se utiliza en esta memoria, es cualquier secuencia de nucleótidos reguladora, tal como una secuencia de promotor o una combinación de promotor-potenciador, que facilita la transcripción y traducción eficientes del ácido nucleico codificante al que está enlazado operativamente. Por ejemplo, la molécula de ácido nucleico aislada de la divulgación se puede enlazar operativamente a al menos una secuencia de control de la transcripción. La secuencia de control de la expresión génica puede ser, por ejemplo, un promotor de mamífero o viral, tal como un promotor constitutivo o inducible. Promotores de mamíferos constitutivos incluyen, pro no se limitan a los promotores de los siguientes genes: hipoxantina fosforribosil transferasa (HPRT, por sus siglas en inglés), adenosina desaminasa, piruvato quinasa, promotor de beta-actina y otros promotores constitutivos. Promotores virales ejemplares que funcionan de forma constitutiva en células eucarióticas incluyen, por ejemplo, promotores del citomegalovirus (CMV), virus de simio (*p. ej.*, SV40), virus del papiloma, adenovirus, virus de la

inmunodeficiencia humana (VIH), virus del sarcoma de Rous, citomegalovirus, las repeticiones terminales largas (LTR, por sus siglas en inglés) del virus de la leucemia de Moloney y otros retrovirus, y el promotor de la timidina quinasa del virus del herpes simple. Los expertos ordinarios en la técnica conocen otros promotores constitutivos. Los promotores útiles como secuencias de expresión génica de la divulgación también incluyen promotores inducibles. Los promotores inducibles se expresan en presencia de un agente inductor. Por ejemplo, se induce al promotor de la metalotioneína para que fomente la transcripción y la traducción en presencia de determinados iones metálicos. Los expertos ordinarios en la técnica conocen otros promotores inducibles.

En una realización, la divulgación incluye la expresión de un transgén bajo el control de un promotor y/o potenciador específico para tejidos. En otra realización, el promotor u otra secuencia de control de la expresión potencian selectivamente la expresión del transgén en células hepáticas. Ejemplos de promotores específicos para el hígado incluyen, pero no se limitan a un promotor de la tiretina de ratón (mTTR, por sus siglas en inglés), un promotor del factor VIII humano endógeno (F8), un promotor de alfa-1-antitripsina humana (hAAT, por sus siglas en inglés), un promotor mínimo de albúmina humana y un promotor de albúmina de ratón. En una realización particular, el promotor comprende un promotor mTTR. El promotor mTTR se describe en R.H. Costa et al., 1986, Mol. Cell. Biol. 6:4697. El promotor F8 se describe en Figueiredo y Brownlee, 1995, J. Biol. Chem. 270: 11828-11838.

Los niveles de expresión se pueden potenciar adicionalmente para lograr la eficacia terapéutica utilizando uno o más potenciadores. Se pueden proporcionar uno o más potenciadores solos o junto con uno o más elementos promotores. Típicamente, la secuencia de control de la expresión comprende una pluralidad de elementos potenciadores y un promotor específico para tejidos. En una realización, un potenciador comprende una o más copias del potenciador de  $\alpha$ -1-microglobulina/bikunina (Rouet et al., 1992, J. Biol. Chem. 267:20765-20773; Rouet et al., 1995, Nucleic Acids Res. 23:395-404, Rouet et al., 1998, Biochem. J. 334:577-584, Ill et al., 1997, Blood Coagulation Fibrinolysis 8:S23-S30). En otra realización, un potenciador se deriva de los sitios de unión al factor de transcripción específico del hígado, tales como EBP, DBP, HNF1, HNF3, HNF4, HNF6, comprendiendo Enh1 HNF1, HNF3-(sentido), HNF4-(sentido), HNF1-(antisentido), HNF6-(antisentido), EBP-(sentido), HNF4-(antisentido) (antisentido).

En un ejemplo particular, un promotor útil para la divulgación comprende SEQ ID NO: 69 (es decir, promotor ET; FIG. 11Y), que también se conoce como GenBank N° AY661265. Véase también Vigna et al., Molecular Therapy 11(5):763 (2005). Ejemplos de otros vectores y elementos reguladores de genes adecuados se describen en los documentos WO 02/092134, EP1395293 o las Patentes de EE.UU. N°s 6.808.905, 7.745.179 o 7.179.903.

En general, las secuencias de control de la expresión incluirán, según sea necesario, secuencias 5' que no se transcriben y 5' que no se traducen involucradas con el inicio de la transcripción y la traducción, respectivamente, tal como una caja TATA, secuencia de cubierta, secuencia CAAT y similares. Especialmente, secuencias 5' no transcripcionales de este tipo incluirán una región de promotor que incluye una secuencia promotora para el control transcripcional del ácido nucleico codificante unido operativamente. Las secuencias de expresión génica incluyen opcionalmente secuencias de potenciador o secuencias de activador aguas arriba, según se desee.

## Sistemas de Vectores

La presente invención describe un vector que comprende la molécula de ácido nucleico como se define en las reivindicaciones adjuntas, una célula huésped que comprende el vector como se define en las reivindicaciones adjuntas y un vector como se define en las reivindicaciones adjuntas para uso en un método de tratamiento de un trastorno de sangrado. La presente divulgación satisface una necesidad importante en la técnica al proporcionar un vector que comprende una secuencia de FVIII optimizada que demuestra una expresión incrementada en un sujeto y potencialmente da como resultado una mayor eficacia terapéutica cuando se utiliza en métodos de terapia génica.

En algunas realizaciones, la presente divulgación proporciona un vector que comprende una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende una secuencia de ácido nucleico que tiene al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 % o al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71 y está enlazado operativamente a un promotor, una secuencia diana o ambos. En otras realizaciones, la secuencia de ácido nucleico comprende (i) los nucleótidos 58-4374 de SEQ ID NO: 71 o (ii) los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71.

Vectores adecuados para la divulgación incluyen vectores de expresión, vectores virales y vectores de plásmidos. En una realización, el vector es un vector viral.

Tal como se utiliza en esta memoria, un vector de expresión se refiere a cualquier construcción de ácido nucleico que contiene los elementos necesarios para la transcripción y traducción de una secuencia codificante insertada, o en el caso de un vector viral de ARN, los elementos necesarios para la replicación y traducción, cuando se introduce en una célula huésped adecuada. Vectores de expresión pueden incluir plásmidos, fagémidos, virus y derivados de los mismos.

5 Vectores de expresión de la divulgación incluirán polinucleótidos optimizados que codifican la proteína BDD FVIII descrita en esta memoria. En una realización, las secuencias codificantes optimizadas para la proteína BDD FVIII se enlazan operativamente a una secuencia de control de la expresión. Tal como se utiliza en esta memoria, dos secuencias de ácido nucleico se enlazan operativamente cuando se enlazan covalentemente de tal manera que permita que cada secuencia de ácido nucleico componente conserve su funcionalidad. Se dice que una secuencia codificante y una secuencia de control de la expresión génica están enlazadas operativamente cuando están enlazadas covalentemente de tal manera que colocan la expresión o transcripción y/o traducción de la secuencia codificante bajo la influencia o el control de la secuencia de control de la expresión génica. Se dice que dos secuencias de ADN están enlazadas operativamente si la inducción de un promotor en la secuencia de expresión del gen 5' da como resultado la transcripción de la secuencia codificante y si la naturaleza del enlace entre las dos secuencias de ADN no (1) da como resultado la introducción de una mutación de desplazamiento de marco, (2) interfieren con la capacidad de la región de promotor para dirigir la transcripción de la secuencia codificante o (3) interfieren con la capacidad de la transcripción de ARN correspondiente para traducirse en una proteína. Por lo tanto, una secuencia de expresión génica estaría enlazada operativamente a una secuencia de ácido nucleico codificante si la secuencia de expresión génica fuera capaz de efectuar la transcripción de esa secuencia de ácido nucleico codificante de modo que el transcrito resultante se traduzca en la proteína o polipéptido deseado.

20 Vectores virales incluyen, pero no se limitan a secuencias de ácido nucleico de los siguientes virus: retrovirus, tales como el virus de la leucemia murina de Moloney, el virus del sarcoma murino de Harvey, el virus del tumor mamario murino y el virus del sarcoma de Rous; lentivirus; adenovirus; virus adeno-asociados; virus tipo SV40; poliomavirus; virus de Epstein-Barr; virus del papiloma; virus del herpes; virus vaccinia; virus de la poliomiéltis; y virus de ARN tal como un retrovirus. Se pueden emplear fácilmente otros vectores bien conocidos en la técnica. Determinados vectores virales se basan en virus eucarióticos no citopáticos en los que los genes no esenciales han sido reemplazados por el gen de interés. Virus no citopáticos incluyen retrovirus, cuyo ciclo de vida implica la transcripción inversa del ARN viral genómico en ADN con la integración proviral subsiguiente en el ADN celular del huésped. Los retrovirus han sido aprobados para ensayos de terapia génica humana. Los más útiles son los retrovirus que tienen una replicación deficiente (*es decir*, son capaces de dirigir la síntesis de las proteínas deseadas, pero incapaces de fabricar una partícula infecciosa). Vectores de expresión retroviral alterados genéticamente de este tipo tienen una utilidad general para la transducción de genes de alta eficacia *in vivo*. Protocolos estándares para producir retrovirus con replicación deficiente (incluyendo las etapas de incorporación de material genético exógeno en un plásmido, transfección de una línea celular empaquetadora con plásmido, producción de retrovirus recombinantes por la línea celular empaquetadora, recogida de partículas virales de medios de cultivo de tejidos e infección de las células diana con partículas víricas) se proporcionan en Kriegler, M., Gene Transfer and Expression, A Laboratory Manual, W.H. Freeman Co., Nueva York (1990) y Murry, E.J., Methods in Molecular Biology, vol. 7, Humana Press, Inc., Clifton, Nueva Jersey (1991).

35 En una realización, el virus es un virus adeno-asociado, un virus de ADN de doble cadena. El virus adeno-asociado puede diseñarse para que sea deficiente en replicación y sea capaz de infectar una amplia gama de tipos de células y especies. Tiene, además, ventajas tales como la estabilidad frente al calor y los disolventes de lípidos; altas frecuencias de transducción en células de diversos linajes, incluyendo células hematopoyéticas; y falta de inhibición de superinfección, lo que permite múltiples series de transducciones. Según se informa, el virus adeno-asociado puede integrarse en el ADN celular humano de una manera específica para el sitio, minimizando así la posibilidad de mutagénesis por inserción y la variabilidad de la expresión génica insertada característica de la infección retroviral. Además, las infecciones por virus adeno-asociados de tipo salvaje se han seguido en cultivos de tejidos durante más de 100 pases en ausencia de presión selectiva, lo que implica que la integración genómica del virus adeno-asociado es un evento relativamente estable. El virus adeno-asociado también puede funcionar de forma extracromosómica.

45 En otra realización, el vector viral es un virus adeno-asociado (AAV, por sus siglas en inglés) que ha sido manipulado para transportar un polinucleótido que codifica una proteína FVIII como se describe en esta memoria. Se han descrito métodos generales para obtener AAV recombinantes (rAAV). Véanse, por ejemplo, los documentos USP 8.734.809, 2013/0195801 así como las referencias allí citadas. En algunas realizaciones, un vector rAAV comprende una o más repeticiones terminales invertidas (ITRs, por sus siglas en inglés) de AAV y un transgén de interés (p. ej., una secuencia de polinucleótidos de FVIII optimizada). En determinadas realizaciones, los métodos para hacer rAAV implican cultivar una célula huésped deseada que contiene una secuencia de ácido nucleico que codifica una proteína de la cápside de AAV o un fragmento de la misma; un gen rep funcional; un vector rAAV compuesto por repeticiones terminales invertidas (ITRs) de AAV y un transgén de interés; y suficientes funciones auxiliares para permitir el empaquetamiento del vector AAV recombinante en las proteínas de la cápside de AAV. Se han descrito materiales y métodos para realizar estos procedimientos y otros relacionados, por ejemplo, en los documentos USP 8.734.809, 2013/0195801, PCT/US1997/015692, PCT/US2002/033692, PCT/US2002/033630, WO2007/148971, WO00/20561, WO03/042361 y WO2007/04670.

60 Una o más secuencias de vectores AAV diferentes derivadas de casi cualquier serotipo pueden utilizarse de acuerdo con la presente divulgación. La elección de una secuencia de vector AAV particular estará guiada por parámetros conocidos tal como el tropismo de interés, los rendimientos de vector requeridos, etc. Generalmente, los serotipos AAV tienen secuencias genómicas de homología significativa en los niveles de aminoácidos y ácidos nucleicos, proporcionan un conjunto relacionado de las funciones genéticas, producen viriones que están relacionados, y se replican y ensamblan de manera similar. Para la secuencia genómica de los diversos serotipos de AAV y una visión general de las similitudes

- genómicas, véase, *p. ej.*, número de acceso de GenBank U89790; número de acceso de GenBank J01901; número de acceso de GenBank AF043303; número de acceso de GenBank AF085716; Chlorini et al. (1997, J. Vir. 71: 6823-33); Srivastava et al. (1983, J. Vir. 45:555-64); Chlorini et al. (1999, J. Vir. 73:1309-1319); Rutledge et al. (1998, J. Vir. 72:309-319); y Wu et al. (2000, J. Vir. 74: 8635-47). Los serotipos 1, 2, 3, 4 y 5 de AAV son una fuente ilustrativa de secuencias de nucleótidos de AAV para uso en el contexto de la presente divulgación. AAV6, AAV7, AAV8 o AAV9 o partículas similares a AAV recientemente desarrolladas, obtenidas, *p. ej.*, mediante técnicas de reordenamiento aleatorio de la cápside y colecciones de cápsides de AAV, o de ITRs recientemente diseñadas, desarrolladas o evolucionadas, también son adecuadas para determinadas aplicaciones de divulgación. Véase Dalkara, D et al. (2013), Sci. Transl. Med. 5(189): 189ra76; Kotterman, MA Nat. Rev. Genet. (2014) 15(7):455.
- 5 Sin embargo, en determinadas realizaciones, los vectores AAV con un tropismo significativo hacia el hígado y tejidos relacionados serán de interés para expresar las proteínas FVIII descritas en esta memoria. Ejemplos no limitativos incluyen los serotipos 1, 2, 6 y 8 de AAV. Véase, *p. ej.*, Torres-Toranteras et al. (2014) 22: 901 y referencias allí citadas.
- En otras realizaciones, el vector se deriva de lentivirus. En determinadas realizaciones, el vector es un vector de un lentivirus recombinante capaz de infectar células que no se dividen.
- 15 El genoma lentiviral y el ADN proviral tienen típicamente los tres genes que se encuentran en los retrovirus: gag, pol y env, que están flanqueados por dos secuencias de repetición terminal larga (LTR). El gen gag codifica las proteínas estructurales internas (matriz, cápside y nucleocápside); el gen pol codifica la ADN polimerasa dirigida por ARN (transcriptasa inversa), una proteasa y una integrasa; y el gen env codifica glicoproteínas de la envoltura viral. Los LTRs 5' y 3' sirven para fomentar la transcripción y poliadenilación de los ARN del virión. La LTR contiene todas las demás
- 20 secuencias que actúan en cis necesarias para la replicación viral. Los lentivirus tienen genes adicionales que incluyen vif, vpr, tat, rev, vpu, nef y vpx (en HIV-1, HIV-2 y/o SIV).
- Adyacentes a la LTR 5' hay secuencias necesarias para la transcripción inversa del genoma (el sitio de unión del cebador de ARNt) y para la encapsidación eficiente del ARN viral en partículas (el sitio Psi). Si las secuencias necesarias para la encapsidación (o empaquetamiento del ARN retroviral en viriones infecciosos) faltan en el genoma viral, el defecto cis evita la encapsidación del ARN genómico.
- 25 Sin embargo, el mutante resultante sigue siendo capaz de dirigir la síntesis de todas las proteínas del virión. La divulgación proporciona un método para producir un lentivirus recombinante capaz de infectar una célula que no se divide, que comprende transfectar una célula huésped adecuada con dos o más vectores que portan las funciones de empaquetamiento, a saber, gag, pol y env, así como rev y tat. Como se describirá más adelante, los vectores que carecen de un gen tat funcional son deseables para determinadas aplicaciones. Así, por ejemplo, un primer vector puede proporcionar un ácido nucleico que codifica una gag viral y una pol viral y otro vector puede proporcionar un ácido nucleico que codifica una env viral para producir una célula de empaquetamiento. La introducción de un vector que proporciona un gen heterólogo, identificado aquí como un vector de transferencia, en esa célula de empaquetamiento produce una célula productora que libera partículas virales infecciosas que portan el gen extraño de interés.
- 30 De acuerdo con la configuración de vectores y genes extraños arriba indicada, el segundo vector puede proporcionar un ácido nucleico que codifica un gen de la cubierta viral (env). El gen env se puede derivar de casi cualquier virus adecuado, incluyendo los retrovirus. En algunas realizaciones, la proteína env es una proteína de envoltura anfotrópica que permite la transducción de células humanas y de otras especies.
- Ejemplos de genes env derivados de retrovirus incluyen, pero no se limitan a: virus de la leucemia murina de Moloney (MoMuLV o MMLV), virus del sarcoma murino de Harvey (HaMuSV o HSV), virus del tumor mamario murino (MuMTV o MMTV), virus de la leucemia del mono gibbon (GaLV o GALV), virus de la inmunodeficiencia humana (VIH) y virus del sarcoma de Rous (RSV). También se pueden utilizar otros genes env, tales como la proteína G (VSV G) del virus de la estomatitis vesicular (VSV), la de los virus de la hepatitis y la de la gripe.
- 40 El vector que proporciona la secuencia de ácido nucleico de env viral está asociado operativamente con secuencias reguladoras descritas en otra parte de esta memoria.
- En determinadas realizaciones, el vector incluye un vector lentiviral en el que se eliminaron los genes de virulencia del VIH env, vif, vpr, vpu y nef sin comprometer la capacidad del vector de transducir células que no se dividen.
- En algunas realizaciones, el vector incluye un vector lentiviral que comprende una delección de la región U3 de la LTR 3'. La delección de la región U3 puede ser una delección completa o una delección parcial.
- 50 En algunas realizaciones, el vector lentiviral de la divulgación que comprende la secuencia de nucleótidos de FVIII descrita en esta memoria puede transfectarse en una célula con (a) una primera secuencia de nucleótidos que comprende un gen gag, un gen pol o genes gag y pol y (b) una segunda secuencia de nucleótidos que comprende un gen env heterólogo; en donde el vector lentiviral carece de un gen tat funcional. En otras realizaciones, la célula se transfecta adicionalmente con

una cuarta secuencia de nucleótidos que comprende un gen rev. En determinadas realizaciones, el vector lentiviral carece de genes funcionales seleccionados de vif, vpr, vpu, vpx y nef, o una combinación de los mismos.

En determinadas realizaciones, un vector lentiviral comprende una o más secuencias de nucleótidos que codifican una proteína gag, un elemento de respuesta a Rev, una pista de polipurina central (cPPT, por sus siglas en inglés) o cualquier combinación de los mismos.

Se describen ejemplos de los vectores lentivirales en los documentos WO9931251, WO9712622, WO9817815, WO9817816 y WO981893.

Otros vectores incluyen vectores de plásmidos. Los vectores de plásmidos se han descrito extensamente en la técnica y son bien conocidos por los expertos en la técnica. Véase, p. ej., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, segunda edición, Cold Spring Harbor Laboratory Press, 1989. En los últimos años, se ha descubierto que los vectores de plásmidos son particularmente ventajosos para suministrar genes a las células *in vivo* debido a su incapacidad de replicarse e integrarse en un genoma huésped. Estos plásmidos, sin embargo, que tienen un promotor compatible con la célula huésped, pueden expresar un péptido de un gen codificado operativamente dentro del plásmido. Algunos plásmidos de uso común disponibles de proveedores comerciales incluyen pBR322, pUC18, pUC19, diversos plásmidos pcDNA, pRC/CMV, diversos plásmidos pCMV, pSV40 y pBlueScript. Ejemplos adicionales de plásmidos específicos incluyen pcDNA3.1, número de catálogo V79020; pcDNA3.1/hygro, número de catálogo V87020; pcDNA4/myc-His, número de catálogo V86320; y pBudCE4.1, número de catálogo V53220, todos de Invitrogen (Carlsbad, CA). Otros plásmidos son bien conocidos por los expertos en la técnica. Adicionalmente, los plásmidos se pueden diseñar a la medida utilizando técnicas estándares de biología molecular para eliminar y/o añadir fragmentos específicos de ADN.

## Expresión Específica de Tejido

En determinadas realizaciones, será útil incluir dentro del vector una o más secuencias diana de miARN que, por ejemplo, estén enlazadas operativamente al transgén de FVIII optimizado. Por lo tanto, la divulgación también proporciona al menos una secuencia diana de miARN enlazada operativamente a la secuencia de nucleótidos de FVIII optimizada o insertada de otro modo dentro de un vector. Más de una copia de una secuencia diana de miARN incluida en el vector puede aumentar la eficacia del sistema. También se incluyen diferentes secuencias diana de miARN. Por ejemplo, los vectores que expresan más de un transgén pueden tener el transgén bajo el control de más de una secuencia diana de miARN, que puede ser igual o diferente. Las secuencias diana de miARN pueden estar en tándem, pero también se incluyen otras disposiciones. El casete de expresión transgénica, que contiene secuencias diana de miARN, también se puede insertar dentro del vector en orientación antisentido. La orientación antisentido puede ser útil en la producción de partículas virales para evitar la expresión de productos génicos que, de lo contrario, pueden ser tóxicos para las células productoras. En otras realizaciones, el vector comprende 1, 2, 3, 4, 5, 6, 7 u 8 copias de la misma o diferente secuencia diana de miARN. Sin embargo, en determinadas otras realizaciones, el vector no incluirá secuencia diana alguna de miARN. La opción de incluir o no una secuencia diana de miARN (y cuántas) se guiará por parámetros conocidos, tales como el tejido objetivo previsto, el nivel de expresión requerido, etc.

En una realización, la secuencia diana es una diana de miR-223 que se ha reseñado que bloquea la expresión más eficazmente en progenitores comprometidos mieloides y al menos parcialmente en el HSPC más primitivo. La diana de miR-223 puede bloquear la expresión en células mieloides diferenciadas, incluyendo granulocitos, monocitos, macrófagos, células dendríticas mieloides. La diana miR-223 también puede ser adecuada para aplicaciones de terapia génica que se basan en una expresión transgénica robusta en el linaje linfóide o eritroide. La diana de miR-223 también puede bloquear la expresión de manera muy efectiva en HSC humana.

En otra realización, la secuencia diana es una diana miR142 (tccataaagt aggaaaacact aca (SEQ ID NO: 43)). En una realización, el vector comprende 4 copias de secuencias diana de miR-142. En determinadas realizaciones, la secuencia complementaria de microARNs hematopoyéticos específicos tal como miR-142 (142T) se incorpora en la región 3' no traducida de un vector, p. ej., vectores lentivirales (LV), lo que hace que el transcrito que codifica el transgén sea susceptible de regulación a la baja mediada por miARN. Mediante este método, la expresión transgénica se puede prevenir en células presentadoras de antígenos de linaje hematopoyético (APC, por sus siglas en inglés), mientras que se mantiene en células no hematopoyéticas (Brown et al., *Nat Med* 2006). Esta estrategia puede imponer un estricto control post-transcripcional sobre la expresión transgénica y, por lo tanto, permite el suministro estable y la expresión a largo plazo de los transgenes. En algunas realizaciones, la regulación de miR-142 evita el aclaramiento inmuno-mediado de células transducidas y/o induce a células T reguladoras específicas para el antígeno (T regs) y media en una tolerancia inmunológica robusta al antígeno codificado por transgén.

En algunas realizaciones, la secuencia diana es una diana miR181. Chen C-Z y Lodish H, *Seminars in Immunology* (2005) 17(2):155-165 describe miR-181, un miARN expresado específicamente en células B dentro de la médula ósea de ratón (Chen y Lodish, 2005). También describe que algunos miARN humanos están ligados a las leucemias.

La secuencia diana puede ser total o parcialmente complementaria al miARN. La expresión "totalmente complementaria" significa que la secuencia diana tiene una secuencia de ácido nucleico que es 100 % complementaria a la secuencia del miARN que la reconoce. La expresión "parcialmente complementaria" significa que la secuencia diana es solo en parte complementaria a la secuencia del miARN que la reconoce, por lo que el miARN sigue reconociendo la secuencia parcialmente complementaria. En otras palabras, una secuencia diana parcialmente complementaria en el contexto de la presente divulgación es eficaz para reconocer el miARN correspondiente y efectuar la prevención o reducción de la expresión transgénica en células que expresan ese miARN. Ejemplos de secuencias diana de miARN se describen en los documentos WO2007/000668, WO2004/094642, WO2010/055413 o WO2010/125471.

### Células Huésped

La presente invención describe una célula huésped que comprende la molécula de ácido nucleico o el vector como se define en las reivindicaciones adjuntas. Tal como se utiliza en esta memoria, el término "transformación" se utilizará en un sentido amplio para referirse a la introducción de ADN en una célula huésped receptora que cambia el genotipo y, en consecuencia, da como resultado un cambio en la célula receptora.

"Células huésped" se refiere a células que han sido transformadas con vectores construidos utilizando técnicas de ADN recombinante y que codifican al menos un gen heterólogo. Las células huésped de la presente divulgación son preferentemente de origen mamífero; lo más preferiblemente de origen humano o de ratón. A los expertos en la técnica se les atribuye la capacidad de determinar preferentemente líneas de células huésped particulares que son las más adecuadas para su propósito. Líneas de células huésped ejemplares incluyen, pero no se limitan a CHO, DG44 y DUXB11 (líneas de ovario de hámster chino, DHFR menos), HELA (carcinoma cervical humano), CVI (línea de riñón de mono), COS (un derivado de CVI con antígeno T SV40), R1610 (fibroblasto de hámster chino), BALBC/3T3 (fibroblasto de ratón), HAK (línea de riñón de hámster), SP2/O (mieloma de ratón), P3.x.63-Ag3.653 (mieloma de ratón), BFA-1c1BPT (células endoteliales bovinas), RAJI (linfocitos humanos), PER.C6<sup>®</sup>, NS0, CAP, BHK21 y HEK 293 (riñón humano). En una realización particular, la célula huésped se selecciona del grupo que consiste en: una célula CHO, una célula HEK293, una célula BHK21, una célula PER.C6<sup>®</sup>, una célula NS0 y una célula CAP. Líneas de células huésped están típicamente disponibles de los servicios comerciales, la Colección Americana de Cultivos de Tejidos o en la bibliografía publicada.

La introducción de las moléculas de ácido nucleico aisladas o los vectores de la divulgación en la célula huésped se puede lograr mediante diversas técnicas bien conocidas por los expertos en la técnica. Estas incluyen, pero no se limitan a transfección (incluyendo electroforesis y electroporación), fusión de protoplastos, precipitación con fosfato de calcio, fusión celular con ADN envuelto, microinyección e infección con virus intacto. Véase, Ridgway, AAG "Mammalian Expression Vectors" Capítulo 24.2, pp. 470-472 Vectors, Rodríguez y Denhardt, Eds. (Butterworths, Boston, Mass. 1988). Lo más preferiblemente, la introducción del plásmido en el huésped se realiza mediante electroporación. Las células transformadas se cultivan en condiciones apropiadas para la producción de cadenas ligeras y cadenas pesadas, y se analizan para determinar la síntesis de proteínas de cadenas pesadas y/o ligeras. Técnicas de ensayo ejemplares incluyen ensayo inmunoabsorbente ligado a enzimas (ELISA), radioinmunoensayo (RIA) o análisis de clasificación de células activadas por fluorescencia (FACS), inmunohistoquímica y similares.

Células huésped que comprenden las moléculas de ácido nucleico aisladas o los vectores de la divulgación se cultivan en un medio de crecimiento apropiado. Tal como se utiliza en esta memoria, la expresión "medio de crecimiento apropiado" significa un medio que contiene los nutrientes necesarios para el crecimiento de las células. Los nutrientes necesarios para el crecimiento celular pueden incluir una fuente de carbono, una fuente de nitrógeno, aminoácidos esenciales, vitaminas, minerales y factores de crecimiento. Opcionalmente, los medios pueden contener uno o más factores de selección. Opcionalmente, los medios pueden contener suero de ternero bovino o suero de ternero fetal (FCS, por sus siglas en inglés). En una realización, el medio no contiene sustancialmente IgG. El medio de crecimiento generalmente seleccionará las células que contienen la construcción de ADN, por ejemplo, por selección de fármacos o deficiencia en un nutriente esencial que se complementa con el marcador seleccionable en la construcción de ADN o se cotransfecta con la construcción de ADN. Las células de mamífero cultivadas se cultivan generalmente en medios que contienen suero o libres de suero comercialmente disponibles (*p. ej.*, MEM, DMEM, DMEM/F12). En una realización, el medio es CDoptiCHO (Invitrogen, Carlsbad, CA.). En otra realización, el medio es CD17 (Invitrogen, Carlsbad, CA.). La selección de un medio apropiado para la línea celular particular utilizada está dentro del nivel de los expertos ordinarios en la técnica.

### Preparación de Polipéptidos

La divulgación también proporciona un polipéptido codificado por una molécula de ácido nucleico de la divulgación. En otras realizaciones, el polipéptido de la divulgación es codificado por un vector que comprende las moléculas nucleicas aisladas de la divulgación. En aún otras realizaciones, el polipéptido de la divulgación es producido por una célula huésped que comprende las moléculas nucleicas aisladas de la divulgación.

La presente invención proporciona un método para producir un polipéptido con actividad FVIII, que comprende: cultivar la célula huésped como se define en las reivindicaciones adjuntas en condiciones en las que se produce un polipéptido con actividad de FVIII y recuperar el polipéptido con actividad de FVIII. En algunas realizaciones, la expresión del polipéptido con actividad de FVIII aumenta con respecto a una célula huésped cultivada en las mismas condiciones, pero que

comprende una secuencia de nucleótidos de referencia que comprende la SEQ ID NO: 16, la secuencia del gen FVIII original.

5 En otras realizaciones, la divulgación proporciona un método para aumentar la expresión de un polipéptido con actividad de FVIII, que comprende cultivar una célula huésped de la divulgación en condiciones en las que la molécula de ácido nucleico expresa un polipéptido con actividad de FVIII, en el que la expresión del polipéptido con actividad de FVIII aumenta en relación con una célula huésped cultivada en las mismas condiciones que comprende una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16.

10 En otras realizaciones, la divulgación proporciona un método para mejorar el rendimiento de un polipéptido con actividad de FVIII, que comprende cultivar una célula huésped en condiciones en las que la molécula de ácido nucleico produce un polipéptido con actividad de FVIII, en el que el rendimiento del polipéptido con actividad de FVIII se incrementa en relación con una célula huésped cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende SEQ ID NO: 16.

15 En otras realizaciones, la divulgación proporciona un método para mejorar el rendimiento de un polipéptido con actividad de FVIII, que comprende cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica el polipéptido, en el que se incrementa el índice de adaptación de codones de una porción 3' de la secuencia de nucleótidos con respecto a una porción 5' de la secuencia de nucleótidos; en el que el rendimiento del polipéptido con actividad de FVIII aumenta en relación con una célula huésped cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16. En algunas realizaciones, el índice de adaptación del codón de la porción 5' de la secuencia de nucleótido aumenta, disminuye o no cambia en relación con el índice de optimización por codones de SEQ ID NO: 16.

20 En otras realizaciones, la divulgación proporciona un método para mejorar el rendimiento de un polipéptido con actividad de FVIII, que comprende cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica el polipéptido, en donde se incrementa el índice de adaptación de codones de una porción 5' de la secuencia de nucleótidos con respecto a una porción 3' de la secuencia de nucleótidos; en el que el rendimiento de polipéptido con actividad de FVIII aumenta con respecto a una célula huésped cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende SEQ ID NO: 16. En algunas realizaciones, el índice de adaptación de codones de la porción 3' de la secuencia de nucleótidos aumenta, disminuye o no cambia en relación con el índice de optimización de codones de SEQ ID NO: 16.

30 En otras realizaciones, la divulgación proporciona un método para mejorar el rendimiento de un polipéptido con actividad de FVIII, que comprende cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica el polipéptido, en el que el índice de adaptación de codones de una porción del nucleótido que codifica una porción C-terminal del polipéptido aumenta con respecto a la porción del nucleótido que codifica una porción N-terminal del polipéptido; en el que el rendimiento del polipéptido con actividad FVIII aumenta en relación con una célula huésped cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16. En algunas realizaciones, el índice de adaptación de codones de la porción del nucleótido que codifica la porción N-terminal del polipéptido aumenta, disminuye o no cambia en relación con el índice de optimización por codones de SEQ ID NO: 16.

40 En otras realizaciones, la divulgación proporciona un método para mejorar el rendimiento de un polipéptido con actividad de FVIII, que comprende cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica el polipéptido, en el que el índice de adaptación de codones de una porción del nucleótido que codifica una porción N-terminal del polipéptido aumenta con respecto a la porción del nucleótido que codifica una porción C-terminal del polipéptido; en el que el rendimiento del polipéptido con actividad de FVIII aumenta en relación con una célula huésped cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende la SEQ ID NO: 16. En algunas realizaciones, el índice de adaptación de codones de la porción del nucleótido que codifica la porción C-terminal del polipéptido aumenta, disminuye o no cambia en relación con el índice de optimización de codones de SEQ ID NO: 16.

45 En otras realizaciones, un polipéptido codificado por la porción 5' del nucleótido, cuando se alinea correctamente, se corresponde con aproximadamente los aminoácidos 1-497 de SEQ ID NO: 17, aminoácidos 20-497 de SEQ ID NO: 17, o un fragmento del mismo. En otras realizaciones, un polipéptido codificado por la porción 3' del nucleótido, cuando se alinea correctamente, se corresponde con aproximadamente los aminoácidos 498-1458 de SEQ ID NO: 17, o un fragmento del mismo.

50 En algunas realizaciones, la expresión del polipéptido FVIII aumenta al menos aproximadamente 1,5 veces, al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 6 veces, al menos aproximadamente 7 veces, al menos aproximadamente 8 veces, al menos aproximadamente 9 veces, al menos aproximadamente 10 veces, al menos aproximadamente 11 veces, al menos aproximadamente 12 veces, al menos aproximadamente 13 veces, al menos aproximadamente 14 veces, al menos aproximadamente 15 veces, al menos aproximadamente 20 veces, al menos aproximadamente 25 veces, al menos aproximadamente 30 veces, al menos aproximadamente 35 veces, al menos aproximadamente 40 veces, al menos aproximadamente 50 veces, al menos aproximadamente 60 veces, al menos

aproximadamente 70 veces, al menos aproximadamente 80 veces, al menos aproximadamente 90 veces, al menos aproximadamente 100 veces, al menos aproximadamente 150 veces, o al menos aproximadamente 200 veces con respecto a una célula huésped cultivada en las mismas condiciones que comprende una secuencia de ácido nucleico de referencia que comprende SEQ ID NO: 16.

- 5 Hay una diversidad de métodos disponibles para producir de forma recombinante una proteína FVIII a partir de la molécula de ácido nucleico optimizada de la divulgación. Se puede producir un polinucleótido de la secuencia deseada mediante síntesis de ADN en fase sólida *de novo* o mediante mutagénesis por PCR de un polinucleótido preparado anteriormente. La mutagénesis mediada por oligonucleótidos es un método para preparar una sustitución, inserción, delección o alteración (p. ej., codón alterado) en una secuencia de nucleótidos. Por ejemplo, el ADN de partida se altera mediante la hibridación de un oligonucleótido que codifica la mutación deseada con un molde de ADN de cadena sencilla. Después de la hibridación, se utiliza una ADN polimerasa para sintetizar una segunda cadena complementaria completa del molde que incorpora el cebador oligonucleotídico. En una realización, la ingeniería genética, p. ej., mutagénesis por PCR basada en cebadores, es suficiente para incorporar una alteración, como se define en esta memoria, para producir un polinucleótido de la divulgación.
- 10
- 15 Para la producción de proteínas recombinantes, se inserta una secuencia de polinucleótidos optimizada de la divulgación que codifica la proteína FVIII en un vehículo de expresión apropiado, es decir, un vector que contiene los elementos necesarios para la transcripción y traducción de la secuencia codificante insertada o, en el caso de un vector viral de ARN, los elementos necesarios para la replicación y traducción.

- 20 La secuencia de polinucleótidos de la divulgación se inserta en el vector en el marco de lectura adecuado. A continuación, el vector de expresión se transfecta en una célula diana adecuada que expresará el polipéptido. Técnicas de transfección conocidas en la técnica incluyen, pero no se limitan a precipitación con fosfato de calcio (Wigler et al. 1978, Cell 14: 725) y electroporación (Neumann et al. 1982, EMBO, J. 1: 841). Se puede utilizar una diversidad de sistemas de vectores de expresión del huésped para expresar las proteínas FVIII descritas en esta memoria en células eucarióticas. En una realización, la célula eucariótica es una célula animal, incluyendo células de mamífero (p. ej., células HEK293, células PER.C6®, CHO, BHK, Cos, HeLa). Una secuencia de polinucleótidos de la divulgación también puede codificar una secuencia señal que permitirá que se secrete la proteína FVIII. Un experto en la técnica entenderá que mientras la proteína FVIII se traduce, la secuencia señal es escindida por la célula para formar la proteína madura. En la técnica se conocen diversas secuencias señal, p. ej., la secuencia señal del factor VII nativo, la secuencia señal del factor IX nativo y la secuencia señal de la cadena ligera de IgK de ratón. Alternativamente, en los casos en los que no se incluye una secuencia señal, la proteína FVIII puede recuperarse lisando las células.
- 25
- 30

- La proteína FVIII de la divulgación se puede sintetizar en un animal transgénico, tal como un roedor, una cabra, una oveja, un cerdo o una vaca. La expresión "animales transgénicos" se refiere a animales no humanos que han incorporado un gen extraño en su genoma. Debido a que este gen está presente en los tejidos de la línea germinal, se transmite de padres a hijos. Los genes exógenos se introducen en embriones unicelulares (Brinster et al. 1985, Proc. Natl. Acad. Sci. USA 82:4438). Se conocen en la técnica métodos para producir animales transgénicos que incluyen transgénicos que producen moléculas de inmunoglobulina (Wagner et al. 1981, Proc. Natl. Acad. Sci. USA 78: 6376; McKnight et al. 1983, Cell 34: 335; Brinster et al. 1983, Nature 306: 332; Ritchie et al. 1984, Nature 312: 517; Baldassarre et al. 2003, Theriogenology 59: 831; Robl et al. 2003, Theriogenology 59: 107; Malassagne et al. 2003, Xenotransplantation 10 (3): 267).
- 35

- Los vectores de expresión pueden codificar etiquetas que permitan una fácil purificación o identificación de la proteína producida de forma recombinante. Ejemplos incluyen, pero no se limitan a, el vector pUR278 (Ruther et al. 1983, EMBO J. 2: 1791) en el que la secuencia codificante de la proteína FVIII descrita en esta memoria se puede ligar en el vector en marco con la región codificante lac Z, de modo que se produce una proteína híbrida; vectores pGEX se pueden utilizar para expresar proteínas con una etiqueta de glutatión S-transferasa (GST). Estas proteínas son habitualmente solubles y se pueden purificar fácilmente de las células mediante adsorción en perlas de glutatión-agarosa seguido de elución en presencia de glutatión libre. Los vectores incluyen sitios de escisión (p. ej., PreCission Protease (Pharmacia, Peapack, NJ)) para facilitar la eliminación de la etiqueta después de la purificación.
- 40
- 45

- Para los fines de esta divulgación, se pueden emplear numerosos sistemas de vectores de expresión. Estos vectores de expresión son típicamente replicables en los organismos huéspedes como episomas o como parte integral del ADN cromosómico del huésped. Vectores de expresión pueden incluir secuencias de control de la expresión que incluyen, pero no se limitan a promotores (p. ej., promotores asociados de forma natural o heterólogos), potenciadores, secuencias señal, señales de corte y empalme, elementos potenciadores y secuencias de terminación de la transcripción. Preferiblemente, las secuencias de control de la expresión son sistemas de promotores eucarióticos en vectores capaces de transformar o transfectar células huésped eucarióticas. Vectores de expresión también pueden utilizar elementos de ADN que se derivan de virus animales tales como virus del papiloma bovino, virus del polio, adenovirus, virus vaccinia, baculovirus, retrovirus (RSV, MMTV o MOMLV), citomegalovirus (CMV) o virus SV40. Otros implican el uso de sistemas policistronicos con sitios de unión a ribosomas internos.
- 50
- 55

Comúnmente, los vectores de expresión contienen marcadores de selección (p. ej., resistencia a ampicilina, resistencia a higromicina, resistencia a tetraciclina o resistencia a neomicina) para permitir la detección de las células transformadas con



las secuencias de ADN deseadas (véase, *p. ej.*, Itakura et al., Patente de EE.UU. 4.704.362). Células que han integrado el ADN en sus cromosomas pueden seleccionarse introduciendo uno o más marcadores que permitan la selección de células huésped transfectadas. El marcador puede proporcionar prototrofia a un huésped auxotrófico, resistencia a biocidas (*p. ej.*, antibióticos) o resistencia a metales pesados tales como el cobre. El gen marcador seleccionable puede enlazarse directamente a las secuencias de ADN a expresar o introducirse en la misma célula mediante cotransformación.

Un ejemplo de un vector útil para expresar una secuencia de FVIII optimizada es NEOSPLA (Patente de EE.UU. Nº 6.159.730). Este vector contiene el promotor/potenciador del citomegalovirus, el promotor principal de la beta globina de ratón, el origen de replicación de SV40, la secuencia de poliadenilación de la hormona de crecimiento bovina, el exón 1 y el exón 2 de neomicina fosfotransferasa, el gen dihidrofolato reductasa y secuencia conductora. Se ha encontrado que este vector da como resultado un nivel muy alto de expresión de anticuerpos tras la incorporación de genes de la región variable y constante, transfección en células, seguida de selección en medio que contiene G418 y amplificación con metotrexato. Los sistemas vectoriales también se enseñan en las Pat. de EE.UU. Nº 5.736.137 y 5.658.570. Este sistema proporciona altos niveles de expresión, *p. ej.*, > 30 pg/célula/día. Otros sistemas de vectores ejemplares se describen, *p. ej.*, en la Patente de EE.UU. Nº 6.413.777.

En otras realizaciones, los polipéptidos de la divulgación de la presente divulgación se pueden expresar utilizando construcciones policistrónicas. En estos sistemas de expresión se pueden producir múltiples productos génicos de interés tales como múltiples polipéptidos de proteína de unión a multímeros a partir de una única construcción policistrónica. Estos sistemas utilizan ventajosamente un sitio de entrada al ribosoma interno (IRES) para proporcionar niveles relativamente altos de polipéptidos en células huésped eucarióticas. Secuencias de IRES compatibles se describen en la Pat. de EE.UU. Nº 6.193.980.

Más generalmente, una vez que se ha preparado el vector o la secuencia de ADN que codifica un polipéptido, el vector de expresión se puede introducir en una célula huésped apropiada. Es decir, las células huésped pueden transformarse. La introducción del plásmido en la célula huésped se puede lograr mediante diversas técnicas bien conocidas por los expertos en la técnica, como se comentó arriba. Las células transformadas se cultivan en condiciones apropiadas para la producción del polipéptido FVIII y se analizan para determinar la síntesis del polipéptido FVIII. Ejemplos de técnicas de ensayo incluyen ensayo inmunoabsorbente ligado a enzimas (ELISA), radioinmunoensayo (RIA) o análisis de clasificación de células activadas por fluorescencia (FACS), inmunohistoquímica y similares.

En descripciones de procesos para el aislamiento de polipéptidos de huéspedes recombinantes, el término "célula" y la expresión "cultivo celular" se utilizan indistintamente para designar la fuente del polipéptido, a menos que se especifique claramente lo contrario. En otras palabras, la recuperación del polipéptido de las "células" puede significar de células completas centrifugadas o del cultivo celular que contiene tanto el medio como las células suspendidas.

La línea de células huésped utilizada para la expresión de proteínas es preferiblemente de origen mamífero; lo más preferiblemente de origen humano o de ratón, ya que los ácidos nucleicos aislados de la divulgación han sido optimizados para la expresión en células humanas. Arriba se han descrito líneas de células huésped ejemplares. En una realización del método para producir un polipéptido con actividad de FVIII, la célula huésped es una célula HEK293. En otra realización del método para producir un polipéptido con actividad de FVIII, la célula huésped es una célula CHO.

Genes que codifican los polipéptidos de la divulgación también se pueden expresar en células no de mamífero tales como bacterias, levaduras o células vegetales. A este respecto, se apreciará que también pueden transformarse diversos microorganismos no mamíferos unicelulares tales como bacterias; *es decir*, aquellos capaces de crecer en cultivos o fermentación. Bacterias que son susceptibles de transformación incluyen miembros de las Enterobacteriaceae, tales como cepas de *Escherichia coli* o *Salmonella*; Bacillaceae, tales como *Bacillus subtilis*; *Pneumococcus*; *Streptococcus* y *Haemophilus influenzae*. Se apreciará además que, cuando se expresan en bacterias, los polipéptidos se convierten típicamente en parte de los cuerpos de inclusión. Los polipéptidos deben aislarse, purificarse y luego ensamblarse en moléculas funcionales.

Alternativamente, las secuencias de nucleótidos optimizadas de la divulgación pueden incorporarse en transgenes para la introducción en el genoma de un animal transgénico y la posterior expresión en la leche del animal transgénico (véase, *p. ej.*, Deboer et al., documento US 5.741.957, Rosen, documento US 5.304.489 y Meade et al., documento US 5.849.992). Transgenes adecuados incluyen secuencias codificantes para polipéptidos en enlace operable con un promotor y potenciador de un gen específico para la glándula mamaria, tal como caseína o beta lactoglobulina.

La producción *in vitro* permite aumentar la escala para dar grandes cantidades de los polipéptidos deseados. Técnicas para el cultivo de células de mamíferos en condiciones de cultivo de tejidos son conocidas en la técnica e incluyen cultivo en suspensión homogénea, *p. ej.*, en un reactor de aire o en un reactor de agitación continua, o cultivo de células inmovilizadas o atrapadas, *p. ej.*, en fibras huecas, microcápsulas, en microperlas de agarosa o cartuchos de material cerámico. Si es necesario y/o deseado, las soluciones de polipéptidos se pueden purificar mediante los métodos cromatográficos habituales, por ejemplo, filtración en gel, cromatografía de intercambio iónico, cromatografía sobre DEAE-celulosa o cromatografía de (inmuno)afinidad, *p. ej.*, después de la biosíntesis preferencial de un polipéptido sintético de la región bisagra o antes o después de la etapa de cromatografía HIC descrita en esta memoria. Una secuencia de etiqueta

de afinidad (p. *ej.*, una etiqueta His( 6)) puede fijarse o incluirse opcionalmente dentro de la secuencia polipeptídica para facilitar la purificación aguas abajo.

- 5 Una vez expresada, la proteína FVIII se puede purificar de acuerdo con procedimientos estándares de la técnica, que incluyen la precipitación con sulfato de amonio, la cromatografía en columna de afinidad, la purificación por HPLC, la electroforesis en gel y similares (véase, *en general*, Scopes, Protein Purification (Springer-Verlag, NY, (1982)). Se prefieren proteínas sustancialmente puras de al menos aproximadamente 90 a 95 % de homogeneidad, y lo más preferido de 98 a 99 % o más de homogeneidad, para usos farmacéuticos.

### Composición Farmacéutica

- 10 Composiciones que contienen una molécula de ácido nucleico aislada, un polipéptido que tiene actividad FVIII codificada por la molécula de ácido nucleico, un vector o una célula huésped de la presente divulgación pueden contener un soporte farmacéuticamente aceptable adecuado. Por ejemplo, pueden contener excipientes y/o sustancias auxiliares que faciliten el procesamiento de los compuestos activos en preparaciones diseñadas para el suministro en el sitio de acción.

- 15 La composición farmacéutica se puede formular para administración parenteral (*es decir*, intravenosa, subcutánea o intramuscular) mediante inyección en bolo. Formulaciones para inyección se pueden presentar en forma de dosificación unitaria, p. *ej.*, en ampollas o en envases multidosis con un conservante añadido. Las composiciones pueden adoptar formas tales como suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos, y pueden contener agentes de formulación tales como agentes de suspensión, estabilización y/o dispersión. Alternativamente, el ingrediente activo puede estar en forma de polvo para reconstitución con un vehículo adecuado, p. *ej.*, agua apirógena.

- 20 Formulaciones adecuadas para administración parenteral también incluyen soluciones acuosas de los compuestos activos en forma hidrosoluble, por ejemplo, sales hidrosolubles. Además, se pueden administrar suspensiones de los compuestos activos como suspensiones de inyección oleosas apropiadas. Disolventes o vehículos lipófilos adecuados incluyen aceites grasos, por ejemplo, aceite de sésamo o ésteres de ácidos grasos sintéticos, por ejemplo, oleato de etilo o triglicéridos. Las suspensiones acuosas para inyección pueden contener sustancias que aumentan la viscosidad de la suspensión, incluyendo, por ejemplo, carboximetilcelulosa sódica, sorbitol y dextrano. Opcionalmente, la suspensión también puede  
25 contener estabilizadores. Los liposomas también se pueden utilizar para encapsular las moléculas de la divulgación para su suministro en células o espacios intersticiales. Soportes farmacéuticamente aceptables ejemplares son disolventes fisiológicamente compatibles, medios de dispersión, revestimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardadores de la absorción, agua, solución salina, solución salina tamponada con fosfato, dextrosa, glicerol, etanol y similares. En algunas realizaciones, la composición comprende agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol o cloruro sódico. En otras realizaciones, las composiciones comprenden  
30 sustancias farmacéuticamente aceptables tales como agentes humectantes o cantidades menores de sustancias auxiliares tales como agentes humectantes o emulsionantes, conservantes o tampones, que potencian la vida útil o la eficacia de los ingredientes activos.

- 35 Composiciones de la divulgación pueden estar en una diversidad de formas, incluyendo, por ejemplo, líquidas (p. *ej.*, soluciones inyectables e infusibles), dispersiones, suspensiones, formas de dosificación semisólidas y sólidas. La forma preferida depende del modo de administración y de la aplicación terapéutica.

- La composición se puede formular como una solución, microemulsión, dispersión, liposoma u otra estructura ordenada adecuada para una alta concentración de fármaco. Soluciones inyectables estériles se pueden preparar incorporando el  
40 ingrediente activo en la cantidad requerida en un disolvente apropiado con uno o una combinación de ingredientes arriba enumerados, según se requiera, seguido de esterilización por filtración. Generalmente, las dispersiones se preparan incorporando el ingrediente activo en un vehículo estéril que contiene un medio de dispersión básico y los demás ingredientes requeridos entre los arriba enumerados. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos de preparación preferidos son el secado al vacío y la liofilización que produce un polvo del ingrediente activo más cualquier ingrediente adicional deseado de una solución filtrada previamente estéril. La fluidez  
45 adecuada de una solución se puede mantener, por ejemplo, mediante el uso de un recubrimiento tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de la dispersión y mediante el uso de tensioactivos. La absorción prolongada de composiciones inyectables puede lograrse incluyendo en la composición un agente que retrase la absorción, por ejemplo, sales de monoestearato y gelatina.

- 50 El ingrediente activo se puede formular con una formulación o dispositivo de liberación controlada. Ejemplos de formulaciones y dispositivos de este tipo incluyen implantes, parches transdérmicos y sistemas de suministro microencapsulados. Se pueden utilizar polímeros biodegradables y biocompatibles, por ejemplo, etileno y acetato de vinilo, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres y ácido poliláctico. Métodos para la preparación de formulaciones y dispositivos de este tipo son conocidos en la técnica. Véase, p. *ej.*, Sustained and Controlled Release Drug Delivery Systems, J. R. Robinson, ed., Marcel Dekker, Inc., Nueva York, 1978.

- Formulaciones de depósito inyectables se pueden preparar formando matrices microencapsuladas del fármaco en polímeros biodegradables tales como polilactida-poliglicolida. Dependiendo de la relación de fármaco a polímero y de la naturaleza del polímero empleado, se puede controlar la velocidad de liberación del fármaco. Otros polímeros biodegradables ejemplares son poliortoésteres y polianhídridos. Formulaciones inyectables de depósito también se pueden preparar atrapando el fármaco en liposomas o microemulsiones.
- Compuestos activos complementarios se pueden incorporar en las composiciones. En una realización, la proteína química de la divulgación se formula con otro factor de coagulación, o una variante, fragmento, análogo o derivado del mismo. Por ejemplo, el factor de coagulación incluye, pero no se limita a factor V, factor VII, factor VIII, factor IX, factor X, factor XI, factor XII, factor XIII, protrombina, fibrinógeno, factor de von Willebrand o factor tisular soluble recombinante. (rsTF, por sus siglas en inglés) o formas activadas de cualquiera de los anteriores. El factor de coagulación del agente hemostático también puede incluir fármacos anti-fibrinolíticos, *p. ej.*, ácido épsilon-amino-caproico, ácido tranexámico.
- Regímenes de dosificación se pueden ajustar para proporcionar la respuesta óptima deseada. Por ejemplo, se puede administrar un solo bolo, se pueden administrar varias dosis divididas a lo largo del tiempo, o la dosis se puede reducir o aumentar proporcionalmente según lo indiquen las exigencias de la situación terapéutica. Es ventajoso formular composiciones parenterales en forma de unidades de dosificación para facilitar la administración y uniformidad de la dosificación. Véase, *p. ej.*, Remington's Pharmaceutical Sciences (Mack Pub. Co., Easton, Pensilvania 1980).
- Además del compuesto activo, la forma de dosificación líquida puede contener ingredientes inertes tales como agua, alcohol etílico, carbonato de etilo, acetato de etilo, alcohol bencílico, benzoato de bencilo, propilenglicol, 1,3-butilenglicol, dimetilformamida, aceites, glicerol, alcohol tetrahidrofurfurílico, polietilenglicoles y ésteres de ácidos grasos de sorbitán.
- Ejemplos no limitativos de soportes farmacéuticos adecuados también se describen en Remington's Pharmaceutical Sciences de E.W. Martin. Algunos ejemplos de excipientes incluyen almidón, glucosa, lactosa, sacarosa, gelatina, malta, arroz, harina, greda, gel de sílice, estearato de sodio, monoestearato de glicerol, talco, cloruro sódico, leche descremada en polvo, glicerol, propilenglicol, agua, etanol, y similares. La composición también puede contener reactivos tamponadores del pH y agentes humectantes o emulsionantes.
- Para la administración oral, la composición farmacéutica puede adoptar la forma de tabletas o cápsulas preparadas por medios convencionales. La composición también se puede preparar como un líquido, por ejemplo, un jarabe o una suspensión. El líquido puede incluir agentes de suspensión (*p. ej.*, jarabe de sorbitol, derivados de celulosa o grasas comestibles hidrogenadas), agentes emulsionantes (lecitina o acacia), vehículos no acuosos (*p. ej.*, aceite de almendras, ésteres oleosos, alcohol etílico o aceites vegetales fraccionados) y conservantes (*p. ej.*, p-hidroxibenzoatos de metilo o propilo o ácido sórbico). Las preparaciones también pueden incluir agentes aromatizantes, colorantes y edulcorantes. Alternativamente, la composición se puede presentar como un producto seco para reconstitución con agua u otro vehículo adecuado.
- Para administración bucal, la composición puede adoptar la forma de tabletas o pastillas de acuerdo con los protocolos convencionales.
- Para la administración por inhalación, los compuestos para uso de acuerdo con la presente divulgación se suministran convenientemente en forma de aerosol nebulizado con o sin excipientes o en forma de un spray de aerosol de un envase presurizado o nebulizador, con opcionalmente un propulsor, *p. ej.*, diclorodifluorometano, triclorofluorometano, diclorotetrafluorometano, dióxido de carbono u otro gas adecuado. En el caso de un aerosol presurizado, la unidad de dosificación se puede determinar proporcionando una válvula para suministrar una cantidad medida. Se pueden formular cápsulas y cartuchos de, *p. ej.*, gelatina para uso en un inhalador o insuflador que contenga una mezcla de polvo del compuesto y una base de polvo adecuada tal como lactosa o almidón.
- La composición farmacéutica también se puede formular para administración rectal como supositorio o enema de retención, *p. ej.*, que contiene bases de supositorios convencionales tales como manteca de cacao u otros gliceridos.
- En una realización, una composición farmacéutica comprende un polipéptido que tiene actividad de Factor VIII, una molécula de ácido nucleico optimizada que codifica el polipéptido que tiene actividad de Factor VIII, el vector que comprende la molécula de ácido nucleico o la célula huésped que comprende el vector, y un soporte farmacéuticamente aceptable. En algunas realizaciones, la composición se administra por una vía seleccionada del grupo que consiste en administración tópica, administración intraocular, administración parenteral, administración intratecal, administración subdural y administración oral. La administración parenteral puede ser administración intravenosa o subcutánea.
- En otras realizaciones, la composición se utiliza para tratar una enfermedad o afección de sangrado en un sujeto que lo necesite. La enfermedad o afección de sangrado se selecciona del grupo que consiste en un trastorno de sangrado de la coagulación, hemartrosis, sangrado muscular, sangrado oral, hemorragia, hemorragia en los músculos, hemorragia oral, traumatismo, trauma capitis, sangrado gastrointestinal, hemorragia intracraneal, hemorragia intra-abdominal, hemorragia intratorácica, fractura ósea, sangrado del sistema nervioso central, sangrado en el espacio retrofaríngeo, sangrado en el

espacio retroperitoneal, sangrado en la vaina del iliopsoas y cualesquiera combinaciones de los mismos. En aún otras realizaciones, el sujeto está programado para someterse a una cirugía. En aún otras realizaciones, el tratamiento es profiláctico o bajo demanda.

### Métodos de Tratamiento

5 La presente invención proporciona una molécula de ácido nucleico o un vector como se define en las reivindicaciones adjuntas para uso en un método para tratar un trastorno de sangrado. En algunas realizaciones, el trastorno de sangrado se caracteriza por una deficiencia de FVIII. En algunas realizaciones, el trastorno de sangrado es hemofilia. En algunas realizaciones, el trastorno de sangrado es la hemofilia A. En algunas realizaciones, la actividad del FVIII en plasma a las 24 horas post-administración aumenta en relación con un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia, o un polipéptido codificado por la molécula de ácido nucleico de referencia.

10 En algunas realizaciones, la actividad de FVIII en plasma aumenta aproximadamente a las 6 horas, aproximadamente a las 12 horas, aproximadamente a las 18 horas, aproximadamente a las 24 horas, aproximadamente a las 36 horas, aproximadamente a las 48 horas, aproximadamente a los 3 días, aproximadamente a los 4 días, aproximadamente a los 5 días, aproximadamente a los 6 días, aproximadamente a los 7 días, aproximadamente a los 8 días, aproximadamente a los 9 días, aproximadamente a los 10 días, aproximadamente a los 11 días, aproximadamente a los 12 días, aproximadamente a los 13 días, aproximadamente a los 14 días, aproximadamente a los 15 días, aproximadamente a los 16 días, aproximadamente a los 17 días, aproximadamente a los 18 días, aproximadamente a los 19 días, aproximadamente a los 20 días, aproximadamente a los 21 días, aproximadamente a los 22 días, aproximadamente a los 23 días, aproximadamente a los 24 días, aproximadamente a los 25 días, aproximadamente a los 26 días, aproximadamente a los 27 días o aproximadamente a los 28 días post-administración en relación con un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector viral que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En determinadas realizaciones, la actividad del FVIII en plasma aumenta aproximadamente 24 horas post-administración en relación con un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector viral que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En otra realización, la actividad del FVIII en plasma aumenta aproximadamente 21 días post-administración en relación con un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector viral que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia.

30 En algunas realizaciones, la actividad post-administración de FVIII en plasma aumenta en al menos aproximadamente 1,5 veces, al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 6 veces, al menos aproximadamente 7 veces, al menos aproximadamente 8 veces, al menos aproximadamente 9 veces, al menos aproximadamente 10 veces, al menos aproximadamente 11 veces, al menos aproximadamente 12 veces, al menos aproximadamente 13 veces, al menos aproximadamente 14 veces, al menos aproximadamente 15 veces, al menos aproximadamente 20 veces, al menos aproximadamente 25 veces, al menos aproximadamente 30 veces, al menos aproximadamente 35 veces, al menos aproximadamente 40 veces, al menos aproximadamente 50 veces, al menos aproximadamente 60 veces, al menos aproximadamente 70 veces, al menos aproximadamente 80 veces, al menos aproximadamente 90 veces, al menos aproximadamente 100 veces, al menos aproximadamente 150 veces, al menos aproximadamente 200 veces, al menos aproximadamente 250 veces, al menos aproximadamente 300 veces, al menos aproximadamente 350 veces, al menos aproximadamente 400 veces, al menos aproximadamente 450 veces o al menos aproximadamente 500 veces en relación con un sujeto al que se le administró una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector viral que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas realizaciones, la actividad del FVIII en plasma post-administración aumenta en al menos aproximadamente un 150 %, al menos aproximadamente un 200 %, al menos aproximadamente un 250 %, al menos aproximadamente un 300 %, al menos aproximadamente un 350 %, al menos aproximadamente un 400 %, al menos aproximadamente un 450 %, al menos aproximadamente un 500 %, al menos aproximadamente un 550 %, al menos aproximadamente un 600 %, al menos aproximadamente un 650 %, al menos aproximadamente un 700 %, al menos aproximadamente un 750 %, al menos aproximadamente un 800 %, al menos aproximadamente un 850 %, al menos aproximadamente un 900 %, al menos aproximadamente un 950 %, al menos aproximadamente un 1000 %, al menos aproximadamente un 1500 %, al menos aproximadamente un 2000 %, al menos aproximadamente un 2500 %, al menos aproximadamente un 3000 %, al menos aproximadamente un 3500 %, al menos aproximadamente un 4000 %, al menos aproximadamente un 4500 %, al menos aproximadamente un 5000 %, al menos aproximadamente un 5500 %, al menos aproximadamente un 6000 %, al menos aproximadamente un 7000 %, al menos aproximadamente un 8000 %, al menos aproximadamente un 9000 %, al menos aproximadamente un 10.000 % en relación con los niveles de FVIII circulantes fisiológicamente normales. En una realización, la actividad de FVIII en plasma después de la administración aumenta en al menos aproximadamente 3000 a aproximadamente 5000 % en relación con los niveles de FVIII circulantes fisiológicamente normales. En algunas realizaciones, a las 24 horas post-administración de un plásmido que comprende un gen optimizado por codones que codifica polipéptidos con actividad de factor VIII (FVIII) descrito en esta memoria, o a los 21 días post-administración de un vector lentiviral o AAV que comprende un gen optimizado por codones que codifica

polipéptidos con la actividad de Factor VIII (FVIII) descrita en esta memoria, la actividad de FVIII en plasma aumenta al menos aproximadamente 6 veces en relación con un sujeto al que se le administra una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas realizaciones, a las 24 horas post-administración de un plásmido que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII) descritos en esta memoria, o a los 21 días post-administración de un vector lentiviral o AAV que comprende un gen optimizado por codones que codifica polipéptidos con la actividad de Factor VIII (FVIII) descritos en esta memoria, la actividad de FVIII en plasma aumenta al menos aproximadamente 10 veces en relación con un sujeto al que se le administra una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia, o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas realizaciones, a las 24 horas post-administración de un plásmido que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII) descritos en esta memoria, o a los 21 días post-administración de un vector lentiviral o AAV que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII) descritos en esta memoria, la actividad de FVIII en plasma aumenta al menos aproximadamente 18 veces en relación con un sujeto al que se le administra una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia, o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas realizaciones, a las 24 horas post-administración de un plásmido que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII) descritos en esta memoria, o a los 21 días post-administración de un vector lentiviral o AAV que comprende un gen optimizado por codones que codifica polipéptidos con la actividad de Factor VIII (FVIII) descritos en esta memoria, la actividad de FVIII en plasma aumenta al menos aproximadamente 30 veces en relación con un sujeto al que se le administra una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia, o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas realizaciones, a las 24 horas post-administración de un plásmido que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII) descritos en esta memoria, o a los 21 días post-administración de un vector lentiviral o AAV que comprende un gen optimizado por codones que codifica polipéptidos con la actividad de Factor VIII (FVIII) descritos en esta memoria, la actividad de FVIII en plasma aumenta al menos aproximadamente 50 veces en relación con un sujeto al que se le administra una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia, o un polipéptido codificado por la molécula de ácido nucleico de referencia. En algunas realizaciones, a las 24 horas post-administración de un plásmido que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII) descritos en esta memoria, o a los 21 días post-administración de un vector lentiviral o AAV que comprende un gen optimizado por codones que codifica polipéptidos con la actividad de Factor VIII (FVIII) descritos en esta memoria, la actividad de FVIII en plasma aumenta al menos aproximadamente 100 veces en relación con un sujeto al que se le administra una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16, un vector que comprende la molécula de ácido nucleico de referencia o un polipéptido codificado por la molécula de ácido nucleico de referencia.

La divulgación también se refiere a una molécula de ácido nucleico aislada descrita en esta memoria o un polipéptido que tiene actividad de FVIII codificada por el ácido nucleico para uso en un método para tratar, mejorar o prevenir un trastorno hemostático en un sujeto, comprendiendo el método administrar un cantidad terapéuticamente eficaz de la molécula de ácido nucleico aislada o del polipéptido. El tratamiento, la mejora y la prevención mediante la molécula de ácido nucleico aislada o el polipéptido codificado pueden ser una terapia de derivación. El sujeto que recibe terapia de derivación ya puede haber desarrollado un inhibidor de un factor de coagulación, *p. ej.*, FVIII, o está sujeto a desarrollar un inhibidor de factor de coagulación.

Las moléculas de ácido nucleico, los vectores o los polipéptidos de la divulgación tratan o previenen un trastorno hemostático al fomentar la formación de un coágulo de fibrina. El polipéptido que tiene actividad de FVIII codificado por la molécula de ácido nucleico de la divulgación puede activar un miembro de una cascada de coagulación. El factor de coagulación puede participar en la vía extrínseca, la vía intrínseca o ambas.

Las moléculas de ácido nucleico, los vectores o los polipéptidos de la divulgación pueden utilizarse para tratar trastornos hemostáticos que se sabe que son tratables con FVIII. Los trastornos hemostáticos que se pueden tratar utilizando los métodos de la divulgación incluyen, pero no se limitan a hemofilia A, hemofilia B, enfermedad de von Willebrand, deficiencia de Factor XI (deficiencia de PTA), deficiencia de Factor XII, así como deficiencias o anomalías estructurales en fibrinógeno, protrombina, Factor V, Factor VII, Factor X o Factor XIII, hemartrosis, sangrado muscular, sangrado oral, hemorragia, hemorragia en los músculos, hemorragia oral, trauma, trauma capitis, sangrado gastrointestinal, hemorragia intracraneal, hemorragia intraabdominal, hemorragia intratorácica, fractura ósea, sangrado del sistema nervioso central, sangrado en el espacio retrofaringeo, sangrado en el espacio retroperitoneal y sangrado en la vaina del iliopsoas. Composiciones para

la administración a un sujeto incluyen moléculas de ácido nucleico que comprenden una secuencia de nucleótidos optimizada de la divulgación que codifica un factor de coagulación FVIII (para aplicaciones de terapia génica) así como moléculas polipeptídicas de FVIII.

5 En algunas realizaciones, el trastorno hemostático es un trastorno hereditario. En una realización, el sujeto tiene hemofilia A. En otras realizaciones, el trastorno hemostático es el resultado de una deficiencia de FVIII. En otras realizaciones, el trastorno hemostático puede ser el resultado de un factor de coagulación FVIII defectuoso.

10 En otra realización, el trastorno hemostático puede ser un trastorno adquirido. El trastorno adquirido puede resultar de una enfermedad o afección secundaria subyacente. La afección no relacionada puede ser, como un ejemplo, pero no como una limitación, cáncer, una enfermedad autoinmune o embarazo. El trastorno adquirido puede resultar de la vejez o de la medicación para tratar un trastorno secundario subyacente (*p. ej.*, quimioterapia contra el cáncer).

15 La divulgación también se refiere al tratamiento de un sujeto que no tiene un trastorno hemostático o una enfermedad o afección secundaria que resulte en la adquisición de un trastorno hemostático. Por lo tanto, la divulgación se refiere a la molécula de ácido nucleico aislada, vector o polipéptido de FVIII descritos en esta memoria para uso en un método para tratar a un sujeto que necesita un agente hemostático general, que comprende administrar una cantidad terapéuticamente eficaz de la molécula de ácido nucleico aislada, vector o polipéptido FVIII. Por ejemplo, el sujeto que necesita un agente hemostático general se somete o está a punto de someterse a cirugía. La molécula de ácido nucleico aislada, el vector o el polipéptido de FVIII de la divulgación se pueden administrar antes o después de la cirugía como profiláctico. La molécula de ácido nucleico aislada, el vector o el polipéptido de FVIII de la divulgación se pueden administrar antes o después de la cirugía para controlar un episodio de sangrado agudo. La cirugía puede incluir, pero no se limita a trasplante de hígado, resección de hígado o trasplante de células madre.

En otra realización, la molécula de ácido nucleico aislada, el vector o el polipéptido de FVIII de la divulgación se pueden utilizar para tratar a un sujeto que tiene un episodio de sangrado agudo que no tiene un trastorno hemostático. El episodio de sangrado agudo puede resultar de un trauma severo, *p. ej.*, cirugía, accidente de automóvil, herida, laceración por arma de fuego o cualquier otro evento traumático que resulte en un sangrado incontrolable.

25 La molécula de ácido nucleico aislada, el vector o la proteína FVIII se pueden utilizar para tratar profilácticamente a un sujeto con un trastorno hemostático. La molécula de ácido nucleico aislada, el vector o la proteína FVIII se pueden utilizar para tratar un episodio de sangrado agudo en un sujeto con un trastorno hemostático.

30 En otra realización, la expresión de la proteína FVIII mediante la administración de la molécula de ácido nucleico aislada o el vector de la divulgación no induce una respuesta inmunitaria en un sujeto. En algunas realizaciones, la respuesta inmune comprende el desarrollo de anticuerpos contra FVIII. En algunas realizaciones, la respuesta inmunitaria comprende la secreción de citoquinas. En algunas realizaciones, la respuesta inmunitaria comprende la activación de células B, células T o tanto células B como células T. En algunas realizaciones, la respuesta inmunitaria es una respuesta inmunitaria inhibidora, en donde la respuesta inmunitaria en el sujeto reduce la actividad de la proteína FVIII en relación con la actividad del FVIII en un sujeto que no ha desarrollado una respuesta inmunitaria. En determinadas realizaciones, la expresión de la proteína FVIII mediante la administración de la molécula de ácido nucleico aislada o el vector de la divulgación previene una respuesta inmunitaria inhibidora contra la proteína FVIII o la proteína FVIII expresada a partir de la molécula de ácido nucleico aislada o el vector.

40 En algunas realizaciones, una composición de molécula de ácido nucleico aislada, vector o proteína FVIII de la divulgación se administra en combinación con al menos otro agente que fomenta la hemostasia. Dicho otro agente que fomenta la hemostasia es un agente terapéutico con actividad coagulante demostrada. Como ejemplo, pero no como limitación, el agente hemostático puede incluir Factor V, Factor VII, Factor IX, Factor X, Factor XI, Factor XII, Factor XIII, protrombina o fibrinógeno o formas activadas de cualquiera de los anteriores. El factor de coagulación o agente hemostático también puede incluir fármacos antifibrinolíticos, *p. ej.*, ácido épsilon-aminocaproico, ácido tranexámico.

45 En una realización de la divulgación, la composición (*p. ej.*, la molécula de ácido nucleico aislada, vector o polipéptido de FVIII) es una en la que el FVIII está presente en forma activable cuando se administra a un sujeto. Una molécula activable de este tipo puede activarse *in vivo* en el sitio de coagulación después de la administración a un sujeto.

50 La molécula de ácido nucleico aislada, vector o polipéptido de FVIII puede administrarse por vía intravenosa, subcutánea, intramuscular o a través de cualquier superficie mucosa, *p. ej.*, por vía oral, sublingual, bucal, sublingual, nasal, rectal, vaginal o pulmonar. La proteína FVIII se puede implantar o enlazar a un soporte sólido de biopolímero que permite la liberación lenta de la proteína quimérica en el sitio deseado.

Para la administración oral, la composición farmacéutica puede adoptar la forma de tabletas o cápsulas preparadas por medios convencionales. La composición también se puede preparar como un líquido, por ejemplo, un jarabe o una suspensión. El líquido puede incluir agentes de suspensión (*p. ej.*, jarabe de sorbitol, derivados de celulosa o grasas comestibles hidrogenadas), agentes emulsionantes (lecitina o acacia), vehículos no acuosos (*p. ej.*, aceite de almendras,

ésteres oleosos, alcohol etílico o aceites vegetales fraccionados) y conservantes (p. ej., p- hidroxibenzoatos de metilo o propilo o ácido sórbico). Las preparaciones también pueden incluir agentes aromatizantes, colorantes y edulcorantes. Alternativamente, la composición se puede presentar como un producto seco para reconstitución con agua u otro vehículo adecuado.

- 5 Para la administración bucal y sublingual, la composición puede adoptar la forma de tabletas, pastillas o películas de disolución rápida de acuerdo con los protocolos convencionales.

10 Para la administración por inhalación, el polipéptido que tiene actividad de FVIII para uso de acuerdo con la presente divulgación se suministra convenientemente en forma de espray de aerosol desde un envase presurizado o nebulizador (p. ej., en PBS), con un propulsor adecuado, p. ej., diclorodifluorometano, triclorofluorometano, diclorotetrafluorometano, dióxido de carbono u otro gas adecuado. En el caso de un aerosol presurizado, la unidad de dosificación se puede determinar proporcionando una válvula para suministrar una cantidad medida. Se pueden formular cápsulas y cartuchos de, p. ej., gelatina para uso en un inhalador o insuflador que contengan una mezcla de polvos del compuesto y una base de polvo adecuada tal como lactosa o almidón.

15 En una realización, la vía de administración de la molécula de ácido nucleico aislada, vector o polipéptido de FVIII es parenteral. El término parenteral, tal como se utiliza en esta memoria, incluye administración intravenosa, intraarterial, intraperitoneal, intramuscular, subcutánea, rectal o vaginal. Se prefiere la forma intravenosa de administración parenteral. Si bien todas estas formas de administración están claramente contempladas dentro del alcance de la divulgación, una forma de administración sería una solución para inyección, en particular para inyección o goteo intravenoso o intraarterial. Habitualmente, una composición farmacéutica adecuada para inyección puede comprender un tampón (p. ej., tampón acetato, fosfato o citrato), un tensioactivo (p. ej., polisorbato), opcionalmente un agente estabilizador (p. ej., albúmina humana), etc. Sin embargo, en otros métodos compatibles con las enseñanzas en esta memoria, la molécula de ácido nucleico aislada, el vector o el polipéptido de FVIII pueden suministrarse directamente en el sitio de la población celular adversa, aumentando con ello la exposición del tejido enfermo al agente terapéutico.

25 Preparaciones para administración parenteral incluyen soluciones, suspensiones y emulsiones acuosas o no acuosas estériles. Ejemplos de disolventes no acuosos son propilenglicol, polietilenglicol, aceites vegetales tales como el aceite de oliva y ésteres orgánicos inyectables tal como el oleato de etilo. Soportes acuosos incluyen agua, soluciones, emulsiones o suspensiones alcohólicas/acuosas, incluyendo solución salina y medios tamponados. En la presente divulgación, soportes farmacéuticamente aceptables incluyen, pero no se limitan a tampón fosfato 0,01-0,1 M y preferiblemente 0,05 M o solución salina al 0,8 %. Otros vehículos parenterales comunes incluyen soluciones de fosfato de sodio, dextrosa de Ringer, dextrosa y cloruro sódico, Ringer lactatado o aceites fijos. Vehículos intravenosos incluyen reponedores de fluidos y nutrientes, reponedores de electrolitos, tales como los basados en dextrosa de Ringer, y similares. También pueden estar presentes conservantes y otros aditivos tales como, por ejemplo, agentes antimicrobianos, antioxidantes, agentes quelantes y gases inertes y similares.

35 Más particularmente, composiciones farmacéuticas adecuadas para uso inyectable incluyen soluciones acuosas estériles (cuando son solubles en agua) o dispersiones y polvos estériles para la preparación extemporánea de soluciones o dispersiones inyectables estériles. En tales casos, la composición debe ser estéril y debe ser fluida en la medida en que exista una fácil jeringabilidad. Debe ser estable bajo las condiciones de fabricación y almacenamiento y preferentemente se preservará contra la acción contaminante de microorganismos, tales como bacterias y hongos. El soporte puede ser un disolvente o medio de dispersión que contenga, por ejemplo, agua, etanol, poliol (p. ej., glicerol, propilenglicol y polietilenglicol líquido, y similares), y mezclas adecuadas de los mismos. La fluidez adecuada se puede mantener, por ejemplo, mediante el uso de un recubrimiento tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de la dispersión y mediante el uso de tensioactivos.

45 La prevención de la acción de microorganismos puede lograrse mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal y similares. En muchos casos, será preferible incluir en la composición agentes isotónicos, por ejemplo, azúcares, polialcoholes, tales como manitol, sorbitol o cloruro sódico. La absorción prolongada de las composiciones inyectables se puede conseguir incluyendo en la composición un agente que retrase la absorción, por ejemplo, monoestearato de aluminio y gelatina.

50 En cualquier caso, soluciones inyectables estériles se pueden preparar incorporando un compuesto activo (p. ej., un polipéptido solo o en combinación con otros agentes activos) en la cantidad requerida en un disolvente apropiado con uno o una combinación de los ingredientes enumerados en esta memoria, según se requiera, seguido de esterilización filtrada. Generalmente, las dispersiones se preparan incorporando el compuesto activo en un vehículo estéril, que contiene un medio de dispersión básico y los otros ingredientes requeridos entre los enumerados arriba. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos preferidos de preparación son secado en vacío y liofilización, lo que produce un polvo de un ingrediente activo más cualquier ingrediente adicional deseado a partir de una solución filtrada previamente en condiciones estériles del mismo. Las preparaciones para inyecciones se procesan, se envasan en recipientes tales como ampollas, bolsas, frascos, jeringas o viales y se sellan en condiciones asépticas de acuerdo con métodos conocidos en la técnica. Además, las preparaciones se pueden envasar y vender en forma de un kit.

Artículos de fabricación de este tipo tendrán preferiblemente etiquetas o prospectos que indiquen que las composiciones asociadas son útiles para tratar a un sujeto que padece o está predispuesto a trastornos de la coagulación.

La composición farmacéutica también se puede formular para administración rectal en forma de un supositorio o enema de retención, *p. ej.*, que contienen bases de supositorios convencionales tales como manteca de cacao u otros gliceridos.

5 Dosis efectivas de las composiciones de la presente divulgación para el tratamiento de afecciones varían dependiendo de muchos factores diferentes, incluyendo medios de administración, el sitio diana, el estado fisiológico del paciente, si el paciente es un ser humano o un animal, otros medicamentos administrado, y si el tratamiento es profiláctico o terapéutico. Habitualmente, el paciente es un ser humano, pero también pueden tratarse mamíferos no humanos, incluyendo los mamíferos transgénicos. Las dosificaciones de tratamiento pueden titularse utilizando métodos de rutina conocidos por los expertos en la técnica para optimizar la seguridad y la eficacia.

10 Dosificaciones para administrar un plásmido que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII), tal como se describe en esta memoria, o un polipéptido codificado por el gen optimizado por codones puede oscilar entre 1000 ug/kg y 0,1 ng/kg de peso corporal. En una realización, el intervalo de dosificación es de 1 ug/kg a 100 ug/kg. Dosificaciones para administrar un vector lentiviral que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII), tal como se describe en esta memoria, pueden oscilar entre  $10^3$  y  $10^{15}$  UT/kg. Dosificaciones para administrar un vector AAV que comprende un gen optimizado por codones que codifica polipéptidos con actividad de Factor VIII (FVIII), tal como se describe en esta memoria, pueden oscilar entre  $10^5$  y  $10^{18}$  VG/kg.

15 La molécula de ácido nucleico aislada, plásmido, vector o polipéptido de FVIII aislado se puede administrar como una dosis única o como dosis múltiples, en donde las dosis múltiples se pueden administrar de forma continua o en intervalos de tiempo específicos. Pueden emplearse ensayos *in vitro* para determinar intervalos de dosis óptimos y/o los programas de administración. Ensayos *in vitro* que miden la actividad del factor de coagulación son conocidos en la técnica. Adicionalmente, dosis efectivas pueden extrapolarse a partir de curvas dosis-respuesta obtenidas de modelos animales, *p. ej.*, un perro hemofílico (Mount et al. 2002, Blood 99 (8): 2670).

20 También se pretende que dosis intermedias en los intervalos anteriores estén dentro del alcance de la divulgación. A los sujetos se les pueden administrar dosis de este tipo diariamente, en días alternos, semanalmente o de acuerdo con cualquier otro programa determinado por análisis empírico. Un tratamiento ejemplar implica la administración en dosificaciones múltiples a lo largo de un período prolongado, por ejemplo, de al menos seis meses. En algunos métodos, se pueden administrar dos o más polipéptidos simultáneamente, en cuyo caso la dosificación de cada uno de los polipéptidos administrado se encuentra dentro de los intervalos indicados.

25 La molécula de ácido nucleico aislada, el vector o los polipéptidos de FVIII de la divulgación se pueden administrar en múltiples ocasiones. Los intervalos entre dosificaciones únicas pueden ser diarios, semanales, mensuales o anuales. Los intervalos también pueden ser irregulares como se indica midiendo los niveles en sangre del polipéptido o antígeno modificado en el paciente. Alternativamente, los polipéptidos se pueden administrar como una formulación de liberación sostenida, en cuyo caso se requiere una administración menos frecuente. La dosificación y la frecuencia varían dependiendo de la semivida del polipéptido o polinucleótido en el paciente.

30 La dosificación y la frecuencia de administración pueden variar dependiendo de si el tratamiento es profiláctico o terapéutico. En aplicaciones profilácticas, las composiciones que contienen la molécula de ácido nucleico aislada, el vector o el polipéptido FVIII o un cóctel de los mismos se administran a un paciente que aún no está en el estado de enfermedad para potenciar la resistencia del paciente o minimizar los efectos de la enfermedad. Una cantidad de este tipo se define como una "dosis efectiva profiláctica". Se administra una dosificación relativamente baja a intervalos relativamente infrecuentes a lo largo de un período de tiempo largo. Algunos pacientes continúan recibiendo tratamiento durante el resto de sus vidas.

35 La molécula de ácido nucleico aislada, el vector o los polipéptidos de FVIII de la divulgación se pueden administrar opcionalmente en combinación con otros agentes que son eficaces para tratar el trastorno o la afección que necesita tratamiento (*p. ej.*, profiláctico o terapéutico).

40 Tal como se utiliza en esta memoria, la administración de moléculas de ácido nucleico aisladas, vectores o polipéptidos de FVIII aislados de la divulgación junto o en combinación con una terapia adjunta significa la administración o aplicación secuencial, simultánea, coextensiva, concurrente, concomitante o contemporánea de la terapia. y los polipéptidos descritos. Los expertos en la técnica apreciarán que la administración o aplicación de los diversos componentes del régimen terapéutico combinado se puede programar para potenciar la eficacia general del tratamiento. Un experto en la materia (*p. ej.*, un médico) podría discernir fácilmente regímenes terapéuticos combinados eficaces sin una experimentación indebida basada en la terapia adjunta seleccionada y las enseñanzas de la presente memoria descriptiva.



Se apreciará, además, que la molécula de ácido nucleico aislada, el vector o el polipéptido de FVIII de la presente divulgación se pueden utilizar junto o en combinación con un agente o agentes (*p. ej.*, para proporcionar un régimen terapéutico combinado). Agentes ejemplares con los que se puede combinar un polipéptido o polinucleótido de la divulgación incluyen agentes que representan el patrón de atención actual para un trastorno particular que se está tratando.

5 Agentes de este tipo pueden ser de naturaleza química o biológica. El término "biológico" o la expresión "agente biológico" se refiere a cualquier agente farmacéuticamente activo elaborado a partir de organismos vivos y/o sus productos que se pretende utilizar como agente terapéutico.

10 La cantidad de agente a utilizar en combinación con los polinucleótidos o polipéptidos de la presente divulgación puede variar según el sujeto o puede administrarse de acuerdo con lo que se conoce en la técnica. Véase, *p. ej.*, Bruce A Chabner et al., *Antineoplastic Agents*, en GOODMAN & GILMAN'S THE PHARMACOLOGICAL BASIS OF THERAPEUTICS 1233-1287 (Joel G. Hardman et al., eds., 9ª ed. 1996). En otra realización, se administra una cantidad de un agente de este tipo consistente con el patrón de cuidado.

15 Tal como se comentó previamente, los polinucleótidos y polipéptidos de la presente divulgación pueden administrarse en una cantidad farmacéuticamente eficaz para el tratamiento *in vivo* de los trastornos de la coagulación. A este respecto, se apreciará que los polipéptidos o polinucleótidos de la divulgación se pueden formular para facilitar la administración y fomentar la estabilidad del agente activo. Preferiblemente, composiciones farmacéuticas de acuerdo con la presente divulgación comprenden un soporte estéril, no tóxico, farmacéuticamente aceptable, tal como solución salina fisiológica, tampones no tóxicos, conservantes y similares. Por supuesto, las composiciones farmacéuticas de la presente divulgación se pueden administrar en dosis únicas o múltiples para proporcionar una cantidad farmacéuticamente eficaz del polipéptido.

20 Hay disponible un cierto número de tests para evaluar la función del sistema de coagulación: test de tiempo de tromboplastina parcial activada (aPTT, por sus siglas en inglés), ensayo cromogénico, ensayo ROTEM®, test de tiempo de protrombina (PT, por sus siglas en inglés) (también utilizado para determinar INR), ensayo de fibrinógeno (a menudo por el método de Clauss), recuento de plaquetas, pruebas de función plaquetaria (a menudo por PFA-100), TCT, tiempo de sangrado, test de mezcla (si una anomalía se corrige si el plasma del paciente se mezcla con plasma normal), ensayo de factores de coagulación, antifosfolípidos anticuerpos, dímero D, tests genéticos (*p. ej.*, factor V de Leiden, mutación de protrombina G20210A), tiempo de veneno de víbora de Russell diluido (dRVVT, por sus siglas en inglés), diversos tests de función plaquetaria, tromboelastografía (TEG o Sonoclot), tromboelastometría (TEM®, *p. ej.*, ROTEM®), o tiempo de lisis de euglobulina (ELT).

30 El test de aPTT es un indicador del rendimiento que mide la eficacia tanto de la vía de coagulación "intrínseca" (a la que también se alude como vía de activación por contacto) como de la común. Este test se utiliza comúnmente para medir la actividad de coagulación de los factores de coagulación recombinantes disponibles comercialmente, *p. ej.*, FVIII o FIX. Se utiliza junto con el tiempo de protrombina (PT), que mide la vía extrínseca.

35 El análisis ROTEM® proporciona información sobre la cinética completa de la hemostasia: tiempo de coagulación, formación de coágulos, estabilidad del coágulo y lisis. Los diferentes parámetros de la tromboelastometría dependen de la actividad del sistema de coagulación plasmática, la función plaquetaria, la fibrinólisis o muchos factores que influyen en estas interacciones. Este ensayo puede proporcionar una visión completa de la hemostasia secundaria.

### Terapia Génica

40 La divulgación proporciona un método para aumentar la expresión de un polipéptido con actividad de FVIII en un sujeto que comprende administrar la molécula de ácido nucleico aislada de la divulgación a un sujeto que lo necesita, en el que la expresión del polipéptido aumenta con respecto a una molécula de ácido nucleico de referencia que comprende SEQ ID NO: 16. La divulgación también proporciona un método para aumentar la expresión de un polipéptido con actividad de FVIII en un sujeto, que comprende administrar un vector de la divulgación a un sujeto que lo necesita, en donde la expresión del polipéptido aumenta en relación a un vector que comprende una molécula de ácido nucleico de referencia.

45 La terapia génica somática se ha explorado como un posible tratamiento para la hemofilia A. La terapia génica es un tratamiento particularmente atractivo para la hemofilia debido a su potencial para curar la enfermedad a través de la producción endógena continua de FVIII tras una única administración de vector. La hemofilia A es muy adecuada para un enfoque de reemplazo de genes porque sus manifestaciones clínicas son completamente atribuibles a la falta de un solo producto génico (FVIII) que circula en cantidades diminutas (200 ng/ml) en el plasma.

50 Una proteína FVIII de la divulgación se puede producir *in vivo* en un mamífero, *p. ej.*, un paciente humano, utilizando un enfoque de terapia génica para el tratamiento de una enfermedad o trastorno de sangrado seleccionado del grupo que consiste en un trastorno hemorrágico de la coagulación, hemartrosis, sangrado muscular, hemorragia oral, hemorragia en los músculos, hemorragia oral, traumatismo, trauma capitis, sangrado gastrointestinal, hemorragia intracraneal, hemorragia intra-abdominal, hemorragia intratorácica, fractura ósea, sangrado del sistema nervioso central, sangrado en el espacio retrofaríngeo, sangrado en el espacio retroperitoneal y el sangrado en la vaina del iliopsoas sería terapéuticamente beneficioso. En una realización, la enfermedad o trastorno de sangrado es hemofilia. En otra realización,

la enfermedad o trastorno de sangrado es la hemofilia A. Esto implica la administración de un ácido nucleico que codifica FVIII optimizado enlazado operativamente a secuencias de control de la expresión adecuadas. En una determinada realización, estas secuencias se incorporan en un vector viral. Vectores virales adecuados para una terapia génica de este tipo incluyen vectores adenovirales, vectores lentivirales, vectores baculovirales, vectores virales de Epstein Barr, vectores papovavirales, vectores virales de vaccinia, vectores virales del herpes simple y vectores de virus adeno-asociados (AAV). El vector viral puede ser un vector viral defectuoso en la replicación. En otras realizaciones, un vector adenoviral tiene una delección en su gen E1 o en su gen E3. En otras realizaciones, las secuencias se incorporan en un vector no viral conocido por los expertos en la materia.

Todos los diversos aspectos, realizaciones y opciones descritos en esta memoria se pueden combinar en todas y cada una de las variaciones.

Habiendo descrito en general esta divulgación, se puede obtener una mayor comprensión por referencia a los ejemplos proporcionados en esta memoria. Estos ejemplos son solo para fines ilustrativos y no pretenden ser limitativos.

## EJEMPLOS

### Ejemplo 1. Estrategia de optimización por codones

Se crearon ocho variantes de BDD FVIII optimizadas por codones controlando el sesgo de uso de codones, incluyendo coFVIII-3 (SEQ ID NO: 1; FIG. 1A), coFVIII-4 (SEQ ID NO: 2; FIG. 1B), coFVIII -5 (SEQ ID NO: 70; FIG. 1C), coFVIII-6 (SEQ ID NO: 71; FIG. 1D), coFVIII-52 (SEQ ID NO: 3; FIG. 1E), coFVIII-62, coFVIII-62 (SEQ ID NO: 4; FIG. 1F), coFVIII-25 (SEQ ID NO: 5; FIG. 1G) y coFVIII-26 (SEQ ID NO: 6; FIG. 1H). La herramienta en línea Eugene se utilizó para facilitar la optimización por codones como se describió anteriormente (Véase Gaspar et al., "EuGene: maximizing synthetic gene design for heterologous expression," *Bioinformatics* 28:2683-84 (2012)), y se monitorizaron varios parámetros de uso de codones, tal como el índice de adaptación de codones (CAI) y el uso relativo de codones sinónimos (RSCU, por sus siglas en inglés) (Tabla 5). Todas las variantes se ajustaron a  $CAI \geq 83\%$  y  $RSCU \geq 1,63$ , mientras que la secuencia de FVIII con dominio B eliminado original, antes de la optimización, tiene un CAI de 74 % y una RSCU de 1,12 (Tabla 5).

**Tabla 5:** Parámetros de Optimización por Codones 70 71 3 4 5 6

	BDD FVIII Parental	coFVIII- 3	coFVIII- 4	coFVIII- 5	coFVIII- 6	coFVIII- 52	coFVIII- 62	coFVIII- 25	coFVIII- 26
Índice de Adaptación de Codones (CAI; %)	74	91	97	83	83	91	91	88	88
Frecuencia de Codones Óptimos (FOP)	39	65	92	64	64	79	79	74	75
Contenido de GC (%)	44,10	52,10	60,80	55,7	55,9%	58,30	58,30	57,30	57,60
Uso Relativo de Codones Sinónimos (RSCU)	1,12	2,32	2,72	1,63	1,63	2,22	2,19	2,04	2,58
Sesgo de Par de Codones	0,19	0,43	0,04	0,11	0,11	0,27	0,27	0,23	0,48
Número efectivo de codones	54,2	25,6	22,8	39,7	39,1	30,9	31,4	34,1	26,7

Además del aumento general del CAI, las ocho variantes se diseñaron en tres clases en función de la distribución del CAI en la región codificante, como se ilustra en la FIG. 2, con respecto a la secuencia de BDD FVIII no optimizada (FIG. 2A). La primera clase comprende variantes de BDD FVIII con una distribución uniforme del CAI alto en toda la región codificante (véanse las FIGs. 2C-2F). La primera clase incluye coFVIII-3 (FIG. 2C), coFVIII-4 (FIG. 2D), coFVIII-5 (FIG. 2E), coFVIII-6 (FIG. 2F), así como el coFVIII-1 descrito anteriormente (FIG. 2B). La segunda clase comprende variantes de BDD-FVIII con un CAI más bajo en la mitad N-terminal de la secuencia codificante y un CAI más alto en la mitad C-terminal de la secuencia codificante (véanse las FIGs. 2G y 2H). La segunda clase incluye coFVIII-52 (FIG. 2G) y coFVIII-62 (FIG. 2H). La tercera clase comprende variantes de BDD FVIII con un CAI superior en la mitad N-terminal de la secuencia codificante y un CAI inferior en la mitad C-terminal de la secuencia codificante (véanse las FIGs. 2I y 2J). La tercera clase incluye coFVIII-25 (FIG. 2I) y coFVIII-26 (FIG. 2J).

Sin estar ligados por teoría alguna, se especuló que un CAI más alto podría correlacionarse con una traducción de proteínas más rápida, y estas tres clases podrían representar diferentes tasas de síntesis de proteínas de principio a fin. Por ejemplo, la traducción de una región que tiene un CAI más bajo podría proceder lentamente en relación con la traducción de una región que tiene un CAI más alto. Si es así, la traducción de, *p. ej.*, la mitad N-terminal de coFVIII-52 de coFVIII-62, que tiene un CAI más bajo, podría proceder inicialmente lentamente seguida de una traducción más rápida de la mitad C-terminal, que tiene un CAI más alto. Esto podría preferirse para el plegamiento de proteínas y la modificación post-traducción durante la traducción sin ralentizar la síntesis de proteínas en general. El efecto opuesto podría verse para las variantes coFVIII-25 y coFVIII-26, que tienen un CAI más alto en la mitad N-terminal y un CAI más bajo en la mitad C-terminal.

Para garantizar la estabilidad del ARNm, todas las variantes optimizadas por codones de FVIII se ajustaron para evitar un cierto número de sitios, incluyendo sitios de corte y empalme crípticos, sitios poliA prematuros, motivos de inestabilidad de ARN (ARE) y secuencias repetidas, y para ajustar el contenido de GC (véase la Tabla 2).

### **Ejemplo 2. Clonación y Expresión de Variantes de coFVIII a partir de un Plásmido pcDNA3**

Se diseñaron plásmidos de expresión que contenían las diversas variantes de FVIII para la expresión *in vivo*. Los polinucleótidos BDD FVIII (FIG. 1I; SEQ ID NO: 16) y coFVIII-1 (FIG. 1J; SEQ ID NO: 68) no optimizados se clonaron en una cadena principal de pcDNA3 (Invitrogen), en donde el promotor CMV se reemplazó por un promotor de ET (véase la FIG. 3). Los plásmidos resultantes, FVIII-311 (BDD FVIII) y FVIII-303 (coFVIII-1), impulsan la expresión de BDD FVIII y coFVIII-1 no optimizados, respectivamente.

La expresión *in vivo* de FVIII-311 y FVIII-303 se evaluó en ratones Hem A mediante inyección hidrodinámica de 5 µg de ADN/ratón de FVIII-303 o FVIII-311. Se recogieron muestras de plasma a las 24, 48 y 72 horas posteriores a la inyección y se determinó la actividad de FVIII mediante un ensayo cromogénico específico para FVIII.

Como se muestra en la FIG. 4, la actividad de FVIII en plasma de ratones tratados con FVIII-311 (BDD FVIII; cuadrados) fue de  $74 \pm 43$  mU/mL a las 72 horas posteriores a la inyección, mientras que la actividad de FVIII en plasma de ratones tratados con FVIII-303 (coFVIII-1; círculos) fue de  $452 \pm 170$  mU/mL a las 72 horas posteriores a la inyección (FIG. 4). Esto representa un aumento de aproximadamente seis veces en la expresión de coFVIII-1 en relación con BDD FVIII no optimizado.

### **Ejemplo 3. Clonación y Expresión de Variantes de coFVIII Utilizando un Sistema de Vector Lentiviral**

Para evaluar adicionalmente el nivel de expresión de las variantes de BDD FVIII optimizadas por codones, las secuencias codificantes se clonaron en plásmidos lentivirales bajo el control de un promotor ET (véase Amendola et al., "Coordinate dual-gene transgenesis by lentiviral vectors carrying synthetic bidirectional promoters," *Nature Biol.* 23:108-16 (2005); Publicación Internacional N° WO 2000/066759 A1). En la FIG. 5 se muestra un mapa de plásmido de pLV-coFVIII-52; y plásmidos que contienen BDD FVIII no optimizado (LV-2116), coFVIII-1 (LV-coFVIII-1), coFVIII-3 (LV-coFVIII-3), coFVIII-4 (LV-coFVIII-4), coFVIII-5 (LV-coFVIII-5) y coFVIII-6 (LV-coFVIII-6), coFVIII-62 (LV-coFVIII-62), coFVIII-25 (LV-coFVIII-25) y coFVIII-26 (LV-coFVIII-26) se construyeron de la misma manera, excepto que el fragmento coFVIII-52 se reemplazó por cada una de las secuencias codificantes indicadas utilizando los sitios NheI y Sall (Tabla 6).

**Tabla 6:** Plásmidos de Expresión que Codifican Variantes de FVIII

ID de plásmido	Descripción
FVIII-303 (coFVIII-1)	coFVIII-1 bajo el promotor ET en pcDNA3
FVIII-311 (BDD FVIII)	BDD-FVIII parental bajo el promotor ET en pcDNA3
LV-2116 (BDD FVIII)	BDD-FVIII parental bajo el promotor ET en plásmido lentiviral
LV-coFVIII-1	coFVIII-1 bajo el promotor ET en plásmido lentiviral
LV-coFVIII-3	coFVIII-3 bajo el promotor ET en plásmido lentiviral
LV-coFVIII-4	coFVIII-4 bajo el promotor ET en plásmido lentiviral
LV-coFVIII-5	coFVIII-5 bajo el promotor ET en plásmido lentiviral
LV-coFVIII-6	coFVIII-6 bajo el promotor ET en plásmido lentiviral
LV-coFVIII-52	coFVIII-52 bajo el promotor ET en plásmido lentiviral
LV-coFVIII-62	coFVIII-62 bajo el promotor ET en plásmido lentiviral
LV-coFVIII-25	coFVIII-25 bajo el promotor ET en plásmido lentiviral
LV-coFVIII-26	coFVIII-26 bajo el promotor ET en plásmido lentiviral

Las variantes de FVIII optimizadas por codones lentivirales se evaluaron en ratones Hema mediante inyección hidrodinámica a una dosis de 5 µg de ADN/ratón (FIGs. 6A, 6B) o 20 µg de ADN/ratón (FIG. 6C). Como se muestra en la FIG. 6, cada uno de coFVIII-3 (FIG. 6A; triángulos), coFVIII-4 (FIG. 6A; triángulos invertidos), coFVIII-5 (FIG. 6A; rombos), coFVIII-6 (FIG. 6A; círculos en blanco), coFVIII-25 (FIG. 6B; triángulos), coFVIII-26 (FIG. 6B; triángulos invertidos), coFVIII-52 (FIG. 6C; cuadrados) y coFVIII-62 (FIG. 6C; círculos negros) exhibieron una mayor actividad de FVIII que coFVIII-1 (FIG. 6A, círculos; FIG. 6B, círculos; y FIG. 6C, triángulos). En particular, coFVIII-25 y coFVIII-26 exhibieron un nivel de expresión similar a las 72 horas posteriores a la inyección, alcanzando una actividad aproximadamente 3 veces mayor que la del coFVIII-1 (FIG. 6B), lo que se traduce en una actividad de FVIII 24 veces mayor en comparación con el BDD FVIII parental no optimizado (véase la FIG. 4). Tanto coFVIII-52 (cuadrados) como coFVIII-62 (círculos negros) lograron una expresión incluso mayor a las 72 horas posteriores a la inyección, exhibiendo una expresión 6 veces y 4 veces mayor, respectivamente, que coFVIII-1 (triángulos) y una expresión 50 veces y 30 veces mayor, respectivamente, que BDD FVIII parental no optimizado (círculos blancos) (FIG. 6C). Estos datos indican que la combinación de un CAI más bajo en la mitad N-terminal de la secuencia codificante y un CAI más alto en la mitad C-terminal de la secuencia codificante podría ser más beneficiosa para la expresión de FVIII en comparación con la distribución inversa de CAI.

#### **Ejemplo 4: Expresión Lentiviral a Largo Plazo de Variantes de FVIII Optimizadas por Codones en Ratones Hema**

Variantes a las que se identificó por impulsar una elevada expresión de FVIII en Hema a las 72 horas posteriores a la inyección hidrodinámica se evaluaron para la expresión de FVIII a largo plazo mediante transferencia génica mediada por vectores lentivirales. Los vectores lentivirales se produjeron en células 293T mediante transfección transitoria y se concentraron mediante ultracentrifugación hasta aproximadamente 5E9 TU/ml. A continuación, los vectores lentivirales se administraron a ratones Hema de 12-14 días de edad mediante inyección retro-orbital a una dosis de 1E8 TU/ratón. A los 21 días después de la inyección lentiviral, la actividad media de FVIII en plasma fue de aproximadamente 0,04 UI/ml para ratones a los que se inyectó LV-2116 (BDD FVIII; FIG. 7). Cada uno de coFVIII-1, coFVIII-5, coFVIII-52, coFVIII-6 y coFVIII-62 dio como resultado un nivel de FVIII circulante más alto a los 21 días posteriores a la inyección en relación con LV-2116 (FVIII con dominio B no optimizado eliminado) control. En particular, la inyección de coFVIII-1 y coFVIII-5 produjo niveles de actividad plasmática de FVIII de aproximadamente 1,8 UI/mL, coFVIII-52 produjo un nivel de actividad plasmática de FVIII de aproximadamente 4,9 UI/mL, coFVIII-6 produjo niveles de actividad plasmática de FVIII de aproximadamente 4,6 UI/mL y coFVIII-62 produjo un nivel de actividad plasmática de FVIII de aproximadamente 2,5 UI/mL a los 21 días posteriores a la inyección (FIG. 7). Los niveles plasmáticos de FVIII observados en ratones a los que se inyectaron LV-coFVIII-6 y LV-coFVIII-52, 4,6 UI/ml y 4,9 UI/ml, respectivamente, son más de 100 veces mayores que los niveles plasmáticos observados en ratones a los que se inyectó el LV-2116 (BDD-FVIII no optimizado) control.

#### **Ejemplo 5. Construcciones de Fusión coFVIII- XTEN**

Se testó la capacidad de XTEN para mejorar la expresión de FVIII en estado estacionario. En primer lugar, la secuencia codificante de un XTEN de 144 aminoácidos ("XTEN<sub>144</sub>"; SEQ ID NO: 18) se insertó en el nucleótido 1193 (o después de los primeros 764 aminoácidos del polipéptido codificado) de coFVIII-52 y coFVIII-1 para generar coFVIII-52-XTEN (FIG. 8A; SEQ ID NO: 19) y coFVIII-1-XTEN (FIG. 8B; SEQ ID NO: 20), respectivamente. La secuencia de coFVIII-1-XTEN se clonó luego en una cadena principal de pcDNA3 (Invitrogen) bajo el control de un promotor ET, como se describió arriba,

para crear el plásmido de expresión de FVIII-306; y la secuencia de coFVIII-52-XTEN se clonó en un plásmido lentiviral bajo el control de un promotor ET, como se describió arriba, para crear el pLV-coFVIII-52-XTEN (FIG. 9). Se administró FVIII-306 (coFVIII-1-XTEN) a ratones HemA a razón de 5 µg de ADN/ratón mediante inyección hidrodinámica. En comparación con FVIII-303 (coFVIII-1; FIG. 10A, círculos pequeños) y FVIII-311 (BDD FVIII; FIG. 10A, cuadrados), la fusión de XTEN<sub>144</sub> a coFVIII-1 (FVIII-306; FIG. 10A, círculos grandes) dio como resultado una expresión de FVIII aproximadamente 5 veces y 33 veces mayor, respectivamente, en ratones HemA a las 72 horas posteriores a la inyección. El efecto de la inserción de XTEN sobre la expresión de FVIII también se evaluó utilizando un vector lentiviral en ratones HemA (FIG. 10B). Se administró LV-coFVIII-52-XTEN a ratones HemA de 12-14 días de edad a razón de 1E8 UT/ratón mediante inyección retro-orbital. En comparación con LV-coFVIII52 y LV-2116 (BDD-FVIII), la fusión de XTEN<sub>144</sub> con coFVIII-52 (FIG. 10B) dio como resultado una expresión de FVIII aproximadamente 4 veces y 450 veces mayor, respectivamente, en ratones HemA a 21 días posteriores a la inyección.

Los vectores lentivirales que comprenden cada uno de coFVIII-3, co-FVIII-4, coFVIII-5, coFVIII-6, coFVIII-62, coFVIII-25 y coFVIII-26 condensados con XTEN<sub>144</sub> y condensados con un promotor ET se harán como se describe arriba. Los vectores se testarán para determinar su expresión de proteínas FVIII.

#### Ejemplo 6. Expresión de Construcciones de coFVIII

Variantes de FVIII optimizadas por codones se clonaron en plásmidos lentivirales, como se ilustra en la FIG. 9, mediante técnicas estándares de clonación molecular. A continuación, se produjeron vectores lentivirales en células HEK293 mediante transfección transitoria y se aislaron mediante ultracentrifugación.

Vectores lentivirales de FVIII se administraron a ratones HemA de 14 días de edad mediante inyección intravenosa a una dosis de 1,5E10 TU/kg variante de LV-FVIII. La actividad plasmática de FVIII se midió el día 21 posterior al tratamiento con LV-FVIII, y el número de copias del vector (VCN, por sus siglas en inglés) por célula se midió en muestras de necropsia hepática recogidas de animales tratados con LV-FVIII el día 150 post-tratamiento con LV-FVIII. Si bien los valores de VCN fueron similares en todos los animales, independientemente de las variantes de LV-FVIII administradas (FIG. 12B), los niveles de actividad de FVIII en los animales tratados con variantes de coFVIII fueron de 30 a 100 veces mayores que en los animales tratados con wtBDD -FVIII (FIGs. 12A y 12C; Tabla 7). Estos datos indican que la optimización por codones de FVIII mejora la expresión de FVIII en un entorno de vector lentiviral.

**Tabla 7:** Expresión relativa de construcciones de FVIII optimizadas por codones

Variantes de LV-FVIII	CoFVIII-1	CoFVIII-3	CoFVIII-4	CoFVIII-5	CoFVIII-6
<b>Mejora de veces de la expresión de FVIII en relación con LV-wtBDD-FVIII</b>	57	34	33	74	107
Variantes de LV-FVIII	CoFVIII-52	CoFVIII-62	CoFVIII-25	CoFVIII-26	
<b>Mejora de veces de la expresión de FVIII en relación con LV-wtBDD-FVIII</b>	96	59	87	83	

#### Ejemplo 7. Respuesta Inmunitaria Mediada por la Expresión Transgénica de FVIII en Ratones HemA Después de Tratamiento Lentiviral

Los ratones tratados con LV-FVIII del Ejemplo 6 se evaluaron para determinar la expresión de FVIII a largo plazo y la formación de anticuerpos anti-FVIII. La expresión de FVIII, como se pone de manifiesto por la actividad plasmática de FVIII, varió entre animales dentro del mismo grupo de tratamiento (FIG. 13A). Por ejemplo, tres ratones (designados 1, 2 y 3) tratados con un vector lentiviral que expresa la variante coFVIII-5 mostraron una expresión constante de FVIII durante aproximadamente 16 semanas, mientras que tres compañeros de camada (designados 4, 5 y 6), que fueron tratados con el mismo vector lentiviral, mostró fuertes descensos en los niveles de actividad plasmática de FVIII en aproximadamente 10 semanas post-tratamiento (Figura 13A). La actividad plasmática constante de FVIII observada en los ratones 1, 2 y 3 se correlacionó con niveles no detectables o muy bajos de anticuerpos anti-FVIII (FIG. 13B; ratones 1, 2 y 3). Por el contrario, los ratones que exhibieron fuertes disminuciones en la actividad plasmática de FVIII también exhibieron niveles incrementados de anticuerpos anti-FVIII (FIG. 13B; ratones 4, 5 y 6). Estos datos sugieren que la expresión transgénica de FVIII induce la formación de anticuerpos anti-FVIII en un subconjunto de animales, y los anticuerpos anti-FVIII resultantes eliminaron la proteína FVIII transgénica de la circulación.

Se evaluó la relación entre la expresión de FVIII y la formación de anticuerpos anti-FVIII. Los ratones tratados con LV-FVIII del Ejemplo 6 se dividieron en dos grupos: ratones que eran negativos para anticuerpos anti-FVIII y ratones que eran

positivos para anticuerpos anti-FVIII. Como se muestra en la FIG. 14, la expresión de FVIII transgénico a niveles fisiológicos no induce una respuesta inmunitaria al FVIII transgénico (FIG. 14, círculos). Sin embargo, los niveles suprafisiológicos de expresión de FVIII parecen inducir la formación de anticuerpos anti-FVIII, de modo que cuanto mayor sea la expresión de FVIII nivel, mayor será la probabilidad de inducción de anticuerpos anti-FVIII. Estos datos sugieren que puede ser beneficioso mantener los niveles fisiológicos de expresión de FVIII en pacientes sometidos a un tratamiento de terapia génica de FVIII.

Para determinar si la respuesta inmunitaria inducida por la expresión de FVIII da como resultado la pérdida de células hepáticas que expresan el transgén, se evaluaron el número de copias del vector (FIG. 15) y el nivel de transcripción de ARN de FVIII (FIG. 16) en muestras de necropsia hepática de ratones con anticuerpos anti-FVIII positivos y negativos. . Como se muestra en la FIG. 15, la distribución del número de copias del vector fue la misma en ratones positivos y negativos para anticuerpos anti-FVIII, lo que indica que las células con integración de LV-FVIII se mantuvieron a pesar del desarrollo de anticuerpos anti-FVIII. Esto sugiere que la dosis de expresión transgénica de FVIII mediada por LV-FVIII no induce una respuesta de linfocitos T citotóxicos (CTL) contra las células hepáticas que expresan FVIII. Para confirmar adicionalmente estos resultados, se evaluó la transcripción de ARN de FVIII mediante hibridación *in situ* de ARN (FIGs. 16C y 16D). En el momento de la extracción del hígado, el coFVIII-52-B de ratón no tenía FVIII circulante detectable y un alto nivel de anticuerpos anti-FVIII (Figuras 16A y 16B). Sin embargo, la señal de transcripción de ARN y el número de células positivas para ARN de FVIII en el tejido hepático del ratón coFVIII-52-B fueron equiparables a los del ratón FVIII-52-A, que tenía aproximadamente 4 UI/ml de FVIII circulante en el momento de la necropsia. Por lo tanto, la expresión de FVIII no indujo la respuesta de CTL en ratones HemA experimentales.

#### **Ejemplo 8. Expresión a Largo Plazo de FVIII en Neonatos de Ratón HemA Tratados con LV-FVIII**

Para evaluar la eficacia del uso de un sistema lentiviral para el tratamiento de pacientes pediátricos con HemA finado como objetivo el hígado, se administró a ratones HemA de 2 días de edad mediante inyección en la vena de la sien aproximadamente 1,5 E10 TU/kg de LV-coFVIII-52XTEN, LV-coFVIII6-XTEN, o un vector lentiviral que expresa wtBDD-FVIII. Se observó una expresión de FVIII constante a largo plazo para ambas variantes y el control, lo que demuestra que la cascada de expresión de FVIII integrada se mantuvo en las células hepáticas en división de los ratones tratados (FIG. 17). Estos datos sugieren que LV-FVIII podría utilizarse potencialmente para tratar pacientes HemA tanto pediátricos como adultos.

#### **Ejemplo 9. Evaluación de LV-FVIII en Neonatos de Perros HemA**

Para evaluar adicionalmente la eficacia de LV-FVIII en modelos animales más grandes, se administraron mediante inyección intravenosa  $1,3 \times 10^9$  TU/kg de LV-coFVIII-6-XTEN a dos perros neonatos HemA de una semana de edad (designados S3 y K4). Esta dosis fue más de 10 veces menor que la dosis utilizada anteriormente en modelos de ratones HemA. Después de la administración de los vectores lentivirales, se controló la actividad del FVIII en plasma mediante un ensayo de coagulación de una etapa (aPTT) (FIG. 18) y se controló la hemostasia de sangre entera mediante un ensayo de tromboelastometría rotacional (ROTEM) (FIG. 19A-19D). Antes del tratamiento con LV-FVIII, el nivel de FVIII para S3 era del 0,7 % del normal (FIG. 18). Después del tratamiento con vector lentiviral, el nivel de FVIII de S3 aumentó al 79 % y al 103 % de lo normal en el día 7 y el día 14, respectivamente (FIG. 18). Para K4, el nivel de FVIII pre-administración fue del 1,4% del normal (FIG. 18). Después del tratamiento con vector lentiviral, el nivel de FVIII aumentó al 22 % y al 25 % de lo normal en el día 6 y el día 14, respectivamente (FIG. 18).

Correlacionado con el nivel de FVIII, se observó ROTEM normalizado para ambos animales 2 semanas después del tratamiento (FIGs. 19A-19C), lo que demuestra el beneficio terapéutico mediado por LV-FVIII. El nivel de expresión de FVIII terapéuticamente beneficioso alcanzado por LV-FVIII en perros HemA confirma el uso potencial de LV-FVIII para el tratamiento de la hemofilia A.

#### **Ejemplo 10. Evaluación de LV-FVIII en Neonatos de Ratón HemA**

La terapia génica *ex vivo* con vectores lentivirales (LV) para el reemplazo de genes ha demostrado eficacia clínica para múltiples indicaciones y con un seguimiento de varios años en pacientes tratados que no muestran evidencia de tumorigénesis. El suministro sistémico de LV-FIX media en la expresión persistente de FIX y se tolera bien en modelos animales con hemofilia. La gran capacidad de empaquetamiento, la capacidad de mantener la expresión transgénica a largo plazo a través de la integración de genes, la falta de anticuerpos anti-LV pre-existentes (abs) en poblaciones humanas y los perfiles *in vivo* alentadores demostrados en entornos preclínicos y clínicos, hacen de LV un vehículo prometedor para la administración de genes *in vivo*, especialmente para genes candidatos con un gran tamaño de ADNc tal como el FVIII.

Para evaluar el uso potencial de LV-FVIII para el tratamiento de la hemofilia A (HemA), variantes de FVIII humano optimizadas por codones (hFVIII) colocadas bajo un promotor específico para hepatocitos se incorporaron en un sistema LV que contiene múltiples copias de microRNA-142 secuencias diana para minimizar la expresión de FVIII en células presentadoras de antígenos y reducir la probabilidad de inducir anticuerpos anti-FVIII. Los vectores LV- hFVIII se

produjeron mediante transfección transitoria de células 293T, seguido de una concentración de 1000 veces mediante ultracentrifugación y evaluados en modelos de ratón HemA. Después de la administración intravenosa de LV- hFVIII, el nivel de hFVIII circulante se controló mediante ensayos de actividad de FVIII y de antígeno, la eficiencia de transducción de LV en el hígado se evaluó midiendo las copias de ADN de LV mediante PCR cuantitativa y el ARN transgénico mediante hibridación in situ, se midieron los anticuerpos anti -hFVIII mediante ELISA de anticuerpos anti-hFVIII totales.

Se observó una expresión persistente de FVIII para todas las variantes de LV- hFVIII en ratones HemA que fueron tratados en la fase neonatal. A una dosis de  $1,5 \times 10^{10}$  unidades de transducción/kg, el hFVIII optimizado por codones codificante de LV (LV-cohFVIII) dio como resultado un FVIII circulante de 30 a 100 veces mayor que el hFVIII de tipo salvaje codificante de LV (FIG. 12C), mientras que el número de copias del vector en células hepáticas y el porcentaje de células positivas para ARN de FVIII fue equiparable en todos los grupos testados (FIG. 12B). La combinación de la optimización por codones con XTEN (LV-cohFVIII-XTEN), un polipéptido hidrofílico no estructurado que supuestamente mejora la semivida circulante al aumentar el tamaño hidrodinámico de la carga útil, dio como resultado una actividad de FVIII de 30-50 UI/mL en plasma, que representa del 3.000 al 5.000 % del nivel de FVIII circulante normal (FIG. 12A, FIG. 17). Además, los abs (siglas inglesas de anticuerpos) anti-hFVIII solo se detectaron en ratones con un nivel suprafisiológico de hFVIII (FIG. 14), pero no se observó respuesta de linfocitos T citotóxicos contra las células transducidas de LV en ratones positivos para anticuerpos anti-hFVIII (FIG.s 15 y 16A-16D). El resultado de los autores de la invención respalda un desarrollo adicional de LV-FVIII para la terapia génica in vivo de la hemofilia A.

#### LISTADO DE SECUENCIAS

<110> BIOVERATIV THERAPEUTICS INC.

<120> GENES DE FACTOR VIII OPTIMIZADOS

<130> 2159.469PC02

<150> 62/409.739 <151> 18-10-2016

<150> 62/289.696 <151> 01-02-2016

<160> 103

<170> PatentIn version 3.5

<210> 1

<211> 4374

<212> ADN

<213> Secuencia Artificial

<220>

<223> coFVIII-3

<400> 1



# ES 2 926 585 T3

atgcagatcg aactgagcac ctgcttcttt ctgtgcctgc tgaggttttg ctttagcgcc	60
accaggagat actatctggg cgccgtggaa ctgagctggg actatatgca gtctgatctg	120
ggcgaaactgc cagtggatgc caggtttccc ccagagtgcc ccaaagctt tccctttaat	180
accagcgtgg tgtataagaa aaccctgttt gtggaattca ctgatcatct gtttaatatc	240
gccaagccca ggccccctg gatgggcctg ctgggccccca ccatccaggc cgaagtgtat	300
gataccgtgg tcatcaccct gaaaaacatg gccagccatc cagtgagcct gcatgctgtg	360
ggcgtctcct attggaaagc ctctgaaggc gccgagtatg atgatcagac cagccagagg	420
gaaaaagaag atgataaagt ctttcctggg ggagccata cctatgtctg gcaggtcctg	480
aaagaaaatg gcccctatggc cagcgatccc ctgtgcctga cctatagcta tctgagccat	540
gtggacctgg tgaaggatct gaacagcggc ctgattgggg ccctgctggt gtgcagggaa	600
ggcagcctgg ccaaagaaaa aaccagacc ctgcataagt ttatcctgct gtttgccgtg	660
tttgatgaag gcaaaagctg gcattctgaa accaaaaaca gcctgatgca ggacagggat	720
gccgcctctg ccagggcctg gcccaaatg cataccgtga atggctatgt gaataggagc	780
ctgcctggcc tgattggctg ccacaggaaa agcgtgtatt ggcatgtgat cggcatgggc	840
accacccccg aagtgcatag catctttctg gaaggccata ccttcctggt caggaaccac	900
aggcaggcca gcctggaaat cagcccatc accttcctga ccgccagac cctgctgatg	960
gatctgggcc agtttctgct gttttgccac atctccagcc atcagcatga tggcatggaa	1020
gcctatgtga aagtcgatag ctgccccgaa gaacccagc tgaggatgaa aaacaatgaa	1080

# ES 2 926 585 T3

gaagccgaag actatgatga tgatctgact gattctgaaa tggatgtggt caggtttgat	1140
gatgataata gccccagctt tatccagatc aggagcgtgg ccaaaaaaca tccaagacc	1200
tgggtgcatt atatcgctgc tgaggaagaa gattgggact atgccccct ggtgctggcc	1260
cctgatgata ggagctataa aagccagtat ctgaacaatg gccccagag gattggcagg	1320
aagtataaaa aagtcagggtt tatggcctac actgatgaaa ctttcaagac caggaagcc	1380
atccagcatg agtctggcat cctgggcccc ctgctgtatg gcgaagtggg ggacaccctg	1440
ctgatcatct ttaaaaatca ggccagcagg cctataata tctatcccca tggcatcact	1500
gatgtgaggg ccctgtacag caggaggctg cccaaaggcg tgaaacatct gaaagatttt	1560
cccatcctgc ctggcgaaat ctttaagtat aaatggactg tgactgtgga agatggcccc	1620
acaaaagcg atcccagggtg cctgaccagg tattattcca gctttgtgaa tatggaacgc	1680
gatctggcct ctggcctgat tggccccctg ctgatctgct ataaagagtc tgtggaccag	1740
aggggcaatc agatcatgag cgataaaagg aatgtcatcc tgttctctgt ctttgatgag	1800
aataggagct ggtacctgac cgaaaacatc cagaggtttc tgcccaatcc cgcggcgtg	1860
cagctggaag atcccaggtt tcaggccagc aatatcatgc atagcatcaa tggctatgtc	1920
tttgatagcc tgcagctgag cgtgtgcctg catgaggtgg cctattggta taccctgagc	1980
atcggcgccc agaccgattt tctgagcgtg ttttctctg gctatacctt taaacataaa	2040
atggtgtatg aggacaccct gaccctgttt cccttctctg gcgaaaccgt gtttatgagc	2100
atgaaaaatc ccggcctgtg gatcctgggc tgccacaaca gcgatttcag gaacaggggc	2160
atgactgccc tgctgaaagt ctccagctgc gataaaaaaca ctggggacta ttatgaggac	2220
agctatgagg acatcagcgc ctatctgctg agcaagaaca atgccatcga acccaggagc	2280
tttagccaga atccccaggt gctgaaaagg catcagaggg aaatcaccag gaccaccctg	2340
cagtctgatc aggaagaaat cgactatgat gataccatca gcgtggaaat gaagaaagaa	2400
gattttgata tctatgatga agatgaaaat cagagcccca ggagctttca gaagaaaacc	2460
aggcattact tcatcgctgc tgtggaagg ctgtgggact atggcatgtc cagcagcccc	2520
catgtgctga ggaacagggc ccagtctggc agcgtgcccc agtttaaaaa agtcgtgttt	2580
caggagttaa ccgatggcag ctttaccag cccctgtata ggggcgaact gaatgaacat	2640
ctgggcctgc tgggccccta catcagggcc gaagtggaag ataatatcat ggtgaccttc	2700
aggaaccagg ccagcaggcc ctacagcttt tattccagcc tgatcagcta tgaggaagat	2760
cagaggcagg gggctgagcc caggaaaaac tttgtgaaac ccaatgaaac caagacctac	2820
ttttggaaag tccagcatca tatggcccc accaaggatg aatttgattg caaagcctgg	2880
gcctacttct ctgatgtgga cctggaaaaa gatgtgcata gcggcctgat tggccccctg	2940
ctggtgtgcc acaccaatac cctgaaccct gcccatggca ggcagggtgac tgtgcaggag	3000

tttgcctgt	tctttacat	ctttgatgaa	acccaaagct	ggtacttcac	cgaaaacatg	3060
gaaaggaact	gcagggcccc	ctgcaacatc	cagatggaag	atcccacctt	taaagaaaat	3120
tataggttcc	atgccatcaa	tggctatatc	atggataccc	tgcttggcct	ggtcatggcc	3180
caggaccaga	ggatcaggtg	gtatctgctg	agcatgggca	gcaatgaaaa	catccatagc	3240
atccatttct	ctggccatgt	ctttaccgtc	aggaaaaaag	aagagtataa	aatggccctg	3300
tataatctgt	accctggggg	gtttgaaacc	gtggaaatgc	tgcccagcaa	agccggcatc	3360
tggaggggtg	aatgcctgat	tggcgaacat	ctgcatgctg	gcatgagcac	cctgtttctg	3420
gtgtatagca	ataagtgcc	gacccccctg	ggcatggcct	ctggccatat	cagggatttt	3480
cagatcactg	cctctggcca	gtatggccag	tgggccccca	aactggccag	gctgcattat	3540
tccggaagca	tcaatgcctg	gagcaccaaa	gaacccttta	gctggatcaa	agtcgatctg	3600
ctggccccc	tgatcatcca	tggcatcaag	accagggggg	ccaggcagaa	attttccagc	3660
ctgtacatca	gccagtttat	catcatgtat	agcctggatg	gcaaaaaatg	gcagacctat	3720
aggggcaata	gcaccggcac	cctgatgggtg	ttctttggca	atgtggacag	cagcggcatc	3780
aaacataata	tctttaatcc	cccatcatc	gccaggtata	tcaggctgca	tcccacccat	3840
tatagcatca	ggagcacctc	gaggatggaa	ctgatgggct	gcgatctgaa	cagctgcagc	3900
atgcccttgg	gcatggaaag	caaagccatc	agcgatgccc	agatcactgc	ctccagctac	3960
ttcactaata	tgtttgccac	ctggagcccc	agcaaagcca	ggctgcatct	gcagggcagg	4020
agcaatgcct	ggaggcccca	ggtcaacaat	cccaaagaat	ggctgcaggt	cgattttcag	4080
aaaaccatga	aagtcaactg	ggtgaccacc	caggggggtga	aaagcctgct	gaccagcatg	4140
tatgtgaaag	agtttctgat	ctccagcagc	caggatggcc	atcagtggac	cctgtttctt	4200
cagaatggca	aagtcaaagt	ctttcagggc	aatcaggaca	gctttacccc	tgtggtgaat	4260
agcctggatc	ccccctgct	gaccaggtat	ctgaggatcc	atccccagag	ctgggtgcat	4320
cagatcgccc	tgaggatgga	agtgctgggc	tgcaagccc	aggacctgta	ctga	4374

<210> 2

<211> 4374

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> coFVIII-4

<400> 2

atgcagatcg	agctgagcac	gtgcttcttc	ctgtgcctgc	tgaggttctg	cttcagcgcc	60
accaggaggt	actacctggg	cgccgtggag	ctgagctggg	actacatgca	gagcgacctg	120
ggcgagctgc	ccgtggacgc	caggttcccc	cccagggtgc	ccaagagctt	ccccttcaac	180

10

# ES 2 926 585 T3

acgagcgtgg	tgtacaagaa	gaccctgttc	gtggagttca	ccgaccatct	gttcaatata	240
gccaagccca	ggccccctg	gatggggctg	ctggggccca	cgatccaggc	cgaggtgtac	300
gacaccgtgg	tcatcacctt	gaagaacatg	gccagccacc	ccgtgagcct	gcacgccctg	360
ggcgtgagct	actggaaggc	cagcgagggc	gccgagtacg	acgaccagac	cagccagagg	420
gagaaggagg	acgacaaggt	gttccccggc	ggcagccaca	cctacgtgtg	gcaggtgctg	480
aaggagaatg	ggcccatggc	cagcgacccc	ctgtgcctga	cctactctta	cctgagccac	540
gtggatctgg	tgaaggacct	gaacagcggc	ctgatcggcg	ccctgctggt	gtgcagggag	600
ggcagcctgg	ccaaggagaa	gaccagacc	ctgcacaagt	tcatcctgct	gttcgccctg	660
ttcgacgagg	gcaagagctg	gcacagcgag	accaagaaca	gcctgatgca	ggatagggac	720
gccgccagcg	ccagggcctg	gcccgaagt	cacaccgtga	acggctacgt	gaacaggctt	780
ctgcccgggc	tgatcggtg	ccacaggaag	agcgtgtact	ggcagctgat	cgccatgggg	840
accacccccg	aggtgcacag	catcttcctg	gagggccaca	cgctcctggt	gaggaatcac	900
aggcaggcca	gcctggagat	cagcccgatc	accttcctga	ccgccagac	cctgctgatg	960
gacctggggc	agttcctgct	gttctgccat	atcagctctc	accagcacga	cgccatggag	1020
gcctacgtga	aggtggatag	ctgccccgag	gagccccagc	tgaggatgaa	gaacaacgag	1080
gaggccgagg	actacgacga	cgacctgacc	gacagcgaga	tggacgtggt	gaggttcgac	1140
gacgacaata	gcccgaagctt	catccagatc	aggagcgtgg	ccaagaagca	ccccagacc	1200
tgggtgcatt	acatcgccgc	cgaggaggag	gattgggact	acgccccctt	ggtgctggcc	1260
cccagcgaca	ggtcttacaa	gagccagtac	ctgaacaacg	ggccccagag	gatcggcagg	1320
aagtacaaga	aggtgaggtt	catggcctac	accgacgaga	ccttcaagac	cagggaggcg	1380
atccagcacg	agagcgggat	cctggggccc	ctgctgtacg	gcgaggtggg	cgacacgctg	1440
ctgatcatct	tcaagaacca	ggccagcagg	ccgtacaata	tctaccccca	cgggatcacc	1500
gacgtgaggc	ccctgtactc	taggaggctg	cccaaggggc	tgaagcacct	gaaggacttc	1560
cccatcctgc	ccggcgagat	cttcaagtac	aagtggaccg	tgaccgtgga	ggacggggcc	1620
acgaagagcg	acccaggtg	cctgaccagg	tactacagct	ctttcgtgaa	catggagagg	1680
gacctggcca	gcggcctgat	cggggccctg	ctgatctgct	acaaggagag	cgtggatcag	1740
aggggcaacc	agatcatgag	cgacaagagg	aacgtgatcc	tgttcagcgt	gttcgacgag	1800
aataggtctt	ggtacctgac	cgagaatata	cagaggttcc	tgcccaaccc	cgccggcgtg	1860
cagctggagg	atcccagatt	ccaggccagc	aacatcatgc	acagcatcaa	cggtacgtg	1920
ttcgacagcc	tgacagctgag	cgtgtgcctg	cacgaggtgg	cctactggta	catcctgagc	1980
atcggcgccc	agaccgactt	cctgagcgtg	ttcttcagcg	gtacacctt	caagcacaag	2040
atggtgtacg	aggataccct	gaccctgttc	cccttcagcg	gcgagaccgt	gttcatgagc	2100

# ES 2 926 585 T3

atggagaacc ccggcctgtg gatcctgggc tgccataact ccgacttcag gaataggggc	2160
atgaccgccc tgctgaaggt gagctcttgc gacaagaaca ccggcgacta ctacgaggat	2220
agctacgagg atatcagcgc ctacctgtg agcaagaaca acgccatcga gccaggtct	2280
ttcagccaga accccccgt gctgaagagg caccagagg agatcaccag gacgacctg	2340
cagagcgacc aggaggagat cgactacgac gacacgatca gcgtggagat gaagaaggag	2400
gatttcgaca tctacgacga ggacgagaat cagagcccca ggtctttcca gaagaagacc	2460
aggcattact tcatcgccgc cgtggagagg ctgtgggact acggcatgag cagctctccc	2520
cacgtgctga ggaatagggc ccagagcggc agcgtgcccc agttcaagaa ggtggtgttc	2580
caggagtcca ccgacggcag cttcaccag cccctgtaca ggggcgagct gaacgagcac	2640
ctgggcctgc tggggcccta catcagggcc gaggtggagg ataacatcat ggtgaccttc	2700
aggaatcagg ccagcaggcc ctatagcttc tatagctctc tgatcagcta cgaggaggat	2760
cagaggcagg gcgccgagcc caggaagaac ttcgtgaagc ccaacgagac caagacctac	2820
ttctggaagg tgcagcacca catggcccc acgaaggacg agttcgactg caaggcctgg	2880
gcctacttca gcgacgtgga tctggagaag gacgtgcaca gcggcctgat cgggcccctg	2940
ctggtgtgcc acaccaacac cctgaacccc gccacggca ggcagggtgac cgtgcaggag	3000
ttcgccctgt tcttcacat cttcgacgag accaagagct ggtacttcac cgagaatatg	3060
gagaggaatt gcagggcccc ctgcaatata cagatggagg acccgacctt caaggagaat	3120
tacaggttcc acgccatcaa cggctacata atggacacgc tgcccggcct ggtcatggcc	3180
caggatcaga ggatcagggt gtatctgtg agcatgggga gcaacgagaa tatccacagc	3240
atccaattca gcggccacgt gttcacctg aggaagaagg aggagtacaa gatggccctg	3300
tacaatctgt accccggcgt gttcgagacc gtggagatgc tgcccagcaa ggccgggac	3360
tggagggagg agtgcctgat cggcgagcac ctgcacgccg gcatgagcac gctgttcctg	3420
gtgtactcta acaagtgcc aacccccctg gggatggcca gcggccacat cagggaactc	3480
cagatcaccc ccagcggcc gtacggccag tggggcccca agctggccag gctgcactat	3540
tccggaagca tcaacgcctg gagcacgaag gagcccttca gctggatcaa ggtggatctg	3600
ctggccccc tgatcatcca cgggatcaag acccagggcg ccaggcagaa gttcagctct	3660
ctgtatatca gccagttcat catcatgtac tctctggacg gcaagaagtg gcagacctac	3720
aggggcaaca gcaccggcac gctgatggtg ttcttcggca acgtggactc tagcgggac	3780
aagcacaata tcttcaaccc ccccatcata gccaggtaga tcagggtgca cccacccat	3840
tactctatca ggtctaccct gaggatggag ctgatgggct gcgacctgaa cagctgcagc	3900
atgccctgg ggatggagag caaggccatc agcgacgcc agatcacgc cagctcttac	3960

# ES 2 926 585 T3

ttcaccaaca	tgttcgccac	ctggagcccg	agcaaggcca	ggctgcacct	gcagggcagg	4020
tctaacgcct	ggaggcccca	ggtgaacaac	cccaaggagt	ggctgcaggt	ggatttccag	4080
aagaccatga	aggtgaccgg	cgtgaccacg	cagggcgtga	agagcctgct	gaccagcatg	4140
tacgtgaagg	agttcctgat	cagctctagc	caggacggcc	accagtggac	cctgttcttc	4200
cagaacggca	aggtgaaggt	gttccagggc	aaccaggata	gcttcacccc	cgtggtgaac	4260
agcctggacc	ccccctgct	gaccaggtat	ctgaggatcc	accccagag	ctgggtgcac	4320
cagatcgccc	tgaggatgga	ggtgctgggc	tgcgaggccc	aggatctgta	ttga	4374

<210> 3

<211> 4374

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> coFVIII-52

10

<400> 3

atgcaaatcg	aactgagcac	ctgtttcttc	ctctgcctgc	tgagattctg	tttctccgcg	60
acccgccgat	actacctggg	agcagtggag	ctctcctggg	attacatgca	gagcgacctt	120
ggggagctgc	cctggtatgc	caggttccct	ccccgggtgc	caaagtcgtt	tccgttcaac	180
acctccgtgg	tgtacaagaa	aactctgttc	gtggagttca	ccgaccacct	gttcaatatc	240
gccaagccca	gacctccctg	gatggggctg	ttgggacctt	ccatccaagc	ggaggtgtac	300
gacactgtgg	tcatactctt	gaagaacatg	gcctcgcata	ccgtgtccct	gcacgccgtg	360
ggagtgtctt	actggaaagc	gtccgagggg	gccgaatacg	acgaccagac	ctcgagagaa	420
gaaaagggaag	atgacaaggt	gttcccagga	ggatcgcaca	cctactgtgt	gcaagtgttg	480
aaggagaacg	gccaatggc	ctccgacctg	ctgtgcctga	cctactcgta	cctgtcccac	540
gtggacctcg	tgaaggacct	caactcggga	ctgattggag	ccctgctggt	ctgcagggaa	600
ggctcactgg	cgaaagaaaa	gactcagacc	ttgcacaagt	tcattctgct	gttcgctgtg	660
ttcgacgagg	ggaagtcgtg	gcacagcgag	actaagaact	ccctgatgca	agatagagat	720
gccgcctccg	cccgggcctg	gcctaagatg	cacaccgtga	acggttacgt	gaaccgctcc	780
ctccctggcc	tgattggatg	ccaccggaag	tccgtgtact	ggcacgtgat	cgggatgggg	840
accacccccg	aggtgcacag	catcttcctg	gaaggtcaca	catttctcgt	gcgcaaccac	900
cggcaggcct	ccctggaaat	cagccccatt	accttcctca	ctgccagac	tctgctgatg	960
gacctgggac	agttcctgct	gttctgccat	atctcctccc	accaacatga	cggaatggag	1020
gcatactgta	aggtcgattc	ctgccctgag	gaaccccagc	tccgcatgaa	gaacaatgag	1080
gaagccgagg	actacgacga	cgacctgacg	gatagcgaga	tggtgtggt	ccggttcgat	1140
gacgataaca	gcccttcctt	catccaaatt	cgctcggtgg	caaagaagca	ccccaagacc	1200

# ES 2 926 585 T3

tgggtgcatt	acattgcggc	ggaagaagag	gactgggatt	atgccccgct	tgtcctcgct	1260
cctgacgacc	ggagctacaa	gagccagtac	ctgaacaacg	gtccacagag	gatcggtaga	1320
aagtacaaga	aggtccgctt	catggcctat	accgacgaaa	ccttcaaaac	tagagaggcc	1380
atccaacacg	aatccggcat	cctgggcccc	ctcttgtacg	gagaagtcgg	cgacaccctt	1440
ctcattatct	tcaagaacca	ggcttcccgg	ccgtacaaca	tctatccgca	tgggatcact	1500
gacgtgcgcc	cactgtactc	gcggcgccctg	cccaagggtg	tcaaacacct	gaaggatttt	1560
ccgatccttc	cgggagaaat	cttcaagtac	aagtggaccg	tgaccgtgga	agatggccca	1620
actaagtctg	accctagatg	cctcaccgcg	tactactcat	ccttcgtcaa	catggagcgc	1680
gacctggcca	gcggactgat	cgggcccgctg	ctgatttgct	acaaggaatc	agtggaccaa	1740
cggggaaacc	agatcatgtc	ggataagagg	aacgtcatcc	tcttctccgt	gtttgacgaa	1800
aaccggtcgt	ggtacctgac	cgagaacatc	cagaggttcc	tgcccaaccc	tgctggggtg	1860
cagctggagg	accccagatt	ccaggccagc	aacatcatgc	acagcatcaa	tggctacgtg	1920
ttcgacagcc	tgcagctgag	cgtgtgcctg	cacgagggtg	cctactggta	catcctgagc	1980
atcggcgccc	agaccgactt	cctgagcgtg	ttcttctctg	gctacacctt	caagcacaag	2040
atggtgtatg	aggacaccct	gaccctgttc	cccttcagcg	gggagactgt	cttcatgagc	2100
atggagaacc	ctggcctgtg	gatcctgggc	tgccacaaca	gcgacttcag	gaacaggggc	2160
atgactgccc	tgttgaaagt	ctccagctgt	gacaagaaca	ccggggacta	ctacgaggac	2220
agctacgagg	acatcagcgc	ctacctgtg	agcaagaaca	atgccatcga	gcccaggagc	2280
ttctctcaga	acccccagtt	gctgaagagg	caccagaggg	agatcaccag	gaccaccctg	2340
cagtctgacc	aggaggagat	cgactatgat	gacaccatca	gcgtggagat	gaagaaggag	2400
gacttcgaca	tctacgacga	ggacgagaac	cagagcccca	ggagcttcca	gaagaagacc	2460
aggcactact	tcattgctgc	tgtggagagg	ctgtgggact	atggcatgtc	cagcagcccc	2520
catgtgctga	ggaacagggc	ccagtctggc	agcgtgcccc	agttcaagaa	agtcgtgttc	2580
caggagtcca	ccgacggcag	cttcacccag	cccctgtaca	gaggggagct	gaacgagcac	2640
ctgggcctgc	tgggcccccta	catcagggcc	gaggtggagg	acaacatcat	ggtgaccttc	2700
aggaaccagg	ccagcaggcc	ctacagcttc	tacagcagcc	tgatcagcta	cgaggaggac	2760
cagaggcagg	gggctgagcc	caggaagaac	tttgtgaagc	ccaatgaaac	caagacctac	2820
ttctggaagg	tgcagcacca	catggccccc	accaaggacg	agttcgactg	caaggcctgg	2880
gcctacttct	ctgacgtgga	cctggagaag	gacgtgcact	ctggcctgat	tggccccctg	2940
ctggtgtgcc	acaccaacac	cctgaaccct	gcccatggca	ggcaggtgac	tgtgcaggag	3000
ttcgccctgt	tcttcaccat	cttcgatgaa	accaagagct	ggtacttcac	tgagaacatg	3060

gagaggaact	gcagggccccc	ctgcaacatc	cagatggagg	acccacacct	caaggagAAC	3120
tacaggttcc	atgccatcaa	tggtacatc	atggacaccc	tgcctggcct	ggcatggcc	3180
caggaccaga	ggatcagggtg	gtatctgctg	agcatgggca	gcaacgagaa	catccacagc	3240
atccacttct	ctggccacgt	gttcaactgt	aggaagaagg	aggagtacaa	gatggccctg	3300
tacaacctgt	accctgggggt	gttcgaaacc	gtggagatgc	tgccagcaa	ggccggcatc	3360
tggaggggtg	agtgcctgat	tgaggagcac	ctgcacgccg	gcatgagcac	cctgttcctg	3420
gtgtacagca	acaagtgcc	gacccccctg	ggcatggcct	ctggccacat	cagggaacttc	3480
cagatcactg	cctctggcca	gtacggccag	tgggccccc	agctggccag	gctgcactac	3540
tccggaagca	tcaatgcctg	gagcaccaag	gagcccttca	gctggatcaa	agtggacctg	3600
ctggccccc	tgatcatcca	cggcatcaag	accaggggg	ccaggcagaa	gttctccagc	3660
ctgtacatca	gccagttcat	catcatgtac	agcctggacg	gcaagaagtg	gcagacctac	3720
aggggcaaca	gcaccggcac	cctgatgggtg	ttcttcggca	acgtggacag	cagcggcatc	3780
aagcacaaca	tcttcaaccc	ccccatcatc	gccagatata	tcaggctgca	ccccaccac	3840
tacagcatca	ggagcaccct	gaggatggag	ctgatgggct	gtgacctgaa	cagctgcagc	3900
atgcccctgg	gcatggagag	caaggccatc	tctgacgcc	agatcactgc	ctccagctac	3960
ttaccaaca	tgtttgccac	ctggagcccc	agcaaggcca	ggctgcacct	gcagggcagg	4020
agcaatgcct	ggaggcccca	ggtcaacaac	cccaaggagt	ggctgcagg	ggacttccag	4080
aagaccatga	aggtgactgg	ggtgaccacc	caggggggtga	agagcctgct	gaccagcatg	4140
tacgtgaagg	agttcctgat	ctccagcagc	caggacggcc	accagtggac	cctgttcttc	4200
cagaatggca	aggtgaaggt	gttccagggc	aaccaggaca	gcttcacccc	tgtggtcaac	4260
agcctggacc	ccccctgct	gaccagatac	ctgaggatcc	accccagag	ctgggtgcac	4320
cagatcgccc	tgaggatgga	ggtgctgggc	tgtgaggccc	aggacctgta	ctga	4374

&lt;210&gt; 4

&lt;211&gt; 4374

5 &lt;212&gt; ADN

&lt;213&gt; Secuencia Artificial

&lt;220&gt;

&lt;223&gt; coFVIII-62

10

&lt;400&gt; 4

atgcagattg	agctgtccac	ttgtttcttc	ctgtgcctcc	tgcgcttctg	tttctccgcc	60
actcgcgggt	actaccttgg	agccgtggag	ctttcatggg	actacatgca	gagcgacctg	120
ggcgaaactcc	ccgtggatgc	cagattcccc	cccccggtgc	caaagtcctt	cccctttaac	180
acctccgtgg	tgtacaagaa	aaccctcttt	gtcgagttca	ctgaccacct	gttcaacatc	240
gccaaagccgc	gcccaccttg	gatgggcctc	ctgggaccga	ccattcaagc	tgaagtgtac	300



# ES 2 926 585 T3

gacaccgtgg tgatcaccct gaagaacatg gcgtcccacc ccgtgtccct gcatgcggtc	360
ggagtgtcct actggaaggc ctccgaagga gctgagtacg acgaccagac tagccagcgg	420
gaaaaggagg acgataaagt gttcccgggc ggctcgcata cttacgtgtg gcaagtcctg	480
aaggaaaacg gacctatggc atccgatcct ctgtgcctga cttactccta cctttcccat	540
gtggacctcg tgaaggacct gaacagcggg ctgattggtg cacttctcgt gtgccgcgaa	600
ggttcgctcg ctaaggaaaa gaccagacc ctccataagt tcatcctttt gttcgctgtg	660
ttcgatgaag gaaagtcatg gcattccgaa actaagaact cgctgatgca ggaccgggat	720
gccgcctcag cccgcgcctg gcctaaaatg catacagtca acggatacgt gaatcgggtca	780
ctgcccgggc tcatcggttg tcacagaaag tccgtgtact ggacggtcat cggcatgggc	840
actacgcctg aagtgcactc catcttcctg gaagggcaca ccttcctcgt ggcgaaccac	900
cgccaggcct ctctggaaat ctccccgatt acctttctga ccgccagac tctgctcatg	960
gacctggggc agttccttct cttctgccac atctccagcc atcagcacga cggaatggag	1020
gcctacgtga aggtggactc atgcccggaa gaacctcagt tgcggatgaa gaacaacgag	1080
gaggccgagg actatgacga cgatttgact gactccgaga tggacgtcgt gcggttcgat	1140
gacgacaaca gcccagctt catccagatt cgcagcgtgg ccaagaagca ccccaaaccc	1200
tgggtgcact acatcgcggc cgaggaagaa gattgggact acgccccgtt ggtgctggca	1260
cccgatgacc ggtcgtacaa gtcccagtat ctgaacaatg gtccgcagcg gattggcaga	1320
aagtacaaga aagtgcggtt catggcgtac actgacgaaa cgtttaagac ccgggaggcc	1380
attcaacatg agagcggcat tctgggacca ctgctgtacg gagaggtcgg cgataccctg	1440
ctcatcatct tcaaaaacca ggctccccg ccttacaaca tctaccctca cggaatcacc	1500
gacgtgcggc cactctactc gcggcgcctg ccgaaggcgg tcaagcacct gaaagacttc	1560
cctatcctgc cgggcgaaat cttcaagtat aagtggaccg tcaccgtgga ggacgggccc	1620
accaagagcg atcctaggtg tctgactcgg tactactcca gcttcgtgaa catggaacgg	1680
gacctggcat cgggactcat tggaccgctg ctgatctgct acaaagagtc ggtggatcaa	1740
cgcggaacc agatcatgtc cgacaagcgc aacgtgatcc tgttctccgt gtttgatgaa	1800
aacagatcct ggtacctgac cgagaacatc cagaggttcc tgcccaaccc tgctggggtg	1860
cagctggagg accccagatt ccaggccagc aacatcatgc acagcatcaa tggctacgtg	1920
ttcgacagcc tgcagctgag cgtgtgcctg cacgaggtgg cctactggta catcctgagc	1980
atcggcgccc agaccgactt cctgagcgtg ttcttctctg gctacacctt caagcacaag	2040
atggtgtatg aggacaccct gacctgttc cccttcagcg gggagactgt cttcatgagc	2100
atggagaacc ctggcctgtg gatcctgggc tgccacaaca gcgacttcag gaacaggggc	2160

# ES 2 926 585 T3

atgactgccc	tgctgaaagt	ctccagctgt	gacaagaaca	ccggggacta	ctacgaggac	2220
agctacgagg	acatcagcgc	ctacctgctg	agcaagaaca	atgccatcga	gcccaggagc	2280
ttctctcaga	acccccagct	gctgaagagg	caccagaggg	agatcaccag	gaccaccctg	2340
cagtctgacc	aggaggagat	cgactatgat	gacaccatca	gcgtggagat	gaagaaggag	2400
gacttcgaca	tctacgacga	ggacgagaac	cagagcccca	ggagcttcca	gaagaagacc	2460
aggcactact	tatttgcctg	tgtggagagg	ctgtgggact	atggcatgtc	cagcagcccc	2520
catgtgctga	ggaacagggc	ccagtctggc	agcgtgcccc	agttcaagaa	agtcgtgttc	2580
caggagtcca	ccgacggcag	cttcacccag	cccctgtaca	gaggggagct	gaacgagcac	2640
ctgggcctgc	tgggccccc	catcagggcc	gaggtggagg	acaacatcat	ggtgaccttc	2700
aggaaccagg	ccagcaggcc	ctacagcttc	tacagcagcc	tgatcagcta	cgaggaggac	2760
cagaggcagg	gggctgagcc	caggaagaac	tttgtgaagc	ccaatgaaac	caagacctac	2820
ttctggaagg	tgcagcacca	catggccccc	accaaggacg	agttcgactg	caaggcctgg	2880
gcctacttct	ctgacgtgga	cctggagaag	gacgtgcact	ctggcctgat	tgccccctg	2940
ctggtgtgcc	acaccaacac	cctgaacctt	gcccattggc	ggcagggtgac	tgtgcaggag	3000
ttcgccctgt	tcttcacccat	cttcgatgaa	accaagagct	ggtacttcac	tgagaacatg	3060
gagagggaact	gcaggggccc	ctgcaacatc	cagatggagg	acccacacct	caaggagaac	3120
tacagggttc	atgccatcaa	tggctacatc	atggacaccc	tgccctggcct	ggtcatggcc	3180
caggaccaga	ggatcagggtg	gtatctgctg	agcatgggca	gcaacgagaa	catccacagc	3240
atccacttct	ctggccacgt	gttcactgtg	aggaagaagg	aggagtacaa	gatggccctg	3300
tacaacctgt	accctgggggt	gttcgaaacc	gtggagatgc	tgcccagcaa	ggccggcatc	3360
tggagggtgg	agtgcctgat	tggggagcac	ctgcacgccg	gcatgagcac	cctgttctctg	3420
gtgtacagca	acaagtgcc	gacccccctg	ggcatggcct	ctggccacat	cagggacttc	3480
cagatcactg	cctctggcca	gtacggccag	tgggccccca	agctggccag	gctgcactac	3540
tccggaagca	tcaatgcctg	gagcaccaag	gagcccttca	gctggatcaa	agtggacctg	3600
ctggccccca	tgatcatcca	cggcatcaag	accagggggg	ccaggcagaa	gttctccagc	3660
ctgtacatca	gccagttcat	catcatgtac	agcctggacg	gcaagaagtg	gcagacctac	3720
aggggcaaca	gcaccggcac	cctgatgggtg	ttcttcggca	acgtggacag	cagcggcatc	3780
aagcacaaca	tcttcaaccc	ccccatcatc	gccagataca	tcaggctgca	ccccaccac	3840
tacagcatca	ggagcacccct	gaggatggag	ctgatgggct	gtgacctgaa	cagctgcagc	3900
atgcccctgg	gcatggagag	caaggccatc	tctgacgccc	agatcactgc	ctccagctac	3960
ttaccaaca	tgtttgccac	ctggagcccc	agcaaggcca	ggctgcacct	gcagggcagg	4020
agcaatgcct	ggaggcccc	ggtcaacaac	cccaaggagt	ggctgcagggt	ggacttccag	4080
aagaccatga	aggtgactgg	ggtgaccacc	caggggggtga	agagcctgct	gaccagcatg	4140
tacgtgaagg	agttcctgat	ctccagcagc	caggacggcc	accagtggac	cctgttcttc	4200
cagaatggca	aggtgaagggt	gttccagggc	aaccaggaca	gcttcacccc	tgtggtcaac	4260
agcctggacc	ccccctgct	gaccagatac	ctgaggatcc	accccagag	ctgggtgcac	4320
cagatcgcct	tgaggatgga	ggtgctgggc	tgtgaggccc	aggacctgta	ctga	4374

<210> 5  
 <211> 4374  
 <212> ADN  
 <213> Secuencia Artificial

5

<220>  
 <223> coFVIII-25

<400> 5

10

atgcagattg agctgagcac ctgcttcttc ctgtgcctgc tgagggttctg cttctctgcc	60
accaggagat actacctggg cgccgtggag ctgagctggg actacatgca gtctgacctg	120
ggcgagctgc cagtggacgc caggttcccc cccagagtgc ccaagagctt ccccttcaac	180
accagcgtgg tgtacaagaa gaccctgttc gtggagttca ctgaccacct gttcaacatc	240
gccaagccca ggccccctg gatgggcctg ctgggcccga ccatccaggc cgaggtgtac	300
gacaccgtgg tcatcacctc gaagaacatg gccagccacc ccgtctccct gcacgccgtg	360
ggggtgagct actggaaggc ctctgagggc gccgagtacg acgaccagac cagccagagg	420
gagaaggagg acgacaaggt gttccctggg ggcagccaca cctacgtgtg gcaggtcctg	480
aaggagaacg gccccatggc ctctgacccc ctgtgcctga cctacagcta cctgagccac	540
gtggacctgg tgaaggacct gaactctggc ctgattgggg ccctgctggt gtgcagggag	600
ggcagcctgg ccaaggagaa gaccagacc ctgcacaagt tcatcctgct gttcgccgtg	660
ttcgacgagg gcaagagctg gcaactctgaa accaagaaca gcctgatgca ggacagggac	720
gccgcctctg ccagggcctg gcccaagatg cacaccgtca acggctacgt caacaggagc	780
ctgcctggcc tgattggctg ccacaggaag agcgtgtact ggcattgtgat cggcatgggc	840
accaccctg aggtgcacag catcttcctg gagggccaca ccttcctggt caggaaccac	900
aggcaggcca gcctggagat cagccccatc accttcctga ccgccagac cctgctgatg	960
gacctgggcc agttcctgct gttctgccac atctccagcc accagcacga cggcatggag	1020
gcctacgtga aagtggacag ctgccctgag gagccccagc tgaggatgaa gaacaacgag	1080
gaggccgagg actatgatga cgacctgacc gacagcgaga tggacgtggt caggttcgac	1140
gacgacaaca gcccagctt catccagatc aggagcgtgg ccaagaagca cccaagacc	1200
tgggtgcact acatcgctgc tgaggaggag gactgggact atgccccctt ggtgctggcc	1260

cctgatgaca ggagctacaa gagccagtac ctgaacaatg gccccagag gattggcagg	1320
aagtacaaga aagtcagggt catggcctac actgatgaaa ccttcaagac caggagaggcc	1380
atccagcatg agtctggcat cctgggcccc ctgctgtacg gggagggtgg ggacaccctg	1440
ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctaccccca tggcatcacc	1500
gacgtgaggc ccctgtacag caggaggctg cctaaggggg tgaagcacct gaaagacttc	1560
cccatcctgc ctggggagat cttcaagtac aagtggactg tgactgtgga ggacggcccc	1620
accaagagcg accccagggt cctgaccaga tactacagca gcttcgtcaa catggagagg	1680
gacctggcct ctggcctgat tggccccctg ctgatctgct acaaggagt c tgtggaccag	1740
aggggcaacc agatcatgag cgacaagagg aacgtgatcc tgttctctgt cttcgacgag	1800
aacaggagct ggtacctgac tgaaaacatc cagcggttcc tccccacccc cgcgggcgtg	1860
cagctggaag atcctgagtt tcaggcatca aacatcatgc actccattaa cggctacgtg	1920
ttcgattcgc tgcagctgag cgtgtgtctg cacgaagtgg cctactggta catcctgtcc	1980
attggtgccc agactgactt cctgtccgtg tttttctccg gctacacgtt caagcacaag	2040
atggtgtacg aggacaccct gaccctcttc ccttttccg gcgaaactgt gtttatgagc	2100
atggagaatc ccggcctgtg gatcttgggc tgccacaaca gcgacttccg taacagagga	2160
atgactgcgc tgctcaaggt gtccagctgc gacaagaaca ccggagacta ttatgaggac	2220
tcatacgagg acatctccgc ctacctctg tccaagaata acgccattga acctcggagc	2280
ttcagccaga acccaccctg gcttaagaga catcaacggg agatcactag gaccaccctg	2340
cagtcagacc aggaggaaat cgactacgat gacaccatct cggtcgagat gaagaaggag	2400
gactttgaca tctacgacga agatgaaaac cagagcccga ggtcgttcca aaagaaaacc	2460
cgccactact ttattgtgc tgtcgagcgg ctgtgggact acggaatgtc gtcctcgcgc	2520
cacgtgtccc gcaaccgagc ccagagcggc tcgggtgccgc aattcaagaa ggtcgtgttc	2580
caggagttca ctgacgggag cttcactcag cctttgtacc ggggagaact caatgaacat	2640
ctcggcctcc tcggacctta catcagagca gaagtggaag ataacatcat ggtcactttc	2700
cgtaaccaag ccagccgccc gtactcgttc tactcctccc tcattttctta cgaagaggac	2760
cagcggcagg gcgcagaacc gcgcaagaac ttcgtgaagc ccaacgaaac caagacctac	2820
ttctggaag tgcagcatca tatggccccg actaaggacg agtttgactg caaagcctgg	2880
gcctactttc ccgatgtgga cttggagaag gacgtccact ccggcctcat cggccccctg	2940
ctcgtgtgcc ataccaatac cctgaacccc gcacacggtc gccaggtcac cgtgcaggag	3000
ttcgtctgt tcttactat cttcgacgaa actaagtcct ggtacttcac cgagaacatg	3060
gagaggaact gcagagcccc ctgtaacatc cagatggagg acccgacgtt caaggaaaac	3120
taccggttcc acgccattaa cggatacatc atggatacgc tgccgggtct tgtgatggcc	3180

# ES 2 926 585 T3

caggatcaac ggatcagatg gtacttattg tcgatgggca gcaacgagaa catccactct	3240
attcacttct ccggtcatgt gttcactgtg cggaagaagg aagagtacaa gatggccctg	3300
tacaaccttt atcccgaggt gttcgaaact gtggaaatgc tgccgtcgaa ggccggcatt	3360
tggcgctgtg agtgtttgat tggagaacat ctccatgcgg ggatgtcaac cctgttcctg	3420
gtgtatagca acaagtgcc aactccgctt gggatggcgt caggacacat tagggatttc	3480
cagatcactg cgtccggcca gtacggccaa tgggccccta agctggcccg cctgcattac	3540
tccggatcca ttaacgcctg gtcaaccaag gagccattct cctggatcaa ggtggacctt	3600
ctggccccc tgattatcca cggaattaag acccaggggg ccgggcagaa gttctcctca	3660
ctgtacatca gccagtcat aatcatgtac tccctggacg gaaagaagtg gcaaaccctac	3720
agggggaaca gcaccggcac actgatggtc tttttcggaa atgtggactc ctccgggatt	3780
aagcataaca tcttcaacct tccgattatc gctcgtgaca ttagacttca ccctaccac	3840
tacagcattc gctccaccct gcggatggaa ctgatgggct gcgatctgaa ctctgcagc	3900
atgccgttgg gaatggagtc caaagcaatt tccgacgcgc agatcaccgc ctctcctac	3960
tttaccaaca tgttcgccac gtggtcaccg tccaaggccc ggctgcacct ccaggaaga	4020
tccaacgcat ggcggccaca ggtcaacaac cctaaggagt ggctccaggt ggacttccag	4080
aaaacctga aggtcaccgg agtcacaacc cagggaagtga agtcgctgct gacttctatg	4140
tacgtcaagg agttcctgat ctccagcagc caggacgggc accagtggac cctgttcttc	4200
caaaatggaa aggtcaaggt gtttcagggc aatcaggatt cattcacccc ggtggtgaac	4260
tcccttgatc caccctcct gaccgcctac cttcgcctcc acccacagtc ctgggtgcac	4320
cagatcgccg tgaggatgga ggtcctggga tgcgaagccc aggacctgta ctga	4374

<210> 6

<211> 4374

5 <212> ADN

<213> Secuencia Artificial

<220>

10 <223> coFVIII-26

<400> 6

atgcagattg agctgagcac ctgcttcttc ctgtgcctgc tgaggttctg cttctctgcc	60
accaggagat actacctggg cgccgtggag ctgagctggg actacatgca gtctgacctg	120
ggcgagctgc cagtggacgc caggttcccc ccagagtgcc ccaagagctt ccccttcaac	180
accagcgtgg tgtacaagaa gaccctgttc gtggagttca ctgaccacct gttcaacatc	240
gccaagccca ggccccctg gatgggcctg ctgggcccga ccatccaggc cgaggtgtac	300
gacaccgtgg tcatcaccct gaagaacatg gccagccacc ccgtctccct gcacgccgtg	360

ggggtgagct actggaaggc ctctgagggc gccgagtacg acgaccagac cagccagagg	420
gagaaggagg acgacaaggt gttccctggg ggcagccaca cctacgtgtg gcaggtcctg	480
aaggagaacg gccccatggc ctctgacccc ctgtgcctga cctacagcta cctgagccac	540
gtggacctgg tgaaggacct gaactctggc ctgattgggg ccctgctggt gtgcaggagg	600
ggcagcctgg ccaaggagaa gaccagacc ctgcacaagt tcatcctgct gttcgccgtg	660
ttcgacgagg gcaagagctg gcactctgaa accaagaaca gcctgatgca ggacagggac	720
gccgcctctg ccagggcctg gcccgaagt caccacctca acggctacgt caacaggagc	780
ctgcctggcc tgattggctg ccacaggaag agcgtgtact ggcattgtgat cggcatgggc	840
accacccctg aggtgcacag catcttctctg gagggccaca ccttctctgt caggaaccac	900
aggcaggcca gcctggagat cagccccatc accttctctga ccgcccagac cctgctgatg	960
gacctgggcc agttctctgt gttctgccac atctccagcc accagcacga cggcatggag	1020
gcctacgtga aagtggacag ctgccctgag gagccccagc tgaggatgaa gaacaacgag	1080
gaggccgagg actatgatga cgacctgacc gacagcgaga tggacgtggt cagggttcgac	1140
gacgacaaca gccccagctt catccagatc aggagcgtgg ccaagaagca cccaagacc	1200
tgggtgcact acatcgctgc tgaggaggag gactgggact atgccccctt ggtgctggcc	1260
cctgatgaca ggagctacaa gagccagtac ctgaacaatg gccccagag gattggcagg	1320
aagtacaaga aagtcagggt catggcctac actgatgaaa ccttcaagac cagggaggcc	1380
atccagcatg agtctggcat cctgggcccc ctgctgtacg gggagggtgg ggacaccctg	1440
ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctaccccc tggcatcacc	1500
gacgtgaggc ccctgtacag caggaggctg cctaaggggg tgaagcacct gaaagacttc	1560
cccatcctgc ctggggagat cttcaagtac aagtggactg tgactgtgga ggacggcccc	1620
accaagagcg accccagggt cctgaccaga tactacagca gcttcgtcaa catggagagg	1680
gacctggcct ctggcctgat tggccccctg ctgatctgct acaaggagtc tgtggaccag	1740
aggggcaacc agatcatgag cgacaagagg aacgtgatcc tgttctctgt cttcgacgag	1800
aacaggagct ggtacctcac tgaaaacatc cagaggttcc tcccaaacc cgcaggagtg	1860
caactggagg accctgagtt tcaggcctcg aatatcatgc actcgattaa cgtttacgtg	1920
ttcgactcgc tgcagctgag cgtgtgcctc catgaagtgc cttactggta cattctgtcc	1980
atcggcgccc agactgactt cctgagcgtg ttcttttccg gttacacctt taagcacaag	2040
atggtgtacg aagataccct gacctgttc cttttctccg gcgaaacggt gttcatgtcg	2100
atggagaacc cgggtctgtg gattctggga tgccacaaca gcgactttcg gaaccgcgga	2160
atgactgccc tgctgaaggt gtcctcatgc gacaagaaca ccggagacta ctacgaggac	2220
tcctacgagg atatctcagc ctacctctg tccaagaaca acgcatcga gccgcgcagc	2280

# ES 2 926 585 T3

ttcagccaga acccgctgt gctgaagagg caccagcgag aaattacccg gaccaccctc	2340
caatcggatc aggaggaaat cgactacgac gacaccatct cggtggaat gaagaaggaa	2400
gatttcgata tctacgacga ggacgaaaat cagtcccctc gctcattcca aaagaaaact	2460
agacactact ttatcgccgc ggtggaaga ctgtgggact atggaatgtc atccagccct	2520
cacgtccttc ggaaccgggc ccagagcgga tcggtgcctc agttcaagaa agtgggtgttc	2580
caggagtcca ccgacggcag cttcaccag ccgctgtacc ggggagaact gaacgaacac	2640
ctgggcctgc tcggtcccta catccgcgcg gaagtggagg ataacatcat ggtgaccttc	2700
cgtaaccaag catccagacc ttactccttc tattcctccc tgatctcata cgaggaggac	2760
cagcgccaag gcgcgagcc ccgcaagaac ttcgtcaagc ccaacgagac taagacctac	2820
ttctggaagg tccaacacca tatggccccg accaaggatg agtttgactg caaggcctgg	2880
gcctacttct ccgacgtgga ccttgagaag gatgtccatt ccggcctgat cgggccgctg	2940
ctcgtgtgtc acaccaacac cctgaacca gcgcatggac gccaggtcac cgtccaggag	3000
tttgctctgt tcttcacat ttttgacgaa actaagtcct ggtacttcac cgagaatatg	3060
gagcgaaaact gtagagcgcc ctgcaatatc cagatggaag atccgacttt caaggagaac	3120
tatagattcc acgccatcaa cgggtacatc atggatactc tgccggggct ggtcatggcc	3180
caggatcaga ggattcgggt gtacttgctg tcaatgggat cgaacgaaaa cattcactcc	3240
attcacttct ccggtcacgt gttcactgtg cgcaagaagg aggagtacaa gatggcgtg	3300
tacaatctgt accccggggt gttcgaaact gtggagatgc tgccgtccaa ggccggcatc	3360
tggagagtgg agtgcctgat cggagagcac ctccacgcgg ggatgtccac cctcttcctg	3420
gtgtactcga ataagtgccg gaccccgctg ggcattggcct cggggccacat cagagacttc	3480
cagatcacag caagcggaca atacggccaa tgggcgccga agctggcccc cttgcactac	3540
tccgcatcga tcaacgcatt gtccaccaag gaaccgttct cgtggattaa ggtggacctc	3600
ctggccctta tgattatcca cggaattaag acccagggcg ccaggcagaa gttctcctcc	3660
ctgtacatct cgcaattcat catcatgtac agcctggacg ggaagaagtg gcagacttac	3720
aggggaaaact ccaccggcac cctgatggtc tttttcggca acgtggattc ctccggcatt	3780
aagcacaaca tcttcaacct accgatcata gccagatata ttaggctcca cccactcac	3840
tactcaatcc gctcaactct tcggatggaa ctcatggggt gcgacctgaa ctctgctcc	3900
atgccgttgg ggatggaatc aaaggctatt agcgacgccc agatcaccgc gagctcctac	3960
ttcactaaca tggtcgccac ctggagcccc tccaaggcca ggctgcactt gcagggacgg	4020
tcaaatgcct ggcggccgca agtgaacaat ccgaaggaaat ggcttcaagt ggatttccaa	4080
aagaccatga aagtgaccgg agtcaccacc caggagtgga agtcccttct gacctcgatg	4140
tatgtgaagg agttcctgat tagcagcagc caggacgggc accagtggac cctgttcttc	4200
caaaacggaa aggtcaagggt gttccagggg aaccaggact cgttcacacc cgtgggtgaa	4260
tccctggacc cccactgct gacgcgttac ttgaggattc atcctcagtc ctgggtccat	4320
cagattgcat tgcgaatgga agtcctgggc tgcgagggcc aggacctgta ctga	4374

<212> ADN  
<213> Secuencia Artificial

<220>  
<223> coFVIII-52-NT58

<400> 7

```

gcgacccgcc gatactacct gggagcagtg gagctctcct gggattacat gcagagcgac      60
cttgggggagc tgcccggtga tgccaggttc cctccccggg tgccaaagtc gtttccgttc      120
aacacctccg tgggtgtacaa gaaaactctg ttcgtggagt tcaccgacca cctgttcaat      180
atcgccaagc ccagacctcc ctggatgggg ctgttggggac ctaccatcca agcggaggtg      240
tacgacactg tggatcatcac tctgaagaac atggcctcgc atcccggtgc cctgcacgcc      300
gtgggagtggt cttactggaa agcgtccgag ggggccgaat acgacgacca gacctcgag      360
agagaaaagg aagatgacaa ggtgttccca ggaggatcgc acacctacgt gtggcaagtg      420
ttgaaggaga acggcccaat ggctccgac ccgctgtgcc tgacctactc gtacctgtcc      480
cacgtggacc tcgtgaagga cctcaactcg ggactgattg gagccctgct ggtctgcagg      540
gaaggctcac tggcgaaaga aaagactcag accttgacac agttcattct gctgttcgct      600
gtgttcgacg aggggaagtc gtggcacagc gagactaaga actccctgat gcaagataga      660
gatgcgcgct ccgcccgggc ctggcctaag atgcacaccg tgaacggtta cgtgaaccgc      720
tccctccctg gcctgattgg atgccaccgg aagtcctgtg actggcacgt gatcgggatg      780
gggaccaccc ccgaggtgca cagcatcttc ctggaaggtc acacatttct cgtgcgcaac      840
caccggcagg cctccctgga aatcagcccc attaccttcc tcaactgcca gactctgctg      900
atggacctgg gacagttcct gctgttctgc catatctcct cccaccaaca tgacggaatg      960
gaggcatacg tgaaggtcga ttccctgcct gaggaacccc agctccgcat gaagaacaat     1020
gaggaagccg aggactacga cgacgacctg acggatagcg agatggatgt ggtccggttc     1080
gatgacgata acagcccttc cttcatccaa attcgtcctg tggcaaagaa gcacccaag      1140
acctgggtgc attacattgc ggcggaagaa gaggactggg attatgcccc gcttgtcctc     1200
gctcctgacg accggagcta caagagccag tacctgaaca acggtccaca gaggatcggt     1260
agaaagtaca agaaggtccg cttcatggcc tataccgacg aaaccttcaa aactagagag     1320
gccatccaac acgaatccgg catcctgggc ccgctcttgt acggagaagt cggcgacacc     1380
cttctcatta tcttcaagaa ccaggcttcc cggcgtaca acatctatcc gcatgggatc     1440
actgacgtgc gcccaactgta ctgcggcgcc ctgcccaagg gtgtcaaaca cctgaaggat     1500
tttccgatcc ttccgggaga aatcttcaag tacaagtgga ccgtgaccgt ggaagatggc     1560
ccaactaagt ctgacctag atgcctcacc cgctactact catccttcgt caacatggag     1620
cgcgacctgg ccagcggact gatcggcccg ctgctgattt gctacaagga atcagtggac     1680
caacggggaa accagatcat gtcggataag aggaacgtca tcctcttctc cgtg          1734

```

<210> 8  
<211> 2583  
<212> ADN  
<213> Secuencia Artificial



&lt;220&gt;

&lt;223&gt; coFVIII-52-CT

5 &lt;400&gt; 8

tttgacgaaa accggtcgtg gtacctgacc gagaacatcc agaggttcct gccaaccct	60
gctggggtgc agctggagga ccccgagttc caggccagca acatcatgca cagcatcaat	120
ggctacgtgt tcgacagcct gcagctgagc gtgtgcctgc acgaggtggc ctactggtac	180
atcctgagca tcggcgccca gaccgacttc ctgagcgtgt tcttctctgg ctacaccttc	240
aagcacaaga tgggtgatga ggacaccctg accctgttcc ccttcagcgg ggagactgtc	300
ttcatgagca tggagaacct tggcctgtgg atcctgggct gccacaacag cgacttcagg	360
aacaggggca tgactgccct gctgaaagtc tccagctgtg acaagaacac cggggactac	420
tacgaggaca gctacgagga catcagcgcc tacctgctga gcaagaacaa tgccatcgag	480
cccaggagct tctctcagaa cccccagtg ctgaagaggc accagaggga gatcaccagg	540
accacctgc agtctgacca ggaggagatc gactatgatg acaccatcag cgtggagatg	600
aagaaggagg acttcgacat ctacgacgag gacgagaacc agagccccag gagcttccag	660
aagaagacca ggcactactt cattgctgct gtggagaggc tgtgggacta tggcatgtcc	720
agcagcccc atgtgctgag gaacagggcc cagtctggca gcgtgcccc gttcaagaaa	780
gtcgtgttcc aggagtccac cgacggcagc ttcacccagc ccctgtacag aggggagctg	840
aacgagcacc tgggcctgct gggccccctac atcagggccg aggtggagga caacatcatg	900
gtgaccttca ggaaccaggc cagcaggccc tacagcttct acagcagcct gatcagctac	960
gaggaggacc agaggcaggg ggctgagccc aggaagaact ttgtgaagcc caatgaaacc	1020
aagacctact tctggaaggt gcagcaccac atggccccca ccaaggacga gttcgactgc	1080
aaggcctggg cctacttctc tgacgtggac ctggagaagg acgtgcactc tggcctgatt	1140
ggccccctgc tgggtgtgcc caccaacacc ctgaaccctg cccatggcag gcaggtgact	1200

# ES 2 926 585 T3

```

gtgcaggagt tcgccctgtt cttcaccatc ttcgatgaaa ccaagagctg gtacttcact 1260
gagaacatgg agaggaactg cagggccccc tgcaacatcc agatggagga cccacacctc 1320
aaggagaact acaggttcca tgccatcaat ggctacatca tggacaccct gcctggcctg 1380
gtcatggccc aggaccagag gatcaggtgg tatctgctga gcatgggcag caacgagaac 1440
atccacagca tccacttctc tggccacgtg ttcactgtga ggaagaagga ggagtacaag 1500
atggccctgt acaacctgta ccctgggggtg ttcgaaaccg tggagatgct gccagcaag 1560
gccggcatct ggagggtgga gtgcctgatt ggggagcacc tgcacgccgg catgagcacc 1620
ctgttcctgg tgtacagcaa caagtgccag acccccctgg gcatggcctc tggccacatc 1680
agggacttcc agatcactgc ctctggccag tacggccagt gggcccccac gctggccagg 1740
ctgcactact ccggaagcat caatgcctgg agcaccaagg agcccttcag ctggatcaaa 1800
gtggacctgc tggcccccac gatcatccac ggcatacaaga cccagggggc caggcagaag 1860
ttctccagcc tgtacatcag ccagttcatc atcatgtaca gcctggacgg caagaagtgg 1920
cagacctaca ggggcaacag caccggcacc ctgatgggtg tcttcggcaa cgtggacagc 1980
agcggcatca agcacaacat cttcaacccc ccatcatcgc ccagatacat caggctgcac 2040
cccaccact acagcatcag gagcaccctg aggatggagc tgatgggctg tgacctgaac 2100
agctgcagca tggccctggg catggagagc aaggccatct ctgacgccca gatcactgcc 2160
tccagctact tcaccaacat gtttgccacc tggagcccca gcaaggccag gctgcacctg 2220
cagggcagga gcaatgcctg gagggcccag gtcaacaacc ccaaggagtg gctgcagggtg 2280
gacttccaga agaccatgaa ggtgactggg gtgaccaccc aggggggtgaa gagcctgctg 2340
accagcatgt acgtgaagga gttcctgatc tccagcagcc aggacggcca ccagtggacc 2400
ctgttcttcc agaatggcaa ggtgaagggtg ttccagggca accaggacag cttcacccct 2460
gtggtcaaca gcctggaccc ccccctgctg accagatacc tgaggatcca ccccagagc 2520
tgggtgcacc agatgcacct gaggatggag gtgtgggct gtgaggcca ggacctgtac 2580
tga 2583

```

<210> 9  
 <211> 1734  
 <212> ADN  
 <213> Secuencia Artificial

<220>  
 <223> coFVIII-62-NT58

<400> 9

```

gccactcgcc ggtactacct tggagccgtg gagctttcat gggactacat gcagagcgac 60
ctgggcgaac tccccgtgga tgccagattc ccccccgcg tgccaaagtc cttccccctt 120
aacacctccg tgggtgtacaa gaaaaccctc tttgtcgagt tcactgacca cctgttcaac 180

```

atcgccaagc cgcgcccacc ttggatgggc ctccctgggac cgaccattca agctgaagtg 240  
 tacgacaccg tggatgatcac cctgaagaac atggcggtccc acccgtgtc cctgcatgcg 300  
 gtcggagtgt cctactggaa ggccctccgaa ggagctgagt acgacgacca gactagccag 360  
 cgggaaaagg aggacgataa agtgttcccg ggcggtcgc atacttacgt gtggcaagtc 420  
 ctgaaggaaa acggacctat ggcatccgat cctctgtgcc tgacttactc ctacctttcc 480  
 catgtggacc tcgtgaagga cctgaacagc gggctgattg gtgcacttct cgtgtgccgc 540  
 gaaggttcgc tcgctaagga aaagaccag accctccata agttcatcct tttgttcgct 600  
 gtgttcgatg aaggaaagtc atggcattcc gaaactaaga actcgctgat gcaggaccgg 660  
 gatgccgcct cagcccgcg ctggcctaaa atgcatacag tcaacggata cgtgaatcgg 720  
 tcactgcccg ggctcatcgg ttgtcacaga aagtccgtgt actggcacgt catcgcatg 780  
 ggcactacgc ctgaagtga ctccatcttc ctggaagggc acaccttcct cgtgcgcaac 840  
 caccgccagg cctctctgga aatctccccg attacctttc tgaccgcca gactctgtc 900  
 atggacctgg ggcagttcct tctcttctgc cacatctcca gccatcagca cgacggaatg 960  
 gaggcctacg tgaaggtgga ctcatgccg gaagaacctc agttgcggat gaagaacaac 1020  
 gaggagggcg aggactatga cgacgatttg actgactccg agatggacgt cgtgcggttc 1080  
 gatgacgaca acagccccag ctcatccag attcgacg cgccaagaa gcaccccaaa 1140  
 acctgggtgc actacatcgc ggccgaggaa gaagattggg actacgccc gttggtgctg 1200  
 gcacccgatg accggtcgta caagtcccag tatctgaaca atggtccgca gcggttggc 1260  
 agaaagtaca agaaagtgcg gttcatggcg tacactgacg aaacgtttaa gaccgggag 1320  
 gccattcaac atgagagcgg cattctggga ccactgctgt acggagaggt cggcgatacc 1380  
 ctgctcatca tcttcaaaaa ccaggcctcc cggccttaca acatctacc tcacggaatc 1440  
 accgacgtgc ggccactcta ctcgcgcgcg ctgccgaagg gcgtcaagca cctgaaagac 1500  
 ttccctatcc tgccggggcg aatcttcaag tataagtggg ccgtcaccgt ggaggacggg 1560  
 cccaccaaga gcgatcctag gtgtctgact cggtaactact ccagcttcgt gaacatggaa 1620  
 cgggacctgg catcgggact cattggaccg ctgctgatct gctacaaaga gtcggtggat 1680  
 caacgcggca accagatcat gtccgacaag cgcaacgtga tcctgttctc cgtg 1734

<210> 10  
 <211> 2583  
 <212> ADN  
 <213> Secuencia Artificial

<220>  
 <223> coFVIII-62-CT

<400> 10

# ES 2 926 585 T3

tttgatgaaa acagatcctg gtacctgacc gagaacatcc agagggttcct gcccaccct	60
gctggggtgc agctggagga ccccgagttc caggccagca acatcatgca cagcatcaat	120
ggctacgtgt tcgacagcct gcagctgagc gtgtgcctgc acgagggtggc ctactggtac	180
atcctgagca tcggcgccca gaccgacttc ctgagcgtgt tcttctctgg ctacaccttc	240
aagcacaaga tgggtgtatga ggacaccctg accctgttcc ccttcagcgg ggagactgtc	300
ttcatgagca tggagaacct tggcctgtgg atcctgggct gccacaacag cgacttcagg	360
aacaggggca tgactgcctt gctgaaagtc tccagctgtg acaagaacac cggggactac	420
tacgaggaca gctacgagga catcagcgcc tacctgctga gcaagaacaa tgccatcgag	480
cccaggagct tctctcagaa ccccccagtg ctgaagaggc accagaggga gatcaccagg	540
accaccctgc agtctgacca ggaggagatc gactatgatg acaccatcag cgtggagatg	600
aagaaggagg acttcgacat ctacgacgag gacgagaacc agagccccag gagcttcag	660
aagaagacca ggcactactt cattgctgct gtggagaggc tgtgggacta tggcatgtcc	720
agcagcccc atgtgtgag gaacagggcc cagtctggca gcgtgcccc gttcaagaaa	780
gtcgtgttcc aggagttcac cgacggcagc ttcaccagc ccctgtacag aggggagctg	840
aacgagcacc tgggcctgct gggcccctac atcagggccg aggtggagga caacatcatg	900
gtgaccttca ggaaccaggc cagcaggccc tacagcttct acagcagcct gatcagctac	960
gaggaggacc agaggcaggg ggctgagccc aggaagaact ttgtgaagcc caatgaaacc	1020
aagacctact tctggaaggt gcagcaccac atggccccca ccaaggacga gttcgactgc	1080
aaggcctggg cctacttctc tgacgtggac ctggagaagg acgtgcactc tggcctgatt	1140
ggccccctgc tgggtgtgca caccaacacc ctgaaccctg cccatggcag gcaggtgact	1200
gtgcaggagt tcgccctgtt cttcaccatc ttcatgaaa ccaagagctg gtacttcact	1260
gagaacatgg agaggaactg cagggcccc tgcaacatcc agatggagga cccacacctc	1320
aaggagaact acaggttcca tgccatcaat ggctacatca tggacaccct gcctggcctg	1380
gtcatggccc aggaccagag gatcagggtg tatctgctga gcatgggcag caacgagaac	1440
atccacagca tccacttctc tggccacgtg ttactgtga ggaagaagga ggagtacaag	1500
atggccctgt acaacctgta ccctggggtg ttcgaaaccg tggagatgct gccagcaag	1560
gccggcatct ggagggtgga gtgcctgatt ggggagcacc tgcacgccgg catgagcacc	1620
ctgttcctgg tgtacagcaa caagtgccag acccccctgg gcatggcctc tggccacatc	1680
agggacttcc agatcactgc ctctggccag tacggccagt gggcccccaa gctggccagg	1740
ctgcactact ccggaagcat caatgcctgg agcaccaagg agcccttcag ctggatcaaa	1800
gtggacctgc tggcccccat gatcatecac ggcataaga cccagggggc caggcagaag	1860
ttctccagcc tgtacatcag ccagttcatc atcatgtaca gcctggacgg caagaagtgg	1920

# ES 2 926 585 T3

cagacctaca	ggggcaacag	caccggcacc	ctgatggtgt	tcttcggcaa	cgtggacagc	1980
agcggcatca	agcacaacat	cttcaacccc	cccatcatcg	ccagatacat	caggctgcac	2040
cccacccact	acagcatcag	gagcaccctg	aggatggagc	tgatgggctg	tgacctgaac	2100
agctgcagca	tgcccctggg	catggagagc	aaggccatct	ctgacgcca	gatcactgcc	2160
tccagctact	tcaccaacat	gtttgccacc	tgagagccca	gcaaggccag	gctgcacctg	2220
cagggcagga	gcaatgcctg	gaggccccag	gtcaacaacc	ccaaggagtg	gctgcagggtg	2280
gacttcaga	agaccatgaa	ggtgactggg	gtgaccaccc	agggggtgaa	gagcctgctg	2340
accagcatgt	acgtgaagga	gttcctgata	tccagcagcc	aggacggcca	ccagtggacc	2400
ctgttcttcc	agaatggcaa	ggtgaagggtg	ttccagggca	accaggacag	cttcaccctt	2460
gtggtcaaca	gcctggaccc	ccccctgctg	accagatacc	tgaggatcca	ccccagagc	2520
tgggtgcacc	agatgcacct	gaggatggag	gtgtgggct	gtgaggccca	ggacctgtac	2580
tga						2583

<210> 11  
 <211> 1734  
 <212> ADN  
 <213> Secuencia Artificial

<220>  
 <223> coFVIII-25-NT58

<400> 11

gccaccagga	gatactacct	gggcgccgtg	gagctgagct	gggactacat	gcagtctgac	60
ctgggcgagc	tgccagtgga	cgccaggttc	ccccccagag	tgcccaagag	cttccccctt	120
aacaccagcg	tgggtgataa	gaagaccctg	ttcgtggagt	tactgacca	cctgttcaac	180
atcgccaagc	ccaggcccc	ctggatgggc	ctgtggggcc	ccaccatcca	ggccgagggtg	240
tacgacaccg	tggatcatcac	cctgaagaac	atggccagcc	accccgcttc	cctgcacgcc	300
gtgggggtga	gctactggaa	ggcctctgag	ggcgccgagt	acgacgacca	gaccagccag	360
agggagaagg	aggacgacaa	ggtgttcctt	gggggcagcc	acacctacgt	gtggcagggtc	420
ctgaaggaga	acggccccat	ggcctctgac	cccctgtgcc	tgacctacag	ctacctgagc	480
cacgtggacc	tggatgaagga	cctgaactct	ggcctgattg	gggccctgct	ggtgtgcagg	540
gagggcagcc	tgcccaagga	gaagaccag	accctgcaca	agttcatcct	gctgttcgcc	600
gtgttcgacg	agggaagag	ctggcactct	gaaaccaaga	acagcctgat	gcaggacagg	660
gacgccgcct	ctgccagggc	ctggcccaag	atgcacaccg	tcaacggcta	cgtcaacagg	720
agcctgcctg	gcctgattgg	ctgccacagg	aagagcgtgt	actggcatgt	gatcggcatg	780
ggcaccaccc	ctgagggtga	cagcatcttc	ctggagggcc	acaccttcct	ggtcaggaac	840

cacaggcagg	ccagcctgga	gatcagcccc	atcaccttcc	tgaccgcca	gaccctgctg	900
atggacctgg	gccagttcct	gctgttctgc	cacatctcca	gccaccagca	cgacggcatg	960
gaggcctacg	tgaagtga	cagctgccct	gaggagcccc	agctgaggat	gaagaacaac	1020
gaggaggccg	aggactatga	tgacgacctg	accgacagcg	agatggacgt	ggtcaggttc	1080
gacgacgaca	acagccccag	cttcatccag	atcaggagcg	tggccaagaa	gcacccaag	1140
acctgggtgc	actacatcgc	tgctgaggag	gaggactggg	actatgcccc	cctggtgctg	1200
gccccgatg	acaggagcta	caagagccag	tacctgaaca	atggcccca	gaggattggc	1260
aggaagtaca	agaaagtcag	gttcatggcc	tacctgatg	aaaccttcaa	gaccagggag	1320
gccatccagc	atgagtctgg	catcctgggc	ccccgtctgt	acggggaggt	gggggacacc	1380
ctgctgatca	tcttcaagaa	ccaggccagc	aggccctaca	acatctaccc	ccatggcatc	1440
accgacgtga	ggccccgtga	cagcaggagg	ctgcctaagg	gggtgaagca	cctgaaagac	1500
ttccccatcc	tgccctggga	gatcttcaag	tacaagtga	ctgtgactgt	ggaggacggc	1560
cccaccaaga	gcgacccag	gtgcctgacc	agatactaca	gcagcttcgt	caacatggag	1620
agggacctgg	cctctggcct	gattggcccc	ctgctgatct	gctacaagga	gtctgtggac	1680
cagaggggca	accagatcat	gagcgacaag	aggaacgtga	tcctgttctc	tgtc	1734

&lt;210&gt; 12

5 &lt;211&gt; 2583

&lt;212&gt; ADN

&lt;213&gt; Secuencia Artificial

&lt;220&gt;

10 &lt;223&gt; coFVIII-25-CT

&lt;400&gt; 12

ttcgacgaga	acaggagctg	gtacctgact	gaaaacatcc	agcggttcct	ccccaacccc	60
gcgggctg	agctggaaga	tcctgagttt	caggcatcaa	acatcatgca	ctccattaac	120
ggctacgtgt	tcgattcgct	gcagctgagc	gtgtgtctgc	acgaagtggc	ctactggtac	180
atcctgtcca	ttgggtgcca	gactgacttc	ctgtccgtgt	ttttctccgg	ctacacgttc	240
aagcacaaga	tgggtgtacga	ggacaccctg	accctcttcc	ctttttccgg	cgaaactgtg	300
tttatgagca	tggagaatcc	ggcctgtgg	atcttgggct	gccacaacag	cgacttccgt	360
aacagaggaa	tgactgctg	gctcaagggt	tccagctg	acaagaacac	cggagactat	420
tatgaggact	catacgagga	catctccgcc	tacctcctgt	ccaagaataa	cgccattgaa	480
cctcggagct	tcagccagaa	cccaccgtg	cttaagagac	atcaacggga	gatcactagg	540
accaccctgc	agtcagacca	ggaggaaatc	gactacgatg	acaccatctc	ggtcgagatg	600
aagaaggagg	actttgacat	ctacgacgaa	gatgaaaacc	agagcccag	gtcgttccaa	660
aagaaaaccc	gccactactt	tattgctgct	gtcgagcggc	tgtgggacta	cggaatgtcg	720

15

# ES 2 926 585 T3

tcctcgccgc acgtgctccg caaccgagcc cagagcggct cggtgccgca attcaagaag	780
gtcgtgttcc aggagttcac tgacgggagc ttactcagc ctttgtaccg gggagaactc	840
aatgaacatc tcggcctcct cggaccttac atcagagcag aagtgggaaga taacatcatg	900
gtcactttcc gtaaccaagc cagccgcccg tactcgttct actcctccct catttcttac	960
gaagaggacc agcggcaggc cgcagaaccg cgcaagaact tcgtgaagcc caacgaaacc	1020
aagacctact tctggaaagt gcagcatcat atggccccga ctaaggacga gtttgactgc	1080
aaagcctggg cctactttct cgatgtggac ttggagaagg acgtccactc cggcctcatc	1140
ggtccccctg tcgtgtgcc aaccaatacc ctgaaccccg cacacggctc ccaggtcacc	1200
gtgcaggagt tcgctctgtt cttcactatc ttcgacgaaa ctaagtcttg gtacttcacc	1260
gagaacatgg agaggaactg cagagccccc tgtaacatcc agatggagga cccgacgttc	1320
aaggaaaact accggttcca cgccattaac ggatacatca tggatacgct gccgggtctt	1380
gtgatggccc aggatcaacg gatcagatgg tactttattgt cgatgggcag caacgagaac	1440
atccactcta ttactttctc cggtcattgt ttactgtgc ggaagaagga agagtacaag	1500
atggccctgt acaaccttta tcccggagtg ttcgaaactg tggaaatgct gccgtcgaag	1560
gccggcattt ggcgcgtgga gtgtttgatt ggagaacatc tccatgcggg gatgtcaacc	1620
ctgttcctgg tgtatagcaa caagtgccag actccgcttg ggatggcgtc aggacacatt	1680
agggatttcc agatcactgc gtccggccag tacggccaat gggcccctaa gctggcccgc	1740
ctgcattact ccggatccat taacgcctgg tcaaccaagg agccattctc ctggatcaag	1800
gtggaccttc tggcccccatt gattatccac ggaattaaga cccagggggc ccggcagaag	1860
ttctcctcac tgtacatcag ccagttcata atcatgtact ccctggacgg aaagaagtgg	1920
caaacctaca gggggaacag caccggcaca ctgatgtct ttttcggaaa tgtggactcc	1980
tccgggatta agcataacat cttcaaccct ccgattatcg ctcggtacat tagacttcac	2040
cctaccctac acagcattcg ctccacctg cggatggaac tgatgggctg cgatctgaac	2100
tcgtgcagca tgccgttggg aatggagtcc aaagcaattt ccgacgcgca gatcaccgcc	2160
tcgtcctact ttaccaacat gttcgccacg tggtcaccgt ccaaggcccc gctgcacctc	2220
cagggaagat ccaacgcatt gcggccacag gtcaacaacc ctaaggagtg gctccagggtg	2280
gacttccaga aaacctgaa ggtcaccgga gtcacaaccc agggagtga gtcgctgctg	2340
acttctatgt acgtcaagga gttcctgatc tccagcagcc aggacgggca ccagtggacc	2400
ctgttcttcc aaaatggaaa ggtcaagggtg tttcagggca atcaggattc attcaccctg	2460
gtggtgaact cccttgatcc acccctcctg acccgtacc ttcgcatcca cccacagtcc	2520
tgggtgcacc agatcgcgct gaggatggag gtcctgggat gcgaagccca ggacctgtac	2580
tga	2583

- 5 <210> 13  
 <211> 1734  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 10 <223> coFVIII-26-NT58

&lt;400&gt; 13

```

gccaccagga gatactacct gggcgccgtg gagctgagct gggactacat gcagtctgac      60
ctgggcgagc tgccagtgga cgccaggttc cccccagag tgcccaagag cttccccttc      120
aacaccagcg tgggtgtacaa gaagaccctg ttcgtggagt tcaactgacca cctgttcaac      180
atcgccaagc ccaggccccc ctggatgggc ctgctggggc ccaccatcca ggccgaggtg      240
tacgacaccg tggtcatcac cctgaagaac atggccagcc acccgtctc cctgcacgcc      300
gtgggggtga gctactggaa ggcctctgag ggcgccgagt acgacgacca gaccagccag      360
agggagaagg aggacgacaa ggtgttccct gggggcagcc acacctacgt gtggcaggtc      420
ctgaaggaga acggccccat ggcctctgac cccctgtgcc tgacctacag ctacctgagc      480
cacgtggacc tgggtgaagga cctgaactct ggcctgattg gggccctgct ggtgtgcagg      540
gagggcagcc tggccaagga gaagaccag accctgcaca agttcatcct gctgttcgcc      600
gtgttcgacg agggcaagag ctggcactct gaaaccaaga acagcctgat gcaggacagg      660
gacgccgcct ctgccagggc ctggcccaag atgcacaccg tcaacggcta cgtcaacagg      720
agcctgcctg gcctgatttg ctgccacagg aagagcgtgt actggcatgt gatcggcatg      780
ggcaccaccc ctgaggtgca cagcatcttc ctggagggcc acaccttcct ggtcaggaac      840
cacaggcagg ccagcctgga gatcagcccc atcaccttcc tgaccgcca gaccctgctg      900
atggacctgg gccagttcct gctgttctgc cacatctcca gccaccagca cgacggcatg      960
gaggcctacg tgaaagtgga cagctgccct gaggagcccc agctgaggat gaagaacaac     1020
gaggaggccg aggactatga tgacgacctg accgacagcg agatggacgt ggtcaggttc     1080
gacgacgaca acagccccag cttcatccag atcaggagcg tggccaagaa gcacccaag     1140
acctgggtgc actacatcgc tgctgaggag gaggactggg actatgcccc cctggtgctg     1200
gcccctgatg acaggagcta caagagccag tacctgaaca atggcccca gaggattggc     1260
aggaagtaca agaaagttag gttcatggcc tacactgatg aaaccttcaa gaccaggag     1320
gccatccagc atgagtctgg catcctgggc cccctgctgt acggggaggt gggggacacc     1380
ctgctgatca tcttcaagaa ccaggccagc aggcctaca acatctacc ccatggcatc     1440
accgacgtga gggccctgta cagcaggagg ctgcctaagg ggggtgaagca cctgaaagac     1500
ttcccatcc tgccctggga gatcttcaag tacaagtgga ctgtgactgt ggaggacggc     1560
cccaccaaga gcgaccccag gtgcctgacc agatactaca gcagcttcgt caacatggag     1620
agggacctgg cctctggcct gattggcccc ctgctgatct gctacaagga gtctgtggac     1680
5 cagaggggca accagatcat gagcgacaag aggaacgtga tcctgttctc tgtc       1734

```

&lt;210&gt; 14

&lt;211&gt; 2583

&lt;212&gt; ADN

10 &lt;213&gt; Secuencia Artificial

&lt;220&gt;

&lt;223&gt; coFVIII-26-CT

&lt;400&gt; 14



ttcgacgaga acaggagctg gtacctcact gaaaacatcc agaggttcct cccaaacccc	60
gcaggagtgc aactggagga ccctgagttt caggcctcga atatcatgca ctcgattaac	120
ggttacgtgt tcgactcgct gcagctgagc gtgtgcctcc atgaagtcgc ttactggtac	180
attctgtcca tcggcgccca gactgacttc ctgagcgtgt tcttttccgg ttacaccttt	240
aagcacaaga tgggtgtacga agataccctg accctgttcc ctttctccgg cgaaacggtg	300
ttcatgtcga tggagaaccc gggctctgtg attctgggat gccacaacag cgactttcgg	360
aaccgcggaa tgactgccct gctgaagggtg tcctcatgcg acaagaacac cggagactac	420
tacgaggact cctacgagga tatctcagcc tacctcctgt ccaagaacaa cgcgatcgag	480
ccgcgagct tcagccagaa cccgcctgtg ctgaagaggc accagcgaga aattaccggg	540
accaccctcc aatcggatca ggaggaaatc gactacgacg acaccatctc ggtggaaatg	600
aagaaggag atttccgatat ctacgacgag gacgaaaatc agtcccctcg ctcattccaa	660
aagaaaaacta gacactactt tatcgccgcg gtggaagac tgtgggacta tggaatgtca	720
tccagccctc acgtccttcg gaaccgggcc cagagcggat cgggtgcctca gttcaagaaa	780
gtggtgttcc aggagtccac cgacggcagc ttcacccagc cgctgtaccg gggagaactg	840
aacgaacacc tgggcctgct cggtcctac atccgcgagg aagtggagga taacatcatg	900
gtgaccttcc gtaaccaagc atccagacct tactccttct attcctccct gatctcatac	960
gaggaggacc agcgccaagg cgccgagccc cgcaagaact tcgtcaagcc caacgagact	1020
aagacctact tctggaaggc ccaacaccat atggccccga ccaaggatga gtttgactgc	1080
aaggcctggg cctacttctc cgacgtggac cttgagaagg atgtccattc cggcctgac	1140
gggccgctgc tcgtgtgtca caccaacacc ctgaaccag cgcatggacg ccaggtcacc	1200
gtccaggagt ttgctctggt cttcaccatt tttgacgaaa ctaagtcctg gtacttcacc	1260
gagaatatgg agcgaaactg tagagcgccc tgcaatatcc agatggaaga tccgactttc	1320
aaggagaact atagattcca cgccatcaac gggtagatca tggatactct gccggggctg	1380

```

gtcatggccc aggatcagag gattcgggtg tacttgctgt caatgggatc gaacgaaaac 1440
attcactcca ttcacttctc cggtcacgtg ttcactgtgc gcaagaagga ggagtacaag 1500
atggcgctgt acaatctgta ccccggggtg ttcgaaactg tggagatgct gccgtccaag 1560
gccggcatct ggagagtgga gtgcctgacg ggagagcacc tccacgcggg gatgtccacc 1620
ctcttcctgg tgtactcgaa taagtgccag accccgctgg gcatggcctc gggccacatc 1680
agagacttcc agatcacagc aagcggacaa tacggccaat gggcgccgaa gctggcccgc 1740
ttgcactact ccgcatcgat caacgcgatg tccaccaagg aaccgttctc gtggattaag 1800
gtggacctcc tggcccttat gattatccac ggaattaaga cccagggcgc caggcagaag 1860
ttctcctccc tgtacatctc gcaattcatc atcatgtaca gcctggacgga gaagaagtgg 1920
cagacttaca ggggaaactc caccggcacc ctgatggctt ttttcggcaa cgtggattcc 1980
tccggcatta agcacaacat cttcaaccca ccgatcatag ccagatatat taggtccac 2040
cccactcact actcaatccg ctcaactctt cggatggaac tcatggggtg cgacctgaac 2100
tcctgctcca tgcggttggg gatggaatca aaggctatta gcgacgcca gatcacgcgc 2160
agctcctact tcaactaat gttcgccacc tggagcccct ccaaggccag gctgcacttg 2220
cagggacggt caaatgcctg gcggccgcaa gtgaacaatc cgaaggaatg gcttcaagtg 2280
gatttccaaa agaccatgaa agtgaccgga gtcaccaccc agggagtga gtccttctg 2340
acctcgatgt atgtgaagga gttcctgatt agcagcagcc aggacgggca ccagtggacc 2400
ctgttcttcc aaaacggaaa ggtcaagggtg ttccagggga accaggactc gttcacaccc 2460
gtggtgaact ccttgacccc cccactgctg acgcggtact tgaggattca tcctcagtec 2520
tgggtccatc agattgcatt gcgaatggaa gtcctgggct gcgaggcca ggacctgtac 2580
tga 2583

```

&lt;210&gt; 15

&lt;211&gt; 2332

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 15

```

Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser Trp Asp Tyr
1           5           10           15

```

```

Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro
          20           25           30

```

```

Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys
          35           40           45

```

```

Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile Ala Lys Pro

```

# ES 2 926 585 T3

50		55		60
Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu Val				
65		70	75	80
Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro Val				
	85	90		95
Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly Ala				
	100	105		110
Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys Val				
	115	120		125
Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu Asn				
130		135		140
Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu Ser				
145		150	155	160
His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala Leu				
	165	170		175
Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr Leu				
	180	185		190
His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser Trp				
	195	200		205
His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp Ala Ala Ser				
	210	215		220
Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr Val Asn Arg				
225		230	235	240
Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His				
	245	250		255
Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu				
	260	265		270
Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile				
	275	280		285
Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly				
290		295	300	

# ES 2 926 585 T3

Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met  
 305 310 315 320  
 Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg  
 325 330 335  
 Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp  
 340 345 350  
 Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser Pro Ser Phe  
 355 360 365  
 Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr Trp Val His  
 370 375 380  
 Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro Leu Val Leu  
 385 390 395 400  
 Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn Asn Gly Pro  
 405 410 415  
 Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met Ala Tyr Thr  
 420 425 430  
 Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu Ser Gly Ile  
 435 440 445  
 Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu Leu Ile Ile  
 450 455 460  
 Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro His Gly Ile  
 465 470 475 480  
 Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys Gly Val Lys  
 485 490 495  
 His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe Lys Tyr Lys  
 500 505 510  
 Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp Pro Arg Cys  
 515 520 525  
 Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg Asp Leu Ala  
 530 535 540  
 Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu Ser Val Asp  
 545 550 555 560

# ES 2 926 585 T3

Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val Ile Leu Phe  
 565 570 575  
 Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile Gln  
 580 585 590  
 Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu Phe  
 595 600 605  
 Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp Ser  
 610 615 620  
 Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile Leu  
 625 630 635 640  
 Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly Tyr  
 645 650 655  
 Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe Pro  
 660 665 670  
 Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu Trp  
 675 680 685  
 Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr Ala  
 690 695 700  
 Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr Glu  
 705 710 715 720  
 Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala  
 725 730 735  
 Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro Ser Thr Arg  
 740 745 750  
 Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp Ile Glu Lys  
 755 760 765  
 Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys Ile Gln Asn  
 770 775 780  
 Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser Pro Thr Pro  
 785 790 795 800  
 His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr Glu Thr Phe  
 805 810 815

# ES 2 926 585 T3

Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn Ser Leu Ser  
 820 825 830  
 Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly Asp Met Val  
 835 840 845  
 Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu Lys Leu Gly  
 850 855 860  
 Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys Val Ser Ser  
 865 870 875 880  
 Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn Leu Ala Ala  
 885 890 895  
 Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met Pro Val His  
 900 905 910  
 Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys Ser Ser Pro  
 915 920 925  
 Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu Asn Asn Asp  
 930 935 940  
 Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu Ser Ser Trp  
 945 950 955 960  
 Gly Lys Asn Val Ser Ser Thr Glu Ser Gly Arg Leu Phe Lys Gly Lys  
 965 970 975  
 Arg Ala His Gly Pro Ala Leu Leu Thr Lys Asp Asn Ala Leu Phe Lys  
 980 985 990  
 Val Ser Ile Ser Leu Leu Lys Thr Asn Lys Thr Ser Asn Asn Ser Ala  
 995 1000 1005  
 Thr Asn Arg Lys Thr His Ile Asp Gly Pro Ser Leu Leu Ile Glu  
 1010 1015 1020  
 Asn Ser Pro Ser Val Trp Gln Asn Ile Leu Glu Ser Asp Thr Glu  
 1025 1030 1035  
 Phe Lys Lys Val Thr Pro Leu Ile His Asp Arg Met Leu Met Asp  
 1040 1045 1050  
 Lys Asn Ala Thr Ala Leu Arg Leu Asn His Met Ser Asn Lys Thr

# ES 2 926 585 T3

1055		1060		1065
Thr Ser 1070	Ser Lys Asn Met	Glu Met Val Gln Gln 1075	Lys Lys Glu Gly 1080	
Pro Ile 1085	Pro Pro Asp Ala	Gln Asn Pro Asp Met 1090	Ser Phe Phe Lys 1095	
Met Leu 1100	Phe Leu Pro Glu	Ser Ala Arg Trp Ile 1105	Gln Arg Thr His 1110	
Gly Lys 1115	Asn Ser Leu Asn	Ser Gly Gln Gly Pro 1120	Ser Pro Lys Gln 1125	
Leu Val 1130	Ser Leu Gly Pro	Glu Lys Ser Val Glu 1135	Gly Gln Asn Phe 1140	
Leu Ser 1145	Glu Lys Asn Lys	Val Val Val Gly Lys 1150	Gly Glu Phe Thr 1155	
Lys Asp 1160	Val Gly Leu Lys	Glu Met Val Phe Pro 1165	Ser Ser Arg Asn 1170	
Leu Phe 1175	Leu Thr Asn Leu	Asp Asn Leu His Glu 1180	Asn Asn Thr His 1185	
Asn Gln 1190	Glu Lys Lys Ile	Gln Glu Glu Ile Glu 1195	Lys Lys Glu Thr 1200	
Leu Ile 1205	Gln Glu Asn Val	Val Leu Pro Gln Ile 1210	His Thr Val Thr 1215	
Gly Thr 1220	Lys Asn Phe Met	Lys Asn Leu Phe Leu 1225	Leu Ser Thr Arg 1230	
Gln Asn 1235	Val Glu Gly Ser	Tyr Asp Gly Ala Tyr 1240	Ala Pro Val Leu 1245	
Gln Asp 1250	Phe Arg Ser Leu	Asn Asp Ser Thr Asn 1255	Arg Thr Lys Lys 1260	
His Thr 1265	Ala His Phe Ser	Lys Lys Gly Glu Glu 1270	Glu Asn Leu Glu 1275	
Gly Leu 1280	Gly Asn Gln Thr	Lys Gln Ile Val Glu 1285	Lys Tyr Ala Cys 1290	

# ES 2 926 585 T3

Thr Thr	Arg Ile Ser Pro	Asn	Thr Ser Gln Gln	Asn	Phe Val Thr
1295		1300		1305	
Gln Arg	Ser Lys Arg Ala	Leu	Lys Gln Phe Arg	Leu	Pro Leu Glu
1310		1315		1320	
Glu Thr	Glu Leu Glu Lys	Arg	Ile Ile Val Asp	Asp	Thr Ser Thr
1325		1330		1335	
Gln Trp	Ser Lys Asn Met	Lys	His Leu Thr Pro	Ser	Thr Leu Thr
1340		1345		1350	
Gln Ile	Asp Tyr Asn Glu	Lys	Glu Lys Gly Ala	Ile	Thr Gln Ser
1355		1360		1365	
Pro Leu	Ser Asp Cys Leu	Thr	Arg Ser His Ser	Ile	Pro Gln Ala
1370		1375		1380	
Asn Arg	Ser Pro Leu Pro	Ile	Ala Lys Val Ser	Ser	Phe Pro Ser
1385		1390		1395	
Ile Arg	Pro Ile Tyr Leu	Thr	Arg Val Leu Phe	Gln	Asp Asn Ser
1400		1405		1410	
Ser His	Leu Pro Ala Ala	Ser	Tyr Arg Lys Lys	Asp	Ser Gly Val
1415		1420		1425	
Gln Glu	Ser Ser His Phe	Leu	Gln Gly Ala Lys	Lys	Asn Asn Leu
1430		1435		1440	
Ser Leu	Ala Ile Leu Thr	Leu	Glu Met Thr Gly	Asp	Gln Arg Glu
1445		1450		1455	
Val Gly	Ser Leu Gly Thr	Ser	Ala Thr Asn Ser	Val	Thr Tyr Lys
1460		1465		1470	
Lys Val	Glu Asn Thr Val	Leu	Pro Lys Pro Asp	Leu	Pro Lys Thr
1475		1480		1485	
Ser Gly	Lys Val Glu Leu	Leu	Pro Lys Val His	Ile	Tyr Gln Lys
1490		1495		1500	
Asp Leu	Phe Pro Thr Glu	Thr	Ser Asn Gly Ser	Pro	Gly His Leu
1505		1510		1515	
Asp Leu	Val Glu Gly Ser	Leu	Leu Gln Gly Thr	Glu	Gly Ala Ile
1520		1525		1530	



# ES 2 926 585 T3

Lys Trp	Asn Glu	Ala Asn	Arg	Pro Gly	Lys Val	Pro	Phe Leu	Arg
1535			1540			1545		
Val Ala	Thr Glu	Ser Ser	Ala	Lys Thr	Pro Ser	Lys	Leu Leu	Asp
1550			1555			1560		
Pro Leu	Ala Trp	Asp Asn	His	Tyr Gly	Thr Gln	Ile	Pro Lys	Glu
1565			1570			1575		
Glu Trp	Lys Ser	Gln Glu	Lys	Ser Pro	Glu Lys	Thr	Ala Phe	Lys
1580			1585			1590		
Lys Lys	Asp Thr	Ile Leu	Ser	Leu Asn	Ala Cys	Glu	Ser Asn	His
1595			1600			1605		
Ala Ile	Ala Ala	Ile Asn	Glu	Gly Gln	Asn Lys	Pro	Glu Ile	Glu
1610			1615			1620		
Val Thr	Trp Ala	Lys Gln	Gly	Arg Thr	Glu Arg	Leu	Cys Ser	Gln
1625			1630			1635		
Asn Pro	Pro Val	Leu Lys	Arg	His Gln	Arg Glu	Ile	Thr Arg	Thr
1640			1645			1650		
Thr Leu	Gln Ser	Asp Gln	Glu	Glu Ile	Asp Tyr	Asp	Asp Thr	Ile
1655			1660			1665		
Ser Val	Glu Met	Lys Lys	Glu	Asp Phe	Asp Ile	Tyr	Asp Glu	Asp
1670			1675			1680		
Glu Asn	Gln Ser	Pro Arg	Ser	Phe Gln	Lys Lys	Thr	Arg His	Tyr
1685			1690			1695		
Phe Ile	Ala Ala	Val Glu	Arg	Leu Trp	Asp Tyr	Gly	Met Ser	Ser
1700			1705			1710		
Ser Pro	His Val	Leu Arg	Asn	Arg Ala	Gln Ser	Gly	Ser Val	Pro
1715			1720			1725		
Gln Phe	Lys Lys	Val Val	Phe	Gln Glu	Phe Thr	Asp	Gly Ser	Phe
1730			1735			1740		
Thr Gln	Pro Leu	Tyr Arg	Gly	Glu Leu	Asn Glu	His	Leu Gly	Leu
1745			1750			1755		
Leu Gly	Pro Tyr	Ile Arg	Ala	Glu Val	Glu Asp	Asn	Ile Met	Val
1760			1765			1770		

# ES 2 926 585 T3

Thr Phe	Arg Asn Gln Ala Ser	Arg Pro Tyr Ser Phe	Tyr Ser Ser
1775	1780	1785	
Leu Ile	Ser Tyr Glu Glu Asp	Gln Arg Gln Gly Ala	Glu Pro Arg
1790	1795	1800	
Lys Asn	Phe Val Lys Pro Asn	Glu Thr Lys Thr Tyr	Phe Trp Lys
1805	1810	1815	
Val Gln	His His Met Ala Pro	Thr Lys Asp Glu Phe	Asp Cys Lys
1820	1825	1830	
Ala Trp	Ala Tyr Phe Ser Asp	Val Asp Leu Glu Lys	Asp Val His
1835	1840	1845	
Ser Gly	Leu Ile Gly Pro Leu	Leu Val Cys His Thr	Asn Thr Leu
1850	1855	1860	
Asn Pro	Ala His Gly Arg Gln	Val Thr Val Gln Glu	Phe Ala Leu
1865	1870	1875	
Phe Phe	Thr Ile Phe Asp Glu	Thr Lys Ser Trp Tyr	Phe Thr Glu
1880	1885	1890	
Asn Met	Glu Arg Asn Cys Arg	Ala Pro Cys Asn Ile	Gln Met Glu
1895	1900	1905	
Asp Pro	Thr Phe Lys Glu Asn	Tyr Arg Phe His Ala	Ile Asn Gly
1910	1915	1920	
Tyr Ile	Met Asp Thr Leu Pro	Gly Leu Val Met Ala	Gln Asp Gln
1925	1930	1935	
Arg Ile	Arg Trp Tyr Leu Leu	Ser Met Gly Ser Asn	Glu Asn Ile
1940	1945	1950	
His Ser	Ile His Phe Ser Gly	His Val Phe Thr Val	Arg Lys Lys
1955	1960	1965	
Glu Glu	Tyr Lys Met Ala Leu	Tyr Asn Leu Tyr Pro	Gly Val Phe
1970	1975	1980	
Glu Thr	Val Glu Met Leu Pro	Ser Lys Ala Gly Ile	Trp Arg Val
1985	1990	1995	
Glu Cys	Leu Ile Gly Glu His	Leu His Ala Gly Met	Ser Thr Leu

# ES 2 926 585 T3

2000	2005	2010
Phe Leu Val Tyr Ser Asn Lys 2015	Cys Gln Thr Pro 2020	Leu Gly Met Ala 2025
Ser Gly His Ile Arg Asp Phe 2030	Gln Ile Thr Ala 2035	Ser Gly Gln Tyr 2040
Gly Gln Trp Ala Pro Lys Leu 2045	Ala Arg Leu His 2050	Tyr Ser Gly Ser 2055
Ile Asn Ala Trp Ser Thr Lys 2060	Glu Pro Phe Ser 2065	Trp Ile Lys Val 2070
Asp Leu Leu Ala Pro Met Ile 2075	Ile His Gly Ile 2080	Lys Thr Gln Gly 2085
Ala Arg Gln Lys Phe Ser Ser 2090	Leu Tyr Ile Ser 2095	Gln Phe Ile Ile 2100
Met Tyr Ser Leu Asp Gly Lys 2105	Lys Trp Gln Thr 2110	Tyr Arg Gly Asn 2115
Ser Thr Gly Thr Leu Met Val 2120	Phe Phe Gly Asn 2125	Val Asp Ser Ser 2130
Gly Ile Lys His Asn Ile Phe 2135	Asn Pro Pro Ile 2140	Ile Ala Arg Tyr 2145
Ile Arg Leu His Pro Thr His 2150	Tyr Ser Ile Arg 2155	Ser Thr Leu Arg 2160
Met Glu Leu Met Gly Cys Asp 2165	Leu Asn Ser Cys 2170	Ser Met Pro Leu 2175
Gly Met Glu Ser Lys Ala Ile 2180	Ser Asp Ala Gln 2185	Ile Thr Ala Ser 2190
Ser Tyr Phe Thr Asn Met Phe 2195	Ala Thr Trp Ser 2200	Pro Ser Lys Ala 2205
Arg Leu His Leu Gln Gly Arg 2210	Ser Asn Ala Trp 2215	Arg Pro Gln Val 2220
Asn Asn Pro Lys Glu Trp Leu 2225	Gln Val Asp Phe 2230	Gln Lys Thr Met 2235

# ES 2 926 585 T3

Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr  
2240 2245 2250

Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly  
2255 2260 2265

His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe  
2270 2275 2280

Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp  
2285 2290 2295

Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp  
2300 2305 2310

Val His Gln Ile Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala  
2315 2320 2325

Gln Asp Leu Tyr  
2330

<210> 16

<211> 4371

<212> ADN

<213> Secuencia Artificial

<220>

<223> BDD-FVIII (no-optimizado; "parental"), Secuencia de Nucleótidos

<400> 16

atgcaaatag agctctccac ctgcttcttt ctgtgccttt tgcgattctg ctttagtgcc	60
accagaagat actacctggg tgcagtgga ctgtcatggg actatatgca aagtgatctc	120
ggtgagctgc ctgtggacgc aagatttcct cctagagtgc caaaatcttt tccattcaac	180
acctcagtcg tgtacaaaaa gactctgttt gtagaattca cggatcacct tttcaacatc	240
gctaagccaa ggccaccctg gatgggtctg ctaggtccta ccatccaggc tgaggtttat	300
gatacagtggt tcattacact taagaacatg gcttcccatc ctgtcagctc tcatgctgtt	360
ggtgtatcct actggaaagc ttctgagga gctgaatatg atgatcagac cagtcaaagg	420
gagaaagaag atgataaagt cttccctggt ggaagccata catatgtctg gcaggtcctg	480
aaagagaatg gtccaatggc ctctgaccca ctgtgcctta cctactcata tctttctcat	540
gtggacctgg taaaagactt gaattcaggc ctcatggag ccctactagt atgtagagaa	600
gggagtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact ttttgctgta	660
tttgatgaag ggaaaagttg gcaactcagaa acaaagaact ccttgatgca ggatagggat	720
gctgcatctg ctcgggcctg gcctaaaaatg cacacagtca atggttatgt aaacaggtct	780

ctgccaggtc tgattggatg ccacaggaaa tcagtctatt ggcattgtgat tggaaatgggc	840
accactcctg aagtgcactc aatattcctc gaaggtcaca catttcttgt gaggaaccat	900
cgccaggcgt ccttggaat ctcgccaata actttcctta ctgctcaaac actcttgatg	960
gaccttgac agtttctact gttttgtcat atctcttccc accaacaatga tggcatggaa	1020
gcttatgtca aagtagacag ctgtccagag gaacccaac tacgaatgaa aaataatgaa	1080
gaagcggag actatgatga tgatcttact gattctgaaa tggatgtggc caggtttgat	1140
gatgacaact ctcttctctt tatccaaatt cgctcagttg ccaagaagca tcctaaaact	1200
tgggtacatt acattgctgc tgaagaggag gactgggact atgctccctt agtctctgcc	1260
cccgatgaca gaagttataa aagtcaatat ttgaacaatg gccctcagcg gattggtagg	1320
aagtacaaaa aagtccgatt tatggcatac acagatgaaa cctttaagac tcgtgaagct	1380
attcagcatg aatcaggaat cttgggacct ttactttatg gggaagttgg agacacactg	1440
ttgattatat ttaagaatca agcaagcaga ccatataaca tctaccctca cggaatcact	1500
gatgtcgcgt ctttgtattc aaggagatta ccaaaagggtg taaaacattt gaaggatttt	1560
ccaattctgc caggagaaat attcaaatat aaatggacag tgactgtaga agatgggcca	1620
actaaatcag atcctcgggtg cctgaccgcg tattactcta gtttcgttaa tatggagaga	1680
gatctagctt caggactcat tggccctctc ctcatctgct acaaagaatc tgtagatcaa	1740
agaggaaacc agataatgtc agacaagagg aatgtcatcc tgttttctgt atttgatgag	1800
aaccgaagct ggtacctcac agagaatata caacgctttc tccccaatcc agctggagtg	1860
cagcttgagg atccagagtt ccaagcctcc aacatcatgc acagcatcaa tggctatgtt	1920
tttgatagtt tgcagttgtc agtttgtttg catgaggtgg catactggta cattctaagc	1980
attggagcac agactgaact cctttctgtc ttcttctctg gatatacctt caaacacaaa	2040
atggtctatg aagacacact caccctattc ccattctcag gagaaactgt cttcatgtcg	2100
atggaaaacc cagggtctatg gattctgggg tgccacaact cagactttcg gaacagaggc	2160
atgaccgcct tactgaaggt ttctagtgtg gacaagaaca ctggtgatta ttacgaggac	2220
agttatgaag atatttcagc atacttgctg agtaaaaaaca atgccattga accaagaagc	2280
ttctctcaaa acccaccagt cttgaaacgc catcaacggg aaataactcg tactactctt	2340
cagtcagatc aagaggaaat tgactatgat gataccatat cagttgaaat gaagaaggaa	2400
gattttgaca tttatgatga ggatgaaaat cagagccccc gcagctttca aaagaaaaca	2460
cgacactatt ttattgctgc agtgagagg ctctgggatt atgggatgag tagctcccca	2520
catgttctaa gaaacagggc tcagagtggc agtgtccctc agttcaagaa agttgttttc	2580
caggaattta ctgatggctc ctttactcag cccttatacc gtggagaact aaatgaacat	2640
ttgggactcc tggggccata tataagagca gaagttgaag ataatatcat ggtaactttc	2700

```

agaaatcagg cctctogtcc ctattccttc tattctagcc ttatttctta tgaggaagat 2760
cagaggcaag gagcagaacc tagaaaaaac tttgtcaagc ctaatgaaac caaaacttac 2820
ttttggaaag tgcaacatca tatggcacc ctaaaagatg agtttgactg caaagcctgg 2880
gcttatttct ctgatgttga cctggaaaa gatgtgcact caggcctgat tggaccctt 2940
ctggtctgcc aactaacac actgaaccct gctcatggga gacaagtga agtacaggaa 3000
tttgctctgt ttttcacat ctttgatgag accaaaagct ggtacttcac tgaaaatatg 3060
gaaagaaact gcagggtcc ctgcaatata cagatggaag atcccacttt taaagagaat 3120
tatcgcttcc atgcaatcaa tggctacata atggatacac tacctggctt agtaatggct 3180
caggatcaaa ggattcgatg gtatctgctc agcatgggca gcaatgaaaa catccattct 3240
attcatttca gtggacatgt gttcactgta cgaaaaaag aggagtataa aatggcactg 3300
tacaatctct atccagggtg ttttgagaca gtggaaatgt taccatccaa agctggaatt 3360
tggcgggtgg aatgccttat tggcgagcat ctacatgctg ggatgagcac actttttctg 3420
gtgtacagca ataagtgtca gactcccctg ggaatggctt ctggacacat tagagatttt 3480
cagattacag cttcaggaca atatggacag tgggccccaa agctggccag acttcattat 3540
tccggatcaa tcaatgcctg gagcaccaag gagccctttt cttggatcaa ggtggatctg 3600
ttggcaccaa tgattattca cggcatcaag acccaggggtg cccgtcagaa gttctccagc 3660
ctctacatct ctcagtttat catcatgtat agtcttgatg ggaagaagtg gcagacttat 3720
cgaggaaatt ccaactggaac cttaatggtc ttctttggca atgtggatc atctgggata 3780
aaacacaata tttttaacct tccaattatt gctcgataca tccgtttgca cccaactcat 3840
tatagcatc gcagcactct tcgcatggag ttgatgggct gtgattttaa tagttgcagc 3900
atgccattgg gaatggagag taaagcaata tcagatgcac agattactgc ttcacctac 3960
tttaccaata tgtttgccac ctggtctcct tcaaaagctc gacttcacct ccaagggagg 4020
agtaatgcct ggagacctca ggtgaataat ccaaagagt ggctgcaagt ggacttccag 4080
aagacaatga aagtcacagg agtaactact caggagtaa aatctctgct taccagcatg 4140
tatgtgaagg agttcctcat ctccagcagt caagatggcc atcagtggac tctctttttt 4200
cagaatggca aagtaaaggt ttttcaggga aatcaagact ccttcacacc tgtggtgaac 4260
tctctagacc caccgttact gactcgctac cttcgaattc acccccagag ttgggtgcac 4320
cagattgccc tgaggatgga ggttctgggc tgcgaggcac aggacctcta c 4371

```

<210> 17

<211> 1438

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> BDD-FVIII (no-optimizado; "parental"), Secuencia de Aminoácidos

10

<400> 17

# ES 2 926 585 T3

Ala	Thr	Arg	Arg	Tyr	Tyr	Leu	Gly	Ala	Val	Glu	Leu	Ser	Trp	Asp	Tyr	1	5	10	15
Met	Gln	Ser	Asp	Leu	Gly	Glu	Leu	Pro	Val	Asp	Ala	Arg	Phe	Pro	Pro	20	25	30	
Arg	Val	Pro	Lys	Ser	Phe	Pro	Phe	Asn	Thr	Ser	Val	Val	Tyr	Lys	Lys	35	40	45	
Thr	Leu	Phe	Val	Glu	Phe	Thr	Asp	His	Leu	Phe	Asn	Ile	Ala	Lys	Pro	50	55	60	
Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln	Ala	Glu	Val	65	70	75	80
Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser	His	Pro	Val	85	90	95	
Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser	Glu	Gly	Ala	100	105	110	
Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp	Asp	Lys	Val	115	120	125	
Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu	Lys	Glu	Asn	130	135	140	
Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser	Tyr	Leu	Ser	145	150	155	160
His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile	Gly	Ala	Leu	165	170	175	
Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr	Gln	Thr	Leu	180	185	190	
His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly	Lys	Ser	Trp	195	200	205	
His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp	Ala	Ala	Ser	210	215	220	
Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr	Val	Asn	Arg	225	230	235	240

# ES 2 926 585 T3

Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His  
 245 250 255  
 Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu  
 260 265 270  
 Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile  
 275 280 285  
 Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly  
 290 295 300  
 Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met  
 305 310 315 320  
 Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg  
 325 330 335  
 Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp  
 340 345 350  
 Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser Pro Ser Phe  
 355 360 365  
 Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr Trp Val His  
 370 375 380  
 Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro Leu Val Leu  
 385 390 395 400  
 Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn Asn Gly Pro  
 405 410 415  
 Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met Ala Tyr Thr  
 420 425 430  
 Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu Ser Gly Ile  
 435 440 445  
 Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu Leu Ile Ile  
 450 455 460  
 Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro His Gly Ile  
 465 470 475 480  
 Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys Gly Val Lys



# ES 2 926 585 T3

485										490					495				
His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	Lys	Tyr	Lys				
			500						505						510				
Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	Pro	Arg	Cys				
		515					520						525						
Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg	Asp	Leu	Ala				
	530					535					540								
Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	Ser	Val	Asp				
545					550					555					560				
Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	Ile	Leu	Phe				
			565						570						575				
Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu	Asn	Ile	Gln				
			580					585						590					
Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp	Pro	Glu	Phe				
		595					600						605						
Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val	Phe	Asp	Ser				
	610						615					620							
Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp	Tyr	Ile	Leu				
625					630					635					640				
Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe	Ser	Gly	Tyr				
				645					650					655					
Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr	Leu	Phe	Pro				
			660					665						670					
Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro	Gly	Leu	Trp				
		675					680					685							
Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly	Met	Thr	Ala				
	690					695					700								
Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp	Tyr	Tyr	Glu				
705					710					715					720				
Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys	Asn	Asn	Ala				
			725						730					735					

# ES 2 926 585 T3

Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu Lys Arg His  
740 745 750

Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile  
755 760 765

Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp  
770 775 780

Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys  
785 790 795 800

Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly  
805 810 815

Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser  
820 825 830

Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser  
835 840 845

Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly Leu  
850 855 860

Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr  
865 870 875 880

Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile  
885 890 895

Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe  
900 905 910

Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His  
915 920 925

Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe  
930 935 940

Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly Pro  
945 950 955 960

Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln  
965 970 975

Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr  
980 985 990

# ES 2 926 585 T3

Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro	995	1000	1005
Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg	1010	1015	1020
Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu	1025	1030	1035
Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met	1040	1045	1050
Gly Ser Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val	1055	1060	1065
Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn	1070	1075	1080
Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys	1085	1090	1095
Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His	1100	1105	1110
Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln	1115	1120	1125
Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile	1130	1135	1140
Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg	1145	1150	1155
Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro	1160	1165	1170
Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His	1175	1180	1185
Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr	1190	1195	1200
Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp	1205	1210	1215
Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe	1220	1225	1230

Gly Asn Val Asp Ser Ser Gly	Ile Lys His Asn Ile Phe Asn Pro
1235	1240 1245
Pro Ile Ile Ala Arg Tyr Ile	Arg Leu His Pro Thr His Tyr Ser
1250	1255 1260
Ile Arg Ser Thr Leu Arg Met	Glu Leu Met Gly Cys Asp Leu Asn
1265	1270 1275
Ser Cys Ser Met Pro Leu Gly	Met Glu Ser Lys Ala Ile Ser Asp
1280	1285 1290
Ala Gln Ile Thr Ala Ser Ser	Tyr Phe Thr Asn Met Phe Ala Thr
1295	1300 1305
Trp Ser Pro Ser Lys Ala Arg	Leu His Leu Gln Gly Arg Ser Asn
1310	1315 1320
Ala Trp Arg Pro Gln Val Asn	Asn Pro Lys Glu Trp Leu Gln Val
1325	1330 1335
Asp Phe Gln Lys Thr Met Lys	Val Thr Gly Val Thr Thr Gln Gly
1340	1345 1350
Val Lys Ser Leu Leu Thr Ser	Met Tyr Val Lys Glu Phe Leu Ile
1355	1360 1365
Ser Ser Ser Gln Asp Gly His	Gln Trp Thr Leu Phe Phe Gln Asn
1370	1375 1380
Gly Lys Val Lys Val Phe Gln	Gly Asn Gln Asp Ser Phe Thr Pro
1385	1390 1395
Val Val Asn Ser Leu Asp Pro	Pro Leu Leu Thr Arg Tyr Leu Arg
1400	1405 1410
Ile His Pro Gln Ser Trp Val	His Gln Ile Ala Leu Arg Met Glu
1415	1420 1425
Val Leu Gly Cys Glu Ala Gln	Asp Leu Tyr
1430	1435

<210> 18

<211> 450

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> Secuencia de nucleótidos de XTEN

10

<400> 18

ggcgcgccaa catcagagag cgccaccctt gaaagtggc ccgggagcga gccagccaca 60  
 tctgggtcgg aaacgccagg cacaagttag tctgcaactc ccgagtcagg acctggctcc 120  
 gagcctgcca ctagcggctc cgagactccg ggaacttccg agagcgctac accagaaagc 180  
 ggacccggaa ccagtaccga acctagcgag ggctctgctc cgggcagccc agccggctct 240  
 cctacatcca cggaggaggg cacttccgaa tccgccacc ccgagtcagg gccaggatct 300  
 gaacccgcta cctcaggcag tgagacgcca ggaacgagcg agtccgctac accggagagt 360  
 gggccaggga gccctgctgg atctcctacg tccactgagg aagggtcacc agcgggctcg 420  
 cccaccagca ctgaagaagg tgctcagagc 450

<210> 19  
 <211> 4824  
 <212> ADN  
 <213> Secuencia Artificial

<220>  
 <223> coFVIII-52-XTEN

<400> 19

atgcaaatcg aactgagcac ctgtttcttc ctctgcctgc tgagattctg tttctccgag 60  
 acccgccgat actacctggg agcagtgagg ctctcctggg attacatgca gagcgacctt 120  
 ggggagctgc ccgtggatgc caggttccct ccccggtgac caaagtcgtt tccgttcaac 180  
 acctccgtgg tgtacaagaa aactctgttc gtggagttca ccgaccacct gttcaatatc 240  
 gccaaagcca gacctccctg gatggggctg ttgggacctt ccatccaagc ggaggtgtac 300  
 gacactgtgg tcatcactct gaagaacatg gcctcgcatc ccgtgtccct gcacgccgtg 360  
 ggagtgtctt actggaaagc gtccgagggg gccgaatacg acgaccagac ctgcagagaa 420  
 gaaaaggaag atgacaaggt gttcccagga ggatcgacaa cctacgtgtg gcaagtgttg 480  
 aaggagaacg gcccaatggc ctccgaccgg ctgtgcctga cctactcgta cctgtccac 540  
 gtggacctcg tgaaggacct caactcggga ctgattggag ccctgctggt ctgcagggaa 600  
 ggctcactgg cgaaagaaaa gactcagacc ttgcacaagt tcattctgct gttcgtctgtg 660  
 ttcgacgagg ggaagtctgt gcacagcgag actaagaact ccctgatgca agatagagat 720  
 gccgcctccg cccgggcctg gcctaagatg cacaccgtga acggttacgt gaaccgctcc 780  
 ctccctggcc tgattggatg ccaccggaag tccgtgtact ggcacgtgat cgggatgggg 840  
 accacccccg aggtgcacag catcttcctg gaaggtcaca catttctcgt gcgcaaccac 900  
 cggcaggcct ccctggaaat cagccccatt accttcctca ctgccagac tctgctgatg 960  
 gacctgggac agttcctgct gttctgceat atctcctccc accaacaatga cggaatggag 1020

# ES 2 926 585 T3

gcatacgtga aggtcgattc ctgccctgag gaaccccagc tccgcatgaa gaacaatgag	1080
gaagccgagg actacgacga cgacctgacg gatagcgaga tggatgtggt ccggttcgat	1140
gacgataaca gcccttcctt catccaaatt cgctcgggtg caaagaagca ccccaagacc	1200
tgggtgcatt acattgcggc ggaagaagag gactgggatt atgccccgct tgtcctcgct	1260
cctgacgacc ggagctacaa gagccagtac ctgaacaacg gtccacagag gatcggtaga	1320
aagtacaaga aggtccgctt catggcctat accgacgaaa ccttcaaaac tagagaggcc	1380
atccaacacg aatccggcat cctgggcccg ctcttgtacg gagaagtcgg cgacaccctt	1440
ctcattatct tcaagaacca ggcttcccgg ccgtacaaca tctatccgca tgggatcact	1500
gacgtgcgcc cactgtactc gcggcgctg cccaaggggtg tcaaacacct gaaggatttt	1560
ccgatccttc cgggagaaat cttcaagtac aagtggaccg tgaccgtgga agatggccca	1620
actaagtctg accctagatg cctcaccgcg tactactcat ccttcgtcaa catggagcgc	1680
gacctggcca gcggactgat cggcccgtg ctgatttgcg acaaggaatc agtggaccaa	1740
cggggaaacc agatcatgtc ggataagagg aacgtcatcc tcttctcgt gtttgacgaa	1800
aaccggctgt ggtacctgac cgagaacatc cagaggttcc tgcccaacct tgctggggtg	1860
cagctggagg accccgagtt ccaggccagc aacatcatgc acagcatcaa tggctacgtg	1920
ttcgacagcc tgcagctgag cgtgtgcctg cacgaggtgg cctactggta catcctgagc	1980
atcggcgccc agaccgactt cctgagcgtg ttcttctctg gctacacctt caagcacaag	2040
atggtgtatg aggacaccct gacctgttc cccttcagcg gggagactgt cttcatgagc	2100
atggagaacc ctggcctgtg gatcctgggc tgccacaaca gcgacttcag gaacaggggc	2160
atgactgccc tgctgaaagt ctccagctgt gacaagaaca ccggggacta ctacgaggac	2220
agctacgagg acatcagcgc ctacctgctg agcaagaaca atgccatcga gcccaggagc	2280
ttctctcaga acggcgcgcc aacatcagag agcgccacct ctgaaagtgg tcccgggagc	2340
gagccagcca catctgggtc ggaaacgcca ggcacaagtg agtctgcaac tcccaggtcc	2400
ggacctggct ccgagcctgc cactagcggc tccgagactc cgggaacttc cgagagcgct	2460
acaccagaaa gcggaccccg aaccagtacc gaacctagcg agggctctgc tccgggcagc	2520
ccagccggct ctctacatc cacggaggag ggcacttccg aatccgccac cccggagtca	2580
gggccaggat ctgaaccgcg tacctcaggc agtgagacgc caggaacgag cgagtccgct	2640
acaccggaga gtgggcccag gagccctgct ggatctccta cgtccactga ggaagggtca	2700
ccagcgggct cggccaccag cactgaagaa ggtgcctcga gccccccagt gctgaagagg	2760
caccagaggg agatcaccag gaccaccctg cagtctgacc aggaggagat cgactatgat	2820
gacaccatca gcgtggagat gaagaaggag gacttcgaca tctacgacga ggacgagaac	2880

# ES 2 926 585 T3

```

cagagcccca ggagcttcca gaagaagacc aggcactact tcattgctgc tgtggagagg 2940
ctgtgggact atggcatgtc cagcagcccc catgtgctga ggaacagggc ccagtctggc 3000
agcgtgcccc agttcaagaa agtcgtgttc caggagtcca ccgacggcag cttcaccag 3060
cccctgtaca gaggggagct gaacgagcac ctgggcctgc tgggccccta catcagggcc 3120
gaggtggagg acaacatcat ggtgacctc aggaaccagg ccagcaggcc ctacagcttc 3180
tacagcagcc tgatcagcta cgaggaggac cagaggcagg gggctgagcc caggaagaac 3240
tttgtgaagc ccaatgaaac caagacctac ttctggaagg tgcagcacca catggccccc 3300
accaaggacg agttcgactg caaggcctgg gcctacttct ctgacgtgga cctggagaag 3360
gacgtgcact ctggcctgat tggcccctg ctggtgtgcc acaccaacac cctgaaccct 3420
gcccattgga ggcaggtgac tgtgcaggag ttgcacctgt tcttcacat cttcgatgaa 3480
accaagagct ggtacttcac tgagaacatg gagaggaact gcagggcccc ctgcaacatc 3540
cagatggagg accccacctt caaggagaac tacaggttcc atgcatcaa tggctacatc 3600
atggacaccc tgccctggcct ggtcatggcc caggaccaga ggatcaggtg gtatctgctg 3660
agcatgggca gcaacgagaa catccacagc atccacttct ctggccacgt gttcactgtg 3720
aggaagaagg aggagtacaa gatggccctg tacaacctgt accctggggt gttcgaaacc 3780
gtggagatgc tgcccagcaa ggccggcatc tggagggtgg agtgccctgat tggggagcac 3840
ctgcacgccg gcatgagcac cctgttctct gtgtacagca acaagtgcc aacccccctg 3900
ggcatggcct ctggccacat cagggaettc cagatcaactg cctctggcca gtacggccag 3960
tgggccccca agctggccag gctgcactac tccggaagca tcaatgcctg gagcaccaag 4020
gagcccttca gctggatcaa agtggacctg ctggccccc tgatcatcca cggcatcaag 4080
accaggggg ccaggcagaa gttctccagc ctgtacatca gccagttcat catcatgtac 4140
agcctggacg gcaagaagtg gcagacctac aggggcaaca gcaccggcac cctgatggtg 4200
ttcttcggca acgtggacag cagcggcatc aagcacaaca tcttcaacct ccccatcatc 4260
gccagatata tcaggctgca cccacccac tacagcatca ggagcaccct gaggatggag 4320
ctgatgggct gtgacctgaa cagctgcagc atgccctgg gcatggagag caaggccatc 4380
tctgacgcc agatcactgc ctccagctac ttcaccaaca tgtttgccac ctggagcccc 4440
agcaaggcca ggctgcacct gcagggcagg agcaatgcct ggaggcccca ggtcaacaac 4500
cccaaggagt ggctgcagggt ggacttccag aagaccatga aggtgactgg ggtgaccacc 4560
cagggggtga agagcctgct gaccagcatg tacgtgaagg agttcctgat ctccagcagc 4620
caggacggcc accagtggac cctgttcttc cagaatggca aggtgaagggt gttccagggc 4680
aaccaggaca gcttcacccc tgtggtcaac agcctggacc cccccctgct gaccagatac 4740
ctgaggatcc acccccagag ctgggtgcac cagatcgccc tgaggatgga ggtgctgggc 4800
tgtgaggccc aggacctgta ctga 4824

```

<210> 20  
 <211> 4824  
 <212> ADN  
 <213> Secuencia Artificial

<220>  
 <223> coFVIII-1-XTEN

&lt;400&gt; 20

atgcagattg agctgtctac ttgctttttc ctgtgcctgc tgaggttttg cttttccgct	60
acacgaaggt attatctggg ggctgtggaa ctgtcttggg attacatgca gagtgacctg	120
ggagagctgc cagtggacgc aagggtttccc cctagagtcc ctaagtcatt ccccttcaac	180
actagcgtgg tctacaagaa aacactgttc gtggagttaa ctgatcacct gttcaacatc	240
gcaaagccta ggccaccctg gatgggactg ctggggccaa caatccaggc cgaggtgtac	300
gacaccgtgg tcattacact taagaacatg gcctcacacc ccgtgagcct gcatgctgtg	360
ggcgtcagct actggaaggc ttccgaagga gcagagtatg acgatcagac ttcccagaga	420
gaaaaagagg acgataaggt gtttcctggc ggatctcata cctacgtgtg gcaggtcctg	480
aaagagaatg gccctatggc ctccgaccct ctgtgcctga cctactctta tctgagtcac	540
gtggacctgg tcaaggatct gaacagcggc ctgatcggag ccctgctggt gtgcagggaa	600
ggaagcctgg ctaaggagaa aaccagaca ctgcataagt tcattctgct gttcgccgtg	660
tttgacgaag ggaaatcatg gcacagcgag acaaagaata gtctgatgca ggacagggat	720
gccgcttcag ccagagcttg gcccaaatg cactctgtga acggctacgt caatcgctca	780
ctgcctgggc tgatcggctg ccaccgaaag agcgtgtatt ggcatgtcat cgggatgggc	840
accacacctg aagtgcactc cttttcctg gagggacata cttttctggt ccgcaaccac	900
cgacaggctt ccttgagat ctctccaatt accttctga cagcacagac tctgctgatg	960
gacctggggc agttcctgct gttttgccac atcagctccc accagcatga tggcatggag	1020
gcttacgtga aagtggactc ttgtcccag gaacctcagc tgcggatgaa gaacaatgag	1080
gaagcagaag actatgacga tgacctgacc gactccgaga tggatgtggt ccgattcgat	1140
gacgataaca gccctcctt tatccagatt agatctgtgg ccaagaaaca ccctaagaca	1200
tgggtccatt acatcgacgc cgaggaagag gactgggatt atgcaccact ggtgctggca	1260
ccagacgatc gctcctacaa atctcagtat ctgaacaatg ggccacagag gattggcaga	1320
aagtacaaga aagtgcggtt catggcatat accgatgaga ctttcaagac tcgcgaagcc	1380
atccagcacg agagcggcat cctgggacca ctgctgtacg gagaagtggg agacacctg	1440
ctgatcattt tcaagaacca ggccagccgg ccttacaata tctatccaca tgggattaca	1500



gatgtgcgcc	ctctgtacag	caggagactg	ccaaaggcgc	tcaaacacct	gaaggacttc	1560
ccaatcctgc	ccggagaaat	cttcaagtac	aagtggactg	tcaccgtcga	ggatggcccc	1620
actaagagcg	accctcggtg	cctgaccgcg	tactattcta	gtttcgtgaa	tatggaaaaga	1680
gatctggcaa	gcggaactgat	cggaccactg	ctgatttggt	acaaagagag	cgtggatcag	1740
agaggcaacc	agatcatgtc	cgacaagcgg	aatgtgattc	tgttcagtgt	ctttgacgaa	1800
aacaggtcac	ggtacctgac	cgagaacatc	cagagattcc	tgcctaatac	agctgggggtg	1860
cagctggaag	atcctgagtt	tcaggcatct	aacatcatgc	atagtattaa	tggctacgtg	1920
ttcgacagtt	tgcagctgag	cgtgtgcctg	cacgaggtcg	cttactggta	tatcctgagc	1980
attggggcac	agacagattt	cctgagcgtg	ttcttttccg	gtacactttt	taagcataaa	2040
atggtctatg	aggacacact	gactctgttc	cccttcagcg	gcgaaaccgt	gtttatgagc	2100
atggagaatc	ccggactgtg	gattctgggg	tgccacaaca	gcgatttcag	aaatcgcgga	2160
atgactgccc	tgctgaaagt	gtcaagctgt	gacaagaaca	ccggggacta	ctatgaagat	2220
tcatacgagg	acatcagcgc	atatctgctg	tccaaaaaca	atgccattga	accccggtct	2280
tttagtcaga	atggcgcgcc	aacatcagag	agcgccaccc	ctgaaagtgg	tcccgggagc	2340
gagccagcca	catctgggtc	ggaaacgcca	ggcacaagtg	agtctgcaac	tcccgagtcc	2400
ggacctggct	ccgagcctgc	cactagcggc	tccgagactc	cgggaacttc	cgagagcgtc	2460
acaccagaaa	gcggaaccgc	aaccagtacc	gaacctagcg	agggctctgc	tccgggcagc	2520
ccagccggct	ctcctacatc	cacggaggag	ggcacttccg	aatccgccac	cccggagtca	2580
gggccaggat	ctgaaccgcg	tacctcaggc	agtgagacgc	caggaacgag	cgagtccgct	2640
acaccggaga	gtgggccagg	gagccctgct	ggatctccta	cgtccactga	ggaagggtca	2700
ccagcgggct	cgcccaccag	cactgaagaa	ggtgcctcga	gccctccagt	gctgaagcgg	2760
caccagcgcg	agatcaccgc	cactaccctg	cagagtgatc	aggaagagat	cgactacgac	2820
gatacaattt	ctgtggaaat	gaagaaagag	gacttcgata	tctatgacga	agatgagaac	2880
cagagtcctc	gatcattcca	gaagaaaacc	aggcattact	ttattgccgc	agtggagcgg	2940
ctgtgggatt	atggcatgtc	ctctagtcc	cacgtgctgc	gaaatagggc	ccagtcagga	3000
agcgtcccac	agttcaagaa	agtgtcttc	caggagttta	cagacgggtc	ctttactcag	3060
ccactgtaca	ggggcgaaact	gaacgagcac	ctgggactgc	tggggcccta	tatcagagca	3120
gaagtggagg	ataacattat	ggtcaccttc	agaaatcagg	cctctcggcc	ttacagtttt	3180
tattcaagcc	tgatctctta	cgaagaggac	cagcgacagg	gagctgaacc	acgaaaaaac	3240
ttcgtgaagc	ctaatgagac	caaaacatac	ttttggaagg	tgcagcacca	tatggcccca	3300
acaaaagacg	agttcgattg	caaggcatgg	gcctattttt	ctgacgtgga	tctggagaag	3360
gacgtgcaca	gtggcctgat	tggcccactg	ctggtgtgcc	atactaacac	cctgaatcca	3420

gcccacggcc ggccaggtcac tgtccaggag ttcgctctgt tctttaccat ctttgatgag 3480  
 acaaagagct ggtacttcac cgaaaacatg gagcgaaatt gcagggctcc atgtaacatt 3540  
 cagatggaag accccacatt caaggagaac taccgctttc atgctatcaa tggatacatc 3600  
 atggatactc tgcccgggct ggtcatggca caggaccaga gaatccggtg gtatctgctg 3660  
 agcatgggca gcaacgagaa tatccactca attcatttca gcgggcacgt gtttactgtc 3720  
 aggaagaaag aagagtacaa gatggccctg tacaacctgt atcccggtg gttcgaaacc 3780  
 gtcgagatgc tgcctagcaa ggccggaatc tggagagtgg aatgcctgat tggagagcac 3840  
 ctgcatgctg ggaagtctac cctgtttctg gtgtacagta ataagtgtca gacaccctg 3900  
 ggaatggcat ccgggcatat cagggtttc cagattaccg catctggaca gtacggacag 3960  
 tgggcacctc agctggctag actgcactat tccggatcta tcaacgcttg gtccacaaaa 4020  
 gagcctttct cttggattaa ggtggacctg ctggcccaa tgatcattca tggcatcaaa 4080  
 actcaggag ctcggcagaa gttctcctct ctgtacatct cacagtttat catcatgtac 4140  
 agcctggatg ggaagaaatg gcagacatac cgcggaata gcacaggaaac tctgatggg 4200  
 ttctttggca acgtggacag cagcggaatc aagcacaaca ttttcaatcc ccctatcatt 4260  
 gctagataca tccggctgca cccaacccat tattctattc gaagtacact gaggatggaa 4320  
 ctgatgggat gcgatctgaa cagttgttca atgcccctgg ggatggagtc caaggcaatc 4380  
 tctgacgcc agattaccgc cagctcctac ttcactaata tgtttgctac ctggagccct 4440  
 tccaaagcaa gactgcacct gcaaggccgc agcaacgcat ggcgaccaca ggtgaacaat 4500  
 cccaaggagt ggttgagggt cgattttcag aaaactatga aggtgaccgg ggtcacaact 4560  
 cagggcgtga aaagtctgct gacctcaatg tacgtcaagg agttcctgat ctctagtcca 4620  
 caggacggac atcagtggac actgttcttt cagaacggga aggtgaaagt cttccagggc 4680  
 aatcaggatt cctttacacc tgtggtcaac agtctagacc ctccactgct gaccagatac 4740  
 ctgagaatcc accctcagtc ctgggtgcac cagattgccc tgagaatgga agtgctggga 4800  
 tgcgaggccc aggatctgta ctga 4824

<210> 21

<211> 6

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> MAR/ARS

10

<400> 21

atattt 6

15 <210> 22

<211> 6

<212> ADN

<213> Secuencia Artificial

<220>

20 <223> MAR/ARS

<400> 22

aaatat 6

5 <210> 23  
 <211> 5  
 <212> ADN  
 <213> Secuencia Artificial

10 <220>  
 <223> elemento desestabilizante

<400> 23

15 attta 5

<210> 24  
 <211> 5  
 <212> ADN  
 <213> Secuencia Artificial

20 <220>  
 <223> elemento desestabilizante

<400> 24

25 taaat 5

<210> 25  
 <211> 6  
 <212> ADN  
 <213> Secuencia Artificial

30 <220>  
 <223> sitio poli-T

35 <400> 25

tttttt 6

40 <210> 26  
 <211> 7  
 <212> ADN  
 <213> Secuencia Artificial

45 <220>  
 <223> sitio poli-A

<400> 26

50 aaaaaaa 7

<210> 27  
 <211> 6  
 <212> ADN  
 <213> Secuencia Artificial

55 <220>  
 <223> sitio de corte y empalme

60 <400> 27

ggtgat 6

<210> 28  
 <211> 5

65

<212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 5 <223> caja TATA  
  
 <400> 28  
  
 tataa 5  
 10  
 <210> 29  
 <211> 5  
 <212> ADN  
 <213> Secuencia Artificial  
 15  
 <220>  
 <223> caja TATA  
  
 <400> 29  
 20  
 ttata 5  
  
 <210> 30  
 <211> 8  
 25 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 <223> elemento de secuencia rico en AU  
 30  
 <400> 30  
  
 attttatt 8  
  
 35 <210> 31  
 <211> 8  
 <212> ADN  
 <213> Secuencia Artificial  
  
 40 <220>  
 <223> elemento de secuencia rico en AU  
  
 <400> 31  
  
 45 atttttaa 8  
  
 <210> 32  
 <211> 13  
 <212> ADN  
 50 <213> Secuencia Artificial  
  
 <220>  
 <223> secuencia consenso kozak  
  
 55 <400> 32  
  
 gccgccacca tgc 13  
  
 <210> 33  
 60 <211> 32  
 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 65 <223> péptido CTP

<400> 33

Asp Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser  
1 5 10 15

Leu Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu  
20 25 30

5 <210> 34  
<211> 28  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> péptido CTP

<400> 34

Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu Pro Ser Pro Ser Arg  
1 5 10 15

15 Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln  
20 25

<210> 35  
<211> 11  
<212> PRT  
20 <213> Secuencia Artificial

<220>  
<223> secuencia del núcleo de péptido de unión a albúmina

25 <400> 35

Asp Ile Cys Leu Pro Arg Trp Gly Cys Leu Trp  
1 5 10

30 <210> 36  
<211> 20  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> secuencia PAS

<400> 36

Ala Ser Pro Ala Ala Pro Ala Pro Ala Ser Pro Ala Ala Pro Ala Pro  
1 5 10 15

40 Ser Ala Pro Ala  
20

<210> 37  
<211> 20  
<212> PRT  
<213> Secuencia Artificial

45 <220>  
<223> secuencia PAS

<400> 37

Ala Ala Pro Ala Ser Pro Ala Pro Ala Ala Pro Ser Ala Pro Ala Pro  
1 5 10 15

Ala Ala Pro Ser  
20

5 <210> 38  
<211> 20  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> secuencia PAS

<400> 38

Ala Pro Ser Ser Pro Ser Pro Ser Ala Pro Ser Ser Pro Ser Pro Ala  
1 5 10 15

Ser Pro Ser Ser  
20

15 <210> 39  
<211> 19  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> secuencia PAS

<400> 39

25 Ala Pro Ser Ser Pro Ser Pro Ser Ala Pro Ser Ser Pro Ser Pro Ala  
1 5 10 15

Ser Pro Ser

30 <210> 40  
<211> 20  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> secuencia PAS

<400> 40

Ser Ser Pro Ser Ala Pro Ser Pro Ser Ser Pro Ala Ser Pro Ser Pro  
1 5 10 15

Ser Ser Pro Ala  
20

40 <210> 41  
<211> 24  
<212> PRT  
<213> Secuencia Artificial

45 <220>  
<223> secuencia PAS

<400> 41

Ala Ala Ser Pro Ala Ala Pro Ser Ala Pro Pro Ala Ala Ala Ser Pro  
1 5 10 15

Ala Ala Pro Ser Ala Pro Pro Ala  
20

5 <210> 42  
<211> 20  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> secuencia PAS

<400> 42

Ala Ser Ala Ala Ala Pro Ala Ala Ala Ser Ala Ala Ala Ser Ala Pro  
1 5 10 15

Ser Ala Ala Ala  
20

15 <210> 43  
<211> 23  
<212> ADN  
<213> Secuencia Artificial

20 <220>  
<223> diana miR142

25 <400> 43  
tccataaagt aggaaacact aca 23

30 <210> 44  
<211> 2332  
<212> PRT  
<213> Homo sapiens

# ES 2 926 585 T3

<400> 44

Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser Trp Asp Tyr  
1 5 10 15

Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro  
20 25 30

Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys  
35 40 45

Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile Ala Lys Pro  
50 55 60

Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu Val  
65 70 75 80

Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro Val  
85 90 95

Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly Ala  
100 105 110

Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys Val  
115 120 125

Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu Asn  
130 135 140

Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu Ser  
145 150 155 160

His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala Leu  
165 170 175

Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr Leu  
180 185 190

His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser Trp  
195 200 205



# ES 2 926 585 T3

His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp	Ala	Ala	Ser	210	215	220	
Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr	Val	Asn	Arg	225	230	235	240
Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val	Tyr	Trp	His	245	250	255	
Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile	Phe	Leu	Glu	260	265	270	
Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser	Leu	Glu	Ile	275	280	285	
Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met	Asp	Leu	Gly	290	295	300	
Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His	Asp	Gly	Met	305	310	315	320
Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro	Gln	Leu	Arg	325	330	335	
Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp	Leu	Thr	Asp	340	345	350	
Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	Pro	Ser	Phe	355	360	365	
Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	Trp	Val	His	370	375	380	
Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	Leu	Val	Leu	385	390	395	400
Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	Asn	Gly	Pro	405	410	415	
Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	Ala	Tyr	Thr	420	425	430	
Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu	Ser	Gly	Ile	435	440	445	
Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	Leu	Ile	Ile	450	455	460	

# ES 2 926 585 T3

Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro His Gly Ile  
465 470 475 480

Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys Gly Val Lys  
485 490 495

His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe Lys Tyr Lys  
500 505 510

Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp Pro Arg Cys  
515 520 525

Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg Asp Leu Ala  
530 535 540

Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu Ser Val Asp  
545 550 555 560

Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val Ile Leu Phe  
565 570 575

Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile Gln  
580 585 590

Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu Phe  
595 600 605

Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp Ser  
610 615 620

Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile Leu  
625 630 635 640

Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly Tyr  
645 650 655

Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe Pro  
660 665 670

Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu Trp  
675 680 685

Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr Ala  
690 695 700

Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr Glu  
705 710 715 720

# ES 2 926 585 T3

Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala  
 725 730 735  
 Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro Ser Thr Arg  
 740 745 750  
 Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp Ile Glu Lys  
 755 760 765  
 Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys Ile Gln Asn  
 770 775 780  
 Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser Pro Thr Pro  
 785 790 795 800  
 His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr Glu Thr Phe  
 805 810 815  
 Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn Ser Leu Ser  
 820 825 830  
 Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly Asp Met Val  
 835 840 845  
 Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu Lys Leu Gly  
 850 855 860  
 Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys Val Ser Ser  
 865 870 875 880  
 Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn Leu Ala Ala  
 885 890 895  
 Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met Pro Val His  
 900 905 910  
 Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys Ser Ser Pro  
 915 920 925  
 Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu Asn Asn Asp  
 930 935 940  
 Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu Ser Ser Trp  
 945 950 955 960  
 Gly Lys Asn Val Ser Ser Thr Glu Ser Gly Arg Leu Phe Lys Gly Lys

# ES 2 926 585 T3

965										970										975											
Arg	Ala	His	Gly	Pro	Ala	Leu	Leu	Thr	Lys	Asp	Asn	Ala	Leu	Phe	Lys																
			980						985						990																
Val	Ser	Ile	Ser	Leu	Leu	Lys	Thr	Asn	Lys	Thr	Ser	Asn	Asn	Ser	Ala																
		995					1000						1005																		
Thr	Asn	Arg	Lys	Thr	His	Ile	Asp	Gly	Pro	Ser	Leu	Leu	Ile	Glu																	
	1010					1015							1020																		
Asn	Ser	Pro	Ser	Val	Trp	Gln	Asn	Ile	Leu	Glu	Ser	Asp	Thr	Glu																	
	1025					1030							1035																		
Phe	Lys	Lys	Val	Thr	Pro	Leu	Ile	His	Asp	Arg	Met	Leu	Met	Asp																	
	1040					1045							1050																		
Lys	Asn	Ala	Thr	Ala	Leu	Arg	Leu	Asn	His	Met	Ser	Asn	Lys	Thr																	
	1055					1060							1065																		
Thr	Ser	Ser	Lys	Asn	Met	Glu	Met	Val	Gln	Gln	Lys	Lys	Glu	Gly																	
	1070					1075							1080																		
Pro	Ile	Pro	Pro	Asp	Ala	Gln	Asn	Pro	Asp	Met	Ser	Phe	Phe	Lys																	
	1085					1090							1095																		
Met	Leu	Phe	Leu	Pro	Glu	Ser	Ala	Arg	Trp	Ile	Gln	Arg	Thr	His																	
	1100					1105							1110																		
Gly	Lys	Asn	Ser	Leu	Asn	Ser	Gly	Gln	Gly	Pro	Ser	Pro	Lys	Gln																	
	1115					1120							1125																		
Leu	Val	Ser	Leu	Gly	Pro	Glu	Lys	Ser	Val	Glu	Gly	Gln	Asn	Phe																	
	1130					1135							1140																		
Leu	Ser	Glu	Lys	Asn	Lys	Val	Val	Val	Gly	Lys	Gly	Glu	Phe	Thr																	
	1145					1150							1155																		
Lys	Asp	Val	Gly	Leu	Lys	Glu	Met	Val	Phe	Pro	Ser	Ser	Arg	Asn																	
	1160					1165							1170																		
Leu	Phe	Leu	Thr	Asn	Leu	Asp	Asn	Leu	His	Glu	Asn	Asn	Thr	His																	
	1175					1180							1185																		
Asn	Gln	Glu	Lys	Lys	Ile	Gln	Glu	Glu	Ile	Glu	Lys	Lys	Glu	Thr																	
	1190					1195							1200																		

# ES 2 926 585 T3

Leu Ile	Gln Glu Asn Val	Val	Leu Pro Gln Ile	His	Thr Val Thr
1205		1210		1215	
Gly Thr	Lys Asn Phe Met	Lys	Asn Leu Phe Leu	Leu	Ser Thr Arg
1220		1225		1230	
Gln Asn	Val Glu Gly Ser	Tyr	Asp Gly Ala Tyr	Ala	Pro Val Leu
1235		1240		1245	
Gln Asp	Phe Arg Ser Leu	Asn	Asp Ser Thr Asn	Arg	Thr Lys Lys
1250		1255		1260	
His Thr	Ala His Phe Ser	Lys	Lys Gly Glu Glu	Glu	Asn Leu Glu
1265		1270		1275	
Gly Leu	Gly Asn Gln Thr	Lys	Gln Ile Val Glu	Lys	Tyr Ala Cys
1280		1285		1290	
Thr Thr	Arg Ile Ser Pro	Asn	Thr Ser Gln Gln	Asn	Phe Val Thr
1295		1300		1305	
Gln Arg	Ser Lys Arg Ala	Leu	Lys Gln Phe Arg	Leu	Pro Leu Glu
1310		1315		1320	
Glu Thr	Glu Leu Glu Lys	Arg	Ile Ile Val Asp	Asp	Thr Ser Thr
1325		1330		1335	
Gln Trp	Ser Lys Asn Met	Lys	His Leu Thr Pro	Ser	Thr Leu Thr
1340		1345		1350	
Gln Ile	Asp Tyr Asn Glu	Lys	Glu Lys Gly Ala	Ile	Thr Gln Ser
1355		1360		1365	
Pro Leu	Ser Asp Cys Leu	Thr	Arg Ser His Ser	Ile	Pro Gln Ala
1370		1375		1380	
Asn Arg	Ser Pro Leu Pro	Ile	Ala Lys Val Ser	Ser	Phe Pro Ser
1385		1390		1395	
Ile Arg	Pro Ile Tyr Leu	Thr	Arg Val Leu Phe	Gln	Asp Asn Ser
1400		1405		1410	
Ser His	Leu Pro Ala Ala	Ser	Tyr Arg Lys Lys	Asp	Ser Gly Val
1415		1420		1425	
Gln Glu	Ser Ser His Phe	Leu	Gln Gly Ala Lys	Lys	Asn Asn Leu
1430		1435		1440	

# ES 2 926 585 T3

Ser	Leu	Ala	Ile	Leu	Thr	Leu	Glu	Met	Thr	Gly	Asp	Gln	Arg	Glu
1445						1450					1455			
Val	Gly	Ser	Leu	Gly	Thr	Ser	Ala	Thr	Asn	Ser	Val	Thr	Tyr	Lys
1460						1465					1470			
Lys	Val	Glu	Asn	Thr	Val	Leu	Pro	Lys	Pro	Asp	Leu	Pro	Lys	Thr
1475						1480					1485			
Ser	Gly	Lys	Val	Glu	Leu	Leu	Pro	Lys	Val	His	Ile	Tyr	Gln	Lys
1490						1495					1500			
Asp	Leu	Phe	Pro	Thr	Glu	Thr	Ser	Asn	Gly	Ser	Pro	Gly	His	Leu
1505						1510					1515			
Asp	Leu	Val	Glu	Gly	Ser	Leu	Leu	Gln	Gly	Thr	Glu	Gly	Ala	Ile
1520						1525					1530			
Lys	Trp	Asn	Glu	Ala	Asn	Arg	Pro	Gly	Lys	Val	Pro	Phe	Leu	Arg
1535						1540					1545			
Val	Ala	Thr	Glu	Ser	Ser	Ala	Lys	Thr	Pro	Ser	Lys	Leu	Leu	Asp
1550						1555					1560			
Pro	Leu	Ala	Trp	Asp	Asn	His	Tyr	Gly	Thr	Gln	Ile	Pro	Lys	Glu
1565						1570					1575			
Glu	Trp	Lys	Ser	Gln	Glu	Lys	Ser	Pro	Glu	Lys	Thr	Ala	Phe	Lys
1580						1585					1590			
Lys	Lys	Asp	Thr	Ile	Leu	Ser	Leu	Asn	Ala	Cys	Glu	Ser	Asn	His
1595						1600					1605			
Ala	Ile	Ala	Ala	Ile	Asn	Glu	Gly	Gln	Asn	Lys	Pro	Glu	Ile	Glu
1610						1615					1620			
Val	Thr	Trp	Ala	Lys	Gln	Gly	Arg	Thr	Glu	Arg	Leu	Cys	Ser	Gln
1625						1630					1635			
Asn	Pro	Pro	Val	Leu	Lys	Arg	His	Gln	Arg	Glu	Ile	Thr	Arg	Thr
1640						1645					1650			
Thr	Leu	Gln	Ser	Asp	Gln	Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile
1655						1660					1665			
Ser	Val	Glu	Met	Lys	Lys	Glu	Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp
1670						1675					1680			

# ES 2 926 585 T3

Glu Asn	Gln Ser Pro Arg Ser	Phe Gln Lys Lys Thr	Arg His Tyr
1685	1690	1695	
Phe Ile	Ala Ala Val Glu Arg	Leu Trp Asp Tyr Gly	Met Ser Ser
1700	1705	1710	
Ser Pro	His Val Leu Arg Asn	Arg Ala Gln Ser Gly	Ser Val Pro
1715	1720	1725	
Gln Phe	Lys Lys Val Val Phe	Gln Glu Phe Thr Asp	Gly Ser Phe
1730	1735	1740	
Thr Gln	Pro Leu Tyr Arg Gly	Glu Leu Asn Glu His	Leu Gly Leu
1745	1750	1755	
Leu Gly	Pro Tyr Ile Arg Ala	Glu Val Glu Asp Asn	Ile Met Val
1760	1765	1770	
Thr Phe	Arg Asn Gln Ala Ser	Arg Pro Tyr Ser Phe	Tyr Ser Ser
1775	1780	1785	
Leu Ile	Ser Tyr Glu Glu Asp	Gln Arg Gln Gly Ala	Glu Pro Arg
1790	1795	1800	
Lys Asn	Phe Val Lys Pro Asn	Glu Thr Lys Thr Tyr	Phe Trp Lys
1805	1810	1815	
Val Gln	His His Met Ala Pro	Thr Lys Asp Glu Phe	Asp Cys Lys
1820	1825	1830	
Ala Trp	Ala Tyr Phe Ser Asp	Val Asp Leu Glu Lys	Asp Val His
1835	1840	1845	
Ser Gly	Leu Ile Gly Pro Leu	Leu Val Cys His Thr	Asn Thr Leu
1850	1855	1860	
Asn Pro	Ala His Gly Arg Gln	Val Thr Val Gln Glu	Phe Ala Leu
1865	1870	1875	
Phe Phe	Thr Ile Phe Asp Glu	Thr Lys Ser Trp Tyr	Phe Thr Glu
1880	1885	1890	
Asn Met	Glu Arg Asn Cys Arg	Ala Pro Cys Asn Ile	Gln Met Glu
1895	1900	1905	
Asp Pro	Thr Phe Lys Glu Asn	Tyr Arg Phe His Ala	Ile Asn Gly

# ES 2 926 585 T3

1910	1915	1920
Tyr Ile Met Asp Thr Leu Pro 1925	Gly Leu Val Met Ala Gln Asp Gln 1930	
Arg Ile Arg Trp Tyr Leu Leu 1940	Ser Met Gly Ser Asn Glu Asn Ile 1945	
His Ser Ile His Phe Ser Gly 1955	His Val Phe Thr Val Arg Lys Lys 1960	
Glu Glu Tyr Lys Met Ala Leu 1970	Tyr Asn Leu Tyr Pro Gly Val Phe 1980	
Glu Thr Val Glu Met Leu Pro 1985	Ser Lys Ala Gly Ile Trp Arg Val 1990	
Glu Cys Leu Ile Gly Glu His 2000	Leu His Ala Gly Met Ser Thr Leu 2005	
Phe Leu Val Tyr Ser Asn Lys 2015	Cys Gln Thr Pro Leu Gly Met Ala 2020	
Ser Gly His Ile Arg Asp Phe 2030	Gln Ile Thr Ala Ser Gly Gln Tyr 2035	
Gly Gln Trp Ala Pro Lys Leu 2045	Ala Arg Leu His Tyr Ser Gly Ser 2050	
Ile Asn Ala Trp Ser Thr Lys 2060	Glu Pro Phe Ser Trp Ile Lys Val 2065	
Asp Leu Leu Ala Pro Met Ile 2075	Ile His Gly Ile Lys Thr Gln Gly 2080	
Ala Arg Gln Lys Phe Ser Ser 2090	Leu Tyr Ile Ser Gln Phe Ile Ile 2095	
Met Tyr Ser Leu Asp Gly Lys 2105	Lys Trp Gln Thr Tyr Arg Gly Asn 2110	
Ser Thr Gly Thr Leu Met Val 2120	Phe Phe Gly Asn Val Asp Ser Ser 2125	
Gly Ile Lys His Asn Ile Phe 2135	Asn Pro Pro Ile Ile Ala Arg Tyr 2140	



# ES 2 926 585 T3

Ile	Arg	Leu	His	Pro	Thr	His	Tyr	Ser	Ile	Arg	Ser	Thr	Leu	Arg
2150						2155					2160			
Met	Glu	Leu	Met	Gly	Cys	Asp	Leu	Asn	Ser	Cys	Ser	Met	Pro	Leu
2165						2170					2175			
Gly	Met	Glu	Ser	Lys	Ala	Ile	Ser	Asp	Ala	Gln	Ile	Thr	Ala	Ser
2180						2185					2190			
Ser	Tyr	Phe	Thr	Asn	Met	Phe	Ala	Thr	Trp	Ser	Pro	Ser	Lys	Ala
2195						2200					2205			
Arg	Leu	His	Leu	Gln	Gly	Arg	Ser	Asn	Ala	Trp	Arg	Pro	Gln	Val
2210						2215					2220			
Asn	Asn	Pro	Lys	Glu	Trp	Leu	Gln	Val	Asp	Phe	Gln	Lys	Thr	Met
2225						2230					2235			
Lys	Val	Thr	Gly	Val	Thr	Thr	Gln	Gly	Val	Lys	Ser	Leu	Leu	Thr
2240						2245					2250			
Ser	Met	Tyr	Val	Lys	Glu	Phe	Leu	Ile	Ser	Ser	Ser	Gln	Asp	Gly
2255						2260					2265			
His	Gln	Trp	Thr	Leu	Phe	Phe	Gln	Asn	Gly	Lys	Val	Lys	Val	Phe
2270						2275					2280			
Gln	Gly	Asn	Gln	Asp	Ser	Phe	Thr	Pro	Val	Val	Asn	Ser	Leu	Asp
2285						2290					2295			
Pro	Pro	Leu	Leu	Thr	Arg	Tyr	Leu	Arg	Ile	His	Pro	Gln	Ser	Trp
2300						2305					2310			
Val	His	Gln	Ile	Ala	Leu	Arg	Met	Glu	Val	Leu	Gly	Cys	Glu	Ala
2315						2320					2325			

Gln Asp Leu Tyr  
2330

<210> 45

<211> 4

5 <212> PRT

<213> Secuencia Artificial

<220>

10 <223> aminoácidos 233-236 de IgG1 humana

<400> 45

Glu Leu Leu Gly

15 1

<210> 46

<211> 42

<212> PRT

<213> Secuencia Artificial

20

<220>

<223> XTEN AE42-4, secuencia de proteína

<400> 46

5

Gly Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly  
1 5 10 15

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala  
20 25 30

Thr Ser Gly Ser Glu Thr Pro Ala Ser Ser  
35 40

<210> 47

<211> 126

10

<212> ADN

<213> Secuencia Artificial

<220>

<223> XTEN AE42-4, secuencia de ADN

15

<400> 47

ggcgcgccag gttctcctgc tggctcccc acctcaacag aagaggggac aagcgaaagc 60

gctacgcctg agagtggccc tggctctgag ccagccacct ccggctctga aaccctgcc 120

tcgagc 126

20

<210> 48

<211> 144

<212> PRT

<213> Secuencia Artificial

25

<220>

<223> XTEN AE144-2A, secuencia de proteína

<400> 48

Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly  
1 5 10 15

Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly  
20 25 30

Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly  
35 40 45

30

# ES 2 926 585 T3

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu  
50 55 60

Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu  
65 70 75 80

Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly  
85 90 95

Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu  
100 105 110

Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu  
115 120 125

Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly  
130 135 140

<210> 49

<211> 450

<212> ADN

<213> Secuencia Artificial

<220>

<223> XTEN AE144-2A, secuencia de ADN

<400> 49

ggcgcgccaa ccagtacgga gccgtccgag gggagcgcac caggaagccc ggctgggagc 60

ccgacttcta ccgaagaggg tacatctacc gaaccaagtg aagggttcagc accaggcacc 120

tcaacagaac cctctgaggg ctcgggcgct ggtacaagtg agtccgccac ccagaaatcc 180

gggcctggga caagcacaga accttcggaa gggagtgcc ctggaacatc cgaatcggca 240

accccagaat cagggccagg atctgagccc gcgacttcgg gctccgagac gcctgggaca 300

tccaccgagc cctccgaagg atcagcccca ggcaccagca cggagccctc tgagggaagc 360

gcacctggta ccagcgaaag cgcaactccc gaatcaggtc ccggtacgag cgagtcggcg 420

accccggaga gcgggcccagg tgcctcgagc 450

<210> 50

<211> 144

<212> PRT

<213> Secuencia Artificial

<220>

<223> XTEN AE144-3B, secuencia de proteína

<400> 50

Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser

1 5 10 15

Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser  
20 25 30

Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly  
35 40 45

Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu  
50 55 60

Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly  
65 70 75 80

Ser Ala Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly  
85 90 95

Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Thr Glu  
100 105 110

Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser  
115 120 125

Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly  
130 135 140

5 <210> 51  
<211> 450  
<212> ADN  
<213> Secuencia Artificial

10 <220>  
<223> XTEN AE144-3B, secuencia de ADN  
<400> 51

ggcgcgccaa gtcgcgctgg aagcccaact agcaccgaag aggggacctc agagtccgcc	60
acccccgagt ccggccctgg ctctgagcct gccactagcg gctccgagac tcctggcaca	120
tccgaaagcg ctacaccgga gagtggaccc ggcacctcta ccgagcccag tgagggtctc	180
gcccctggaa caagcaccga gccagcgaa ggcagcgccc caggacctc cacagagccc	240
agtgaaggca gtgctcctgg caccagcacc gaaccaagcg agggctctgc acccgggacc	300
tccaccgagc caagcgaagg ctctgcccct ggcacttcca ccgagcccag cgaaggcagc	360
gcccctggga gcccgcctgg ctctcccacc agcactgagg agggcacatc taccgaacca	420
agtgaaggct ctgcaccagg tgcctcgagc	450

15 <210> 52  
<211> 144  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> XTEN AE144-4A, secuencia de proteína  
<400> 52

# ES 2 926 585 T3

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala  
1 5 10 15

Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu  
20 25 30

Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly  
35 40 45

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu  
50 55 60

Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu  
65 70 75 80

Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly  
85 90 95

Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ser Pro Ala Gly  
100 105 110

Ser Pro Thr Ser Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu  
115 120 125

Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly  
130 135 140

<210> 53

5 <211> 450

<212> ADN

<213> Secuencia Artificial

<220>

10 <223> XTEN AE144-4A, secuencia de ADN

<400> 53

ggcgcgccaa cgtccgaaag tgctaccct gagtcaggcc ctggtagtga gcctgccaca 60

agcggaaagc aaactccggg gacctcagag tctgccactc ccgaatcggg gccaggetct 120

gaaccggcca cttcaggagg cgaaacacca ggaacatcgg agagcgctac cccggagagc 180

gggccaggaa ctagtactga gcctagcgag ggaagtgcac ctggtacaag cgagtccgcc 240

acacccgagt ctggccctgg ctctccagcg ggctcaccca cgagcactga agagggetct 300

cccgtggca gcccaacgtc gacagaagaa ggatcaccag caggctcccc cacatcaaca 360

15 gaggagggtg catcagaatc tgctactccc gagagtggac ccggtacctc cactgagccc 420

agcgagggga gtgcaccagg tgccctcgagc 450

<210> 54

<211> 144

20 <212> PRT

<213> Secuencia Artificial

<220>

<223> XTEN AE144-5A, secuencia de proteína

# ES 2 926 585 T3

<400> 54

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala  
1 5 10 15

Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu  
20 25 30

Ser Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly  
35 40 45

Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu  
50 55 60

Pro Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser  
65 70 75 80

Thr Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly  
85 90 95

Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser  
100 105 110

Ala Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser  
115 120 125

Thr Glu Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly  
130 135 140

5

<210> 55

<211> 450

<212> ADN

<213> Secuencia Artificial

10

<220>

<223> XTEN AE144-5A, secuencia de ADN

<400> 55

15

ggcgcgccaa catcagagag cgccaccct gaaagtggc ccgggagcga gccagccaca 60

tctgggtcgg aaacgccagg cacaagtgag tctgcaactc ccgagtcgag acctggctcc 120

gagcctgcca ctacgggctc cgagactccg ggaacttccg agagcgctac accagaaagc 180

ggaccgggaa ccagtaccga acctagcgag ggctctgctc cgggcagccc agccggctct 240

cctacatcca cggaggagg cacttccgaa tccgccacc cggagtcagg gccaggatct 300

gaaccgcta cctcaggcag tgagacgcca ggaacgagcg agtcgctac accggagagt 360

gggccaggga gccctgctgg atctcctacg tccactgagg aagggtcacc agcgggctcg 420

cccaccagca ctgaagaagg tgccctcgagc 450

20

<210> 56

<211> 144

<212> PRT

<213> Secuencia Artificial

<220>

<223> XTEN AE144-6B, secuencia de proteína

<400> 56

5

```

Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Thr Ser Glu Ser
1          5          10          15

Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu
          20          25          30

Ser Gly Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly
          35          40          45

Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Ser Glu Pro Ala
          50          55          60

Thr Ser Gly Ser Glu Thr Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser
65          70          75          80

Thr Glu Glu Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly
          85          90          95

Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly Ser Glu Pro Ala
          100          105          110

Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu
          115          120          125

Ser Gly Pro Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro Gly
          130          135          140

```

<210> 57

<211> 450

10

<212> ADN

<213> Secuencia Artificial

<220>

<223> XTEN AE144-6B, secuencia de ADN

15

<400> 57

```

ggcgcgccaa catctaccga gccttcgaa ggctctgcc ctgggacctc agaatctgca      60
accctgaaa gcggccctgg aacctccgaa agtgccactc ccgagagcgg ccagggaca      120
agcgagtcag caaccctga gtctggaccc ggcagcgagc ctgcaacctc tggctcagag      180
actcccggct cagaaccgc tacctcaggc tccgagacac ccggctctcc tgctgggagt      240
cccacttcca ccgaggaagg aacatccact gagcctagtg agggctctgc ccctggaacc      300
agcacagagc caagtgaggg cagtgcacca ggatccgagc cagcaaccag cgggtccgag      360
actccgggga cctctgagtc tgccacccca gagagcggac ccggcacttc aaccgagccc      420
tccgaaggat cagcaccagg tgcctcgagc                                450

```

20

<210> 58

<211> 144

<212> PRT

<213> Secuencia Artificial

<220>

<223> XTEN AG144-1, secuencia de proteína

5 <400> 58

Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser  
1 5 10 15

Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Thr Pro Gly Ser Gly Thr  
20 25 30

Ala Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser  
35 40 45

Pro Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Ala Ser  
50 55 60

Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr  
65 70 75 80

Ala Ser Ser Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser  
85 90 95

Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser  
100 105 110

Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser  
115 120 125

Ser Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser  
130 135 140

10

<210> 59

<211> 450

<212> ADN

<213> Secuencia Artificial

15

<220>

<223> XTEN AG144-1, secuencia de ADN

<400> 59

20

ggcgcgccac ccgggtcgtc cccgtcggcg tccaccggaa cagggccagg gtcacccccg 60

tcagcgtcga ctgggacggg acccgggaca cccggttcgg ggactgcatc ctctcgcct 120

ggttcgtcca ccccgtcagg agccacgggt tcgcccggaa gcagcccaag cgcacccact 180

ggtacagggc ctggggcttc accgggtact tcatccacgg ggtcaccggg aacgcccgga 240

tcggggacgg ctctctcatc accaggatcg tcaacaccct cgggcgcaac gggcagcccc 300

ggaacccctg gttcgggtac ggcgtcgtcg agccccggtg cgagcccggg aacaagctcg 360

acaggatcgc ctggggcgtc acccggcacg tcgagcacag gcagccccgg aaccctgga 420

tcgggaaccg cgtcgtcaag cgcctcgagc 450

<210> 60

<211> 144

<212> PRT

<213> Secuencia Artificial

25



<220>

<223> XTEN AG144-A, secuencia de proteína

5 <400> 60

Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ser Ser Pro  
1 5 10 15

Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser Thr  
20 25 30

Gly Thr Gly Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro  
35 40 45

Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Pro  
50 55 60

Ser Ala Ser Thr Gly Thr Gly Pro Gly Ala Ser Pro Gly Thr Ser Ser  
65 70 75 80

Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro  
85 90 95

Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Thr Pro Gly  
100 105 110

Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser  
115 120 125

Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro  
130 135 140

10

<210> 61

<211> 450

<212> ADN

<213> Secuencia Artificial

15

<220>

<223> XTEN AG144-A, secuencia de ADN

<400> 61

20

ggcgcgccag gtgcctcgcc gggaacatca tcaactggtt caccgcgggtc atccccctcg	60
gcctcaaccg ggacgggtcc cggctcatcc cccagcgcca gcaactggaac aggtcctggc	120
actcctggtt ccggtacggc atcgctcatcc ccgggaagct caacaccgtc cggagcgaca	180
ggatcacctg gctcgtcacc ttcggcgta actggaacgg ggccaggggc ctcacccgga	240
acgtcctcga ctgggtcgcc tggtagccg ggatcaggaa cggectcatc ctgcctggg	300
tcctcaacgc cctcgggtgc gactggttcg ccgggaactc ctggctcggg gacggcctcg	360
tcgtcgctg gggcatcacc ggggacgagc tccacggggt cccctggagc gtcaccgggg	420
acctcctcga caggtagccc ggctcagac	450

<210> 62

<211> 144

<212> PRT  
 <213> Secuencia Artificial

<220>

5 <223> XTEN AG144-B, secuencia de proteína

<400> 62

Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ser Ser Thr  
 1 5 10 15

Pro Ser Gly Ala Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser  
 20 25 30

Thr Gly Ser Pro Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro  
 35 40 45

10

Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Pro  
 50 55 60

Ser Ala Ser Thr Gly Thr Gly Pro Gly Ser Ser Pro Ser Ala Ser Thr  
 65 70 75 80

Gly Thr Gly Pro Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro  
 85 90 95

Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro Gly Ala Ser Pro  
 100 105 110

Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser  
 115 120 125

Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro  
 130 135 140

<210> 63

<211> 450

15

<212> ADN

<213> Secuencia Artificial

<220>

<223> XTEN AG144-B, secuencia de ADN

20

<400> 63

ggcgcgccag gtacaccggg cagcggcacg gcttcgtcgt caccgcggctc gtccacaccg 60

tccgggagcta cgggaagccc aggagcgtca ccgggaacgt cgtcaacggg gtcaccgggt 120

acgccaggta gcggcacggc cagcagctcg ccaggttcat cgaccccgctc gggagcgact 180

gggtcgcccg gatcaagccc gtcagcttcc actggaacag gaccgcgggtc gtcgcccgtca 240

gcctcaacgg ggacaggacc tggttcatcg acgcgctcag gggcgacagg ctgcccggga 300

tcgtcaaacac cctcgggggc aacggggagc cctgggtgct cgcttggaac ctcattccacc 360

ggaagcccg gggcctcgcc gggtacgagc tccacgggat cgcccgagc gtcccccgga 420

acttcaagca caggagagccc tgcctcgagc 450

25

<210> 64

<211> 144

<212> PRT  
<213> Secuencia Artificial

<220>

5 <223> XTEN AG144-C, secuencia de proteína

<400> 64

```

Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro
1           5           10           15

Gly Thr Ser Ser Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser
20           25           30

Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro
35           40           45

Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro Gly Thr Pro Gly
50           55           60

Ser Gly Thr Ala Ser Ser Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser
65           70           75           80

Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro
85           90           95

Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro Gly Ser Ser Thr
100          105          110

Pro Ser Gly Ala Thr Gly Ser Pro Gly Ser Ser Thr Pro Ser Gly Ala
115          120          125

Thr Gly Ser Pro Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro
130          135          140

```

<210> 65

<211> 450

15 <212> ADN

<213> Secuencia Artificial

<220>

20 <223> XTEN AG144-C, secuencia de ADN

<400> 65

```

ggcgcgccag gtacacccgg atcgggtaca gcgtcatcga gcccgggtgc gtcacctggt      60
acgtcgagca cggggctgcc agggcgctcc cctgggacgt cctcaacagg ctgccccggt      120
gcgtcaccgg gcaactcggt caagggttca cctggtagct ccccttcgcg gtccactggc      180
accgggcctg gaactccggg gagcggcaca gcgagctcgt cgccgggagc atgcctggg      240
acatcgagca ccgggtcgcc aggagcatcg cccggaacat ccagcacagg aagccccggc      300
gcgtcgcccc ggacatcaag cacaggttcc ccgggatcga gcacgccgtc cggagccact      360
ggatcaccag ggagctcgac accttcgggc gcaacgggat cgcccgagc cagccccggg      420
acgtcaagca ctggctcccc tgctcgagc                                450

```

<210> 66  
 <211> 144  
 <212> PRT  
 <213> Secuencia Artificial

5

<220>  
 <223> XTEN AG144-F, secuencia de proteína

<400> 66

10

Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ser	Ser	Pro	
1				5				10					15			
Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	
			20				25					30				
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	
		35				40					45					
Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro	
	50					55					60					
Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	
65					70				75					80		
Thr	Gly	Ser	Pro	Gly	Ser	Ser	Pro	Ser	Ala	Ser	Thr	Gly	Thr	Gly	Pro	
				85				90						95		
Gly	Thr	Pro	Gly	Ser	Gly	Thr	Ala	Ser	Ser	Ser	Pro	Gly	Ser	Ser	Thr	
		100					105					110				
Pro	Ser	Gly	Ala	Thr	Gly	Ser	Pro	Gly	Ser	Ser	Thr	Pro	Ser	Gly	Ala	
		115					120					125				
Thr	Gly	Ser	Pro	Gly	Ala	Ser	Pro	Gly	Thr	Ser	Ser	Thr	Gly	Ser	Pro	
	130					135					140					

<210> 67  
 <211> 450  
 <212> ADN  
 <213> Secuencia Artificial

15

<220>  
 <223> XTEN AG144-F, secuencia de ADN

<400> 67

20

ggcgcgccag gctccagccc ctccgcgagc acgggaaccg gaccagggtc gtcaccctca	60
gcatcaacgg ggacgggacc gggggcgta ccaggaacgt cctccaccgg ctgcgcgggt	120
gcatcaccgg gaacgtcatc gaccggatcg ccaggagct cgacgccatc aggcgcaaca	180
ggatcacctg gctcaagccc tagcgcgtca accggcacgg gtccgggtgc ctcccctggc	240
acgtccagca ccggatcacc cggatcgagc ccatccgcct caaccggaac cggaccgggt	300
acaccagggt cggaacagc ctctctgtca ccaggctcct caaccctctc gggagccacg	360

	ggttcgcccg gttcgtcaac gccttcgga gcaactggta gccccggagc atcgccagga	420
	acttcgagca cggggtcgcc cgcctcgagc	450
5	<210> 68 <211> 4374 <212> ADN <213> Secuencia Artificial	
10	<220> <223> coFVIII-1 Secuencia de ADN	
	<400> 68	
	atgcagattg agctgtctac ttgctttttc ctgtgcctgc tgaggttttg cttttccgct	60
	acacgaaggt attatctggg ggctgtggaa ctgtcttggg attacatgca gagtgcctg	120
	ggagagctgc cagtggacgc aaggtttccc cctagagtcc ctaagtcatt ccccttcaac	180
	actagcgtgg tctacaagaa aacactgttc gtggagttaa ctgatcacct gttcaacatc	240
	gcaaagccta ggccaccctg gatgggactg ctggggccaa caatccaggc cgagggtgtac	300
	gacaccgtgg tcattacact taagaacatg gcctcacacc ccgtgagcct gcatgctgtg	360
	ggcgtcagct actggaaggc ttccgaagga gcagagtatg acgatcagac ttcccagaga	420
	gaaaaagagg acgataaggt gtttcctggc ggatctcata cctacgtgtg gcaggtcctg	480
	aaagagaatg gccctatggc ctccgaccct ctgtgcctga cctactctta tctgagtcac	540
	gtggacctgg tcaaggatct gaacagcggc ctgatcggag ccctgctggt gtgcagggaa	600
	ggaagcctgg ctaaggagaa aaccagaca ctgcataagt tcattctgct gttcgccgtg	660
	tttgacgaag ggaaatcatg gcacagcgag acaaagaata gtctgatgca ggacagggat	720
	gccgcttcag ccagagcttg gcccaaatg cactctgtga acggctacgt caatcgctca	780
	ctgcctgggc tgatcggctg ccaccgaaag agcgtgtatt ggcatgtcat cgggatgggc	840
	accacacctg aagtgcactc cttttcctg gagggacata cttttctggt ccgcaaccac	900
	cgacaggctt cctggagat ctctccaatt accttcctga cagcacagac tctgctgatg	960
	gacctggggc agttcctgct gttttgccac atcagctccc accagcatga tggcatggag	1020
	gcttacgtga aagtggactc ttgtcccgag gaacctcagc tgcggatgaa gaacaatgag	1080
	gaagcagaag actatgacga tgacctgacc gactccgaga tggatgtggt ccgattcgat	1140
	gacgataaca gcccctcctt tatccagatt agatctgtgg ccaagaaaca ccctaagaca	1200
	tgggtccatt acatcgagc cgaggaagag gactgggatt atgcaccact ggtgctggca	1260
	ccagacgacg gctcctacaa atctcagtat ctgaacaatg ggccacagag gattggcaga	1320
	aagtacaaga aagtgcggtt catggcatat accgatgaga ctttcaagac tcgcgaagcc	1380
	atccagcacg agagcggcat cctgggacca ctgctgtacg gagaagtggg agacaccctg	1440
	ctgatcattt tcaagaacca ggccagccgg ccttacaata tctatccaca tgggattaca	1500

15

# ES 2 926 585 T3

gatgtgcgcc ctctgtacag caggagactg ccaaagggcg taaaacacct gaaggacttc	1560
ccaatcctgc ccggagaaat cttcaagtac aagtggactg tcaccgtcga ggatggcccc	1620
actaagagcg accctcgggtg cctgaccgc tactattcta gtttcgtgaa tatggaaaga	1680
gatctggcaa gcggactgat cggaccactg ctgatttggtt acaaagagag cgtggatcag	1740
agaggcaacc agatcatgtc cgacaagcgg aatgtgattc tgttcagtgt ctttgacgaa	1800
aacaggctcat ggtacctgac cgagaacatc cagagattcc tgcctaatacc agctgggggtg	1860
cagctggaag atcctgagtt tcaggcatct aacatcatgc atagtattaa tggctacgtg	1920
ttcgacagtt tgcagctgag cgtgtgcctg cagaggtcg cttactggta taccctgagc	1980
attggggcac agacagattt cctgagcgtg ttcttttccg gctacacttt taagcataaa	2040
atggtctatg aggacacact gactctgttc cccttcagcg gcgaaaccgt gtttatgagc	2100
atggagaatc ccggactgtg gattctgggg tgccacaaca gcgatttcag aaatcgcgga	2160
atgactgccc tgctgaaagt gtcaagctgt gacaagaaca ccggggacta ctatgaagat	2220
tcatacagag acatcagcgc atatctgctg tccaaaaaca atgccattga accccggtct	2280
tttagtcaga atcctccagt gctgaagagg caccagaggg agatcacccg cactaccctg	2340
cagagtgatc aggaagagat cgactacgac gatacaattt ctgtggaaat gaagaaagag	2400
gacttcgata tctatgacga agatgagaac cagagtcctc gatcattcca gaagaaaacc	2460
aggcattact ttattgccgc agtggagcgg ctgtgggatt atggcatgtc ctctagtcct	2520
cacgtgctgc gaaatagggc ccagtcagga agcgtccac agttcaagaa agtggctctc	2580
caggagttta cagacgggtc ctttactcag ccactgtaca ggggcgaact gaacgagcac	2640
ctgggactgc tggggcccta tatcagagca gaagtggagg ataacattat ggtcaccttc	2700
agaaatcagg cctctcggcc ttacagtttt tattcaagcc tgatctctta cgaagaggac	2760
cagcgacagg gagctgaacc acgaaaaaac ttcgtgaagc ctaatgagac caaacatac	2820
ttttggaagg tgcagacca tatggcccca acaaaagacg agttcgattg caaggcatgg	2880
gcctattttt ctgacgtgga tctggagaag gacgtgcaca gtggcctgat tggccactg	2940
ctggtgtgcc atactaacac cctgaatcca gccacggcc gccaggtcac tgtccaggag	3000
ttcgctctgt tctttaccat ctttgatgag acaaagagct ggtacttcac cgaaaacatg	3060
gagcgaaatt gcagggtcc atgtaacatt cagatggaag accccacatt caaggagaac	3120
taccgctttc atgctatcaa tggatacatc atggatactc tgcccggtct ggtcatggca	3180
caggaccaga gaatccggtg gtatctgctg agcatgggca gcaacgagaa tatccactca	3240
attcatttca gcgggcacgt gtttactgtc aggaagaaag aagagtacaa gatggccctg	3300
tacaacctgt atcccgcggt gttcgaaacc gtcgagatgc tgcctagcaa ggccggaatc	3360

tggagagtgg aatgcctgat tggagagcac ctgcatgctg ggatgtctac cctgtttctg 3420  
 gtgtacagta ataagtgtca gacaccctg ggaatggcat ccgggcatat cagggatttc 3480  
 cagattaccg catctggaca gtacggacag tgggacaccta agctggctag actgcactat 3540  
 tccggatcta tcaacgcttg gtccacaaaa gaggcctttct cttggattaa ggtggacctg 3600  
 ctggcccca tgatcattca tggcatcaaa actcagggag ctccggcagaa gttctcctct 3660  
 ctgtacatct cacagtttat catcatgtac agcctggatg ggaagaaatg gcagacatac 3720  
 cgcggaata gcacaggaac tctgatggtg ttctttggca acgtggacag cagcggaatc 3780  
 aagcacaaca ttttcaatcc ccctatcatt gctagataca tccggctgca cccaacccat 3840  
 tattctattc gaagtacact gaggatggaa ctgatgggat gcgatctgaa cagttgttca 3900  
 atgcccctgg ggatggagtc caaggcaatc tctgacgcc agattaccgc cagctcctac 3960  
 ttcactaata tgtttgctac ctggagccct tccaaagcaa gactgcacct gcaaggccgc 4020  
 agcaacgcat ggcgaccaca ggtgaacaat cccaaggagt ggttgcaggt cgattttcag 4080  
 aaaactatga aggtgaccgg ggtcacaact cagggcgtga aaagtctgct gacctcaatg 4140  
 tacgtcaagg agttcctgat ctctagttca caggacggac atcagtgagc actgttcttt 4200  
 cagaacggga aggtgaaagt cttccagggc aatcaggatt cctttacacc tgttgtcaac 4260  
 agtctagacc ctccactgct gaccagatac ctgagaatcc accctcagtc ctgggtgcac 4320  
 cagattgccc tgagaatgga agtgctggga tgcgagggcc aggatctgta ctga 4374

<210> 69

<211> 577

5 <212> ADN

<213> Secuencia Artificial

<220>

10 <223> Promotor ET, secuencia de ADN

<400> 69

ctcgaggtca attcacgcga gttaataatt accagcgcg gccaataaaa taatccgcga 60  
 ggggcaggtg acgtttgcc agcgcgcgct ggtaattatt aacctcgca atattgattc 120  
 gaggccgcga ttgccgcaat cgcgaggggc aggtgacctt tgcccagcgc gcgttcgccc 180  
 cgccccggac ggtatcgata agcttaggag cttgggctgc aggtcgaggg cactgggagg 240  
 atgttgagta agatggaaaa ctactgatga cccttcgaga gacagagtat taggacatgt 300  
 ttgaacaggg gccggcgat cagcaggtag ctctagagga tccccgtctg tctgcacatt 360  
 tcgtagagcg agtggtccga tactctaate tccctaggca aggttcatat ttgtgtagg 420  
 tacttattct ccttttgttg actaagtcaa taatcagaat cagcaggttt ggagtcagct 480  
 tggcagggat cagcagcctg ggttggaagg aggggggata aaagcccctt caccaggaga 540  
 agccgtcaca cagatccaca agctcctgcc accatgg 577

15 <210> 70

<211> 4374

<212> ADN

<213> Secuencia Artificial

&lt;220&gt;

&lt;223&gt; coFVIII-5

&lt;400&gt; 70

5

atgcaaatcg aactgagcac ctgtttcttc ctctgcctgc tgagattctg tttctccgcg	60
acccgccgat actacctggg agcagtggag ctctcctggg attacatgca gagcgacctt	120
ggggagctgc ccgtggatgc caggttccct ccccggtgc caaagtcgtt tccgttcaac	180
acctccgtgg tgtacaagaa aactctgttc gtggagtcca ccgaccacct gttcaatatc	240
gccaagccca gacctccctg gatggggctg ttgggacctt ccatccaagc ggaggtgtac	300
gacactgtgg tcatcactct gaagaacatg gcctcgcac ccgtgtccct gcacgccgtg	360
ggagtgtctt actggaaagc gtccgagggg gccgaatacg acgaccagac ctgcgagaga	420
gaaaaggaag atgacaaggt gttcccagga ggatcgacac cctacgtgtg gcaagtgttg	480
aaggagaacg gcccaatggc ctccgaccgc ctgtgcctga cctactcgtt cctgtcccac	540
gtggacctcg tgaaggacct caactcggga ctgattggag ccctgctggt ctgcagggaa	600
ggctcactgg cgaagaaaaa gactcagacc ttgcacaagt tcattctgct gttcgtctgtg	660
ttcgacgagg ggaagtcgtg gcacagcgag actaagaact ccctgatgca agatagagat	720
gccgcctccg cccgggcctg gcctaagatg cacaccgtga acggttacgt gaaccgctcc	780
ctccctggcc tgattggatg ccaccggaag tccgtgtact ggacagtgat cgggatgggg	840
accacccccg aggtgcacag catcttcctg gaaggtcaca catttctcgt gcgcaaccac	900
cggcaggcct ccctggaaat cagccccatt accttcctca ctgccagac tctgctgatg	960
gacctgggac agttcctgct gttctgccat atctcctccc accaacaatga cggaatggag	1020
gcatacgtga aggtcgattc ctgccctgag gaacccagc tccgcatgaa gaacaatgag	1080
gaagccgagg actacgacga cgacctgacg gatagcgaga tggatgtggt ccggttcgat	1140
gacgataaca gcccttcctt catccaaatt cgctcgggtg caaagaagca cccaagacc	1200
tgggtgcatt acattgcggc ggaagaagag gactgggatt atgccccgct tgtcctcgct	1260
cctgacgacc ggagctacaa gagccagtac ctgaacaacg gtccacagag gatcggtaga	1320
aagtacaaga aggtccgctt catggcctat accgacgaaa cttcaaaac tagagaggcc	1380
atccaacacg aatccggcat cctgggcccgc ctctgtacg gagaagtcgg cgacaccctt	1440
ctcattatct tcaagaacca ggcttcccgc ccgtacaaca tctatccgca tgggatcact	1500
gacgtgcgcc cactgtactc ggggcgcctg cccaagggtg taaaacacct gaaggatatt	1560



ccgatccttc	cgaggagaaat	cttcaagtac	aagtggaccg	tgaccgtgga	agatggccca	1620
actaagtctg	accctagatg	cctcaccgcg	tactactcat	ccttcgtcaa	catggagcgc	1680
gacctggcca	gaggactgat	cgcccgctg	ctgatttgct	acaaggaatc	agtggaccaa	1740
cggggaaacc	agatcatgtc	ggataagagg	aacgtcatcc	tcttctccgt	gtttgacgaa	1800
aaccggctcg	ggtacctgac	tgaaaacatc	cagcggttcc	tccccaaccc	cgcgggctg	1860
cagctggaag	atcctgagtt	tcaggcatca	aacatcatgc	actccattaa	cggctacgtg	1920
ttcgattcgc	tgacgtgag	cgtgtgtctg	cacgaagtgg	cctactggtg	catcctgtcc	1980
attggtgccc	agactgactt	cctgtccgtg	tttttctccg	gctacacgtt	caagcacaag	2040
atggtgtacg	aggacaccct	gaccctcttc	ccttttccg	gcgaaactgt	gtttatgagc	2100
atggagaatc	ccggcctgtg	gatcttgggc	tgccacaaca	gcgacttccg	taacagagga	2160
atgactgcgc	tgctcaaggt	gtccagctgc	gacaagaaca	ccggagacta	ttatgaggac	2220
tcatacgagg	acatctccgc	ctacctcctg	tccaagaata	acgccattga	acctcggagc	2280
ttcagccaga	accacccctg	gcttaagaga	catcaacggg	agatcactag	gaccaccctg	2340
cagtcagacc	aggaggaaat	cgactacgat	gacaccatct	cggtcgagat	gaagaaggag	2400
gactttgaca	tctacgacga	agatgaaaac	cagagcccga	ggtcgttcca	aaagaaaacc	2460
cgccactact	ttattgctgc	tgctgagcgg	ctgtgggact	acggaatgtc	gtcctcgccg	2520
cacgtgctcc	gcaaccgagc	ccagagcggc	tcgggtgccgc	aattcaagaa	ggtcgtgttc	2580
caggagtcca	ctgacgggag	cttcaactcag	cctttgtacc	ggggagaact	caatgaacat	2640
ctcggcctcc	tcggacotta	catcagagca	gaagtggaag	ataacatcat	ggtcactttc	2700
cgtaaccaag	ccagccgccc	gtactcgttc	tactcctccc	tcatttctta	cgaagaggac	2760
cagcggcagg	gcgcagaacc	gcgcaagaac	ttcgtgaagc	ccaacgaaac	caagacctac	2820
ttctggaaag	tgacagcatca	tatggccccg	actaaggacg	agtttgactg	caaagcctgg	2880
gcctactttc	ccgatgtgga	cttgagagaag	gacgtccact	ccggcctcat	cggccccctg	2940
ctcgtgtgcc	ataccaatac	cctgaacccc	gcacacggtc	gccaggtcac	cgtgcaggag	3000
ttcgtctctg	tcttcaactat	cttcgacgaa	actaagtcct	ggtacttcac	cgagaacatg	3060
gagagggaact	gcagagcccc	ctgtaacatc	cagatggagg	acccgacgtt	caaggaaaac	3120
taccggttcc	acgccattaa	cggatacatc	atggatacgc	tgccgggtct	tgtgatggcc	3180
caggatcaac	ggatcagatg	gtacttattg	tcgatgggca	gcaacgagaa	catccactct	3240
attcactttc	ccggctcatgt	gttcaactgtg	cggaagaagg	aagagtacaa	gatggccctg	3300
tacaaccttt	atcccgaggt	gttcgaaaact	gtggaaatgc	tgccgtcgaa	ggccggcatt	3360
tggcgcgtgg	agtgtttgat	tggaagaacat	ctccatgcgg	ggatgtcaac	cctgttcctg	3420
gtgtatagca	acaagtgccca	gactccgctt	gggatggcgt	caggacacat	tagggatttc	3480

cagatcactg	cgccgggcca	gtacggccaa	tgggccccta	agctggcccg	cctgcattac	3540
tccggatcca	ttaacgcctg	gtcaaccaag	gagccattct	cctggatcaa	ggtggacctt	3600
ctggccccc	tgattatcca	cggaattaag	accaggggg	ccggcagaa	gttctcctca	3660
ctgtacatca	gccagttcat	aatcatgtac	tccctggacg	gaaagaagtg	gcaaacctac	3720
agggggaaca	gcaccggcac	actgatggtc	tttttcggaa	atgtggactc	ctccgggatt	3780
aagcataaca	tcttcaaccc	tccgattatc	gctcgggtaca	ttagacttca	ccctaccac	3840
tacagcatto	gctccaccct	gcggatggaa	ctgatgggct	gcgatctgaa	ctcgtgcagc	3900
atgccgttgg	gaatggagtc	caaagcaatt	tccgacgcgc	agatcacccg	ctcgtcctac	3960
tttaccaaca	tggtcgccac	gtggtcaccg	tccaaggccc	ggctgcacct	ccaggaaga	4020
tccaacgcat	ggcgggccaca	ggtcaacaac	cctaaggagt	ggctccaggt	ggacttccag	4080
aaaaccatga	aggtcaccgg	agtcacaacc	cagggagtga	agtcgctgct	gacttctatg	4140
tacgtcaagg	agttcctgat	ctccagcagc	caggacgggc	accagtggac	cctgttcttc	4200
caaaatggaa	aggtcaaggt	gtttcagggc	aatcaggatt	cattcacccc	ggtggtgaac	4260
tcccttgatc	cacccctcct	gacccgctac	cttcgcatcc	acccacagtc	ctgggtgcac	4320
cagatcgcg	tgaggatgga	ggtcctggga	tgcgaaagccc	aggacctgta	ctga	4374

&lt;210&gt; 71

&lt;211&gt; 4374

&lt;212&gt; ADN

&lt;213&gt; Secuencia Artificial

&lt;220&gt;

&lt;223&gt; coFVIII-6

&lt;400&gt; 71

atgcagattg	agctgtccac	ttgtttcttc	ctgtgcctcc	tgcgcttctg	tttctccgcc	60
actcgccggt	actaccttgg	agccgtggag	ctttcatggg	actacatgca	gagcgacctg	120
ggcgaaactcc	ccgtggatgc	cagattcccc	ccccgcgtgc	caaagtcctt	cccctttaac	180
acctccgtgg	tgtacaagaa	aacctctttt	gtcgagttca	ctgaccacct	gttcaacatc	240
gccaagccgc	gcccaccttg	gatgggcctc	ctgggaccga	ccattcaagc	tgaagtgtac	300
gacaccgtgg	tgatcacctc	gaagaacatg	gcgtcccacc	ccgtgtccct	gcctgcggtc	360
ggagtgtcct	actggaaggc	ctccgaagga	gctgagtacg	acgaccagac	tagccagcgg	420
gaaaaggagg	acgataaagt	gttcccgggc	ggctcgcata	cttacgtgtg	gcaagtcctg	480
aaggaaaacg	gacctatggc	atccgatcct	ctgtgcctga	cttactccta	cctttcccat	540
gtggacctcg	tgaaggacct	gaacagcggg	ctgattgggtg	cacttctcgt	gtgccgcgaa	600
ggttcgctcg	ctaaggaaaa	gacccagacc	ctccataagt	tcctcctttt	gttcgctgtg	660

ttcgatgaag gaaagtcattg gcattccgaa actaagaact cgctgatgca ggaccgggat	720
gccgcctcag cccgcgcctg gcctaaaatg catacagtca acggatacgt gaatcgggtca	780
ctgcccgggc tcatcggttg tcacagaaag tccgtgtact ggcacgtcat cggcatgggc	840
actacgcctg aagtgcactc catcttctctg gaagggcaca ccttctctctg gcgcaaccac	900
cgccaggcct ctctggaaat ctccccgatt acctttctga ccgccagac tctgctcatg	960
gacctggggc agttccttct cttctgccac atctccagcc atcagcacga cggaatggag	1020
gcctacgtga aggtggactc atgccggaa gaacctcagt tgcggatgaa gaacaacgag	1080
gaggccgagg actatgacga cgatttgact gactccgaga tggacgtcgt gcggttcgat	1140
gacgacaaca gcccagctt catccagatt cgcagcgtgg ccaagaagca ccccaaaacc	1200
tgggtgcact acatcgcggc cgaggaagaa gattgggact acgccccgtt ggtgctggca	1260
cccgatgacc ggtcgtacaa gtcccagtat ctgaacaatg gtccgcagcg gattggcaga	1320
aagtacaaga aagtgcggtt catggcgtag actgacgaaa cgtttaagac ccgggaggcc	1380
attcaacatg agagcgcat tctgggacca ctgctgtacg gagaggtcgg cgataccctg	1440
ctcatcatct tcaaaaacca ggctcccgg ccttacaaca tctaccctca cggaatcacc	1500
gacgtggggc cactctactc gggcgccctg ccgaagggcg tcaagcacct gaaagacttc	1560
cctatcctgc cgggcgaaat cttcaagtat aagtggaccg tcaccgtgga ggacggggccc	1620
accaagagcg atcctagggtg tctgactcgg tactactcca gcttcgtgaa catggaacgg	1680
gacctggcat cgggactcat tggaccgctg ctgatctgct acaaagagtc ggtggatcaa	1740
cgcggaacc agatcatgtc cgacaagcgc aacgtgatcc tgttctccgt gtttgatgaa	1800
aacagatcct ggtacctcac tgaaaacatc cagaggttcc tcccaaacc cgcaggagtg	1860
caactggagg accctgagtt tcaggcctcg aatatcatgc actcgattaa cggttacgtg	1920
ttcgactcgc tgcagctgag cgtgtgcctc catgaagtgc cttactggta cattctgtcc	1980
atcggcgccc agactgactt cctgagcgtg ttcttttccg gttacacett taagcacaag	2040
atggtgtacg aagataccct gacctgttc cctttctccg gcgaaacggt gttcatgtcg	2100
atggagaacc cgggtctgtg gattctggga tgccacaaca gcgactttcg gaaccgcgga	2160
atgactgcc tgctgaagggt gtcctcatgc gacaagaaca ccggagacta ctacgaggac	2220
tcctacgagg atatctcagc ctacctctg tccaagaaca acgcatcga gccgcgcagc	2280
ttcagccaga accgcctgt gctgaagagg caccagcgag aaattaccgg gaccaccctc	2340
caatcgatc aggaggaat cgactacgac gacaccatct cggtggaat gaagaaggaa	2400
gatttcgata tctacgacga ggacgaaaat cagtccccctc gtcatttcca aaagaaaact	2460
agacactact ttatcgccgc ggtggaaga ctgtgggact atggaatgtc atccagccct	2520
cacgtccttc ggaaccgggc ccagagcgga tcggtgcctc agttcaagaa agtgggtgttc	2580

caggagttca ccgacggcag cttcaccacag ccgctgtacc ggggagaact gaacgaacac 2640  
ctgggcctgc tcggtcccta catccgcgcg gaagtggagg ataacatcat ggtgacottc 2700  
cgtaaccaag catccagacc ttactccttc tattcctccc tgatctcata cgaggaggac 2760  
cagcgccaag gcgccgagcc ccgcaagaac ttcgtcaagc ccaacgagac taagacctac 2820  
ttctggaagg tccaacacca tatggccccg accaaggatg agtttgactg caaggcctgg 2880  
gcctacttct ccgacgtgga ccttgagaag gatgtccatt ccggcctgat cgggccgctg 2940  
ctcgtgtgtc acaccaacac cctgaaccca gcgcatggac gccaggtcac cgtccaggag 3000  
tttgctctgt tcttcacat ttttgacgaa actaagtcct ggtacttcac cgagaatatg 3060  
gagcgaaaact gttagagcgc ctgcaatata cagatggaag atccgacttt caaggagaac 3120  
tatagattcc acgccatcaa cgggtacatc atggatactc tgccggggct ggtcatggcc 3180  
caggatcaga ggattcgggt gtacttgctg tcaatgggat cgaacgaaaa cattcactcc 3240  
attcacttct ccggtcacgt gttcactgtg cgcaagaagg aggagtacaa gatggcgctg 3300  
tacaatctgt accccggggt gttcgaaaact gtggagatgc tgccgtccaa ggccggcatc 3360  
tgagagtggt agtgccctgat cggagagcac ctccacgcgg ggatgtccac cctcttcctg 3420  
gtgtactoga ataagtgcc gaccccgctg ggcatggcct cgggccacat cagagacttc 3480  
cagatcacag caagcggaca atacggccaa tgggcgccga agctggcccc cttgcactac 3540  
tccggatcga tcaacgcagt gtccaccaag gaaccgttct cgtggattaa ggtggacctc 3600  
ctggccccta tgattatcca cggaattaag acccagggcg ccaggcagaa gttctcctcc 3660  
ctgtacatct cgcaattcat catcatgtac agcctggacg ggaagaagtg gcagacttac 3720  
aggggaaaact ccaccggcac cctgatggtc tttttcgcca acgtggattc ctccggcatt 3780  
aagcacaaca tcttcaaccc accgatcata gccagatata ttaggctcca cccactcac 3840  
tactcaatcc gctcaactct tcggatggaa ctcatggggg gcgacctgaa ctctgctcc 3900  
atgccgttgg ggatggaatc aaaggctatt agcgacgccc agatcaccgc gagctcctac 3960  
ttactaaca tgttcgccac ctggagcccc tccaaggcca ggctgcactt gcagggacgg 4020  
tcaaagtccg ggcggccgca agtgaacaat ccgaaggaat ggcttcaagt ggatttccaa 4080  
aagaccatga aagtgaccgg agtcaccacc caggagtgta agtccttct gacctcgatg 4140  
tatgtgaagg agttcctgat tagcagcagc caggacgggc accagtggac cctgtttctc 4200  
caaaacggaa aggtcaaggt gttccagggg aaccaggact cgttcacacc cgtggtgaac 4260  
tccctggacc cccactgct gacgcgttac ttgaggattc atcctcagtc ctgggtccat 4320  
cagattgcat tgcgaatgga agtcctgggc tgcgaggccc aggacctgta ctga 4374

<210> 72  
<211> 4824  
5 <212> ADN  
<213> Secuencia Artificial

<220>  
10 <223> coFVIII-6-XTEN  
<400> 72

atgcagattg agctgtccac ttgtttcttc ctgtgcctcc tgcgcttctg tttctccgcc	60
actcgccggt actaccttgg agccgtggag ctttcatggg actacatgca gagcgacctg	120
ggcgaactcc ccgtggatgc cagattcccc ccccgctgac caaagtcctt cccctttaac	180
acctccgtgg tgtacaagaa aaccctcttt gtcgagttca ctgaccacct gttcaacatc	240
gccaaagccg gccacacttg gatgggcctc ctgggaccga ccattcaagc tgaagtgtac	300
gacaccgtgg tgatcacctt gaagaacatg gcgtcccacc ccgtgtccct gcatgcggtc	360
ggagtgtcct actggaaggc ctccgaagga gctgagtacg acgaccagac tagccagcgg	420
gaaaaggagg acgataaagt gttcccgggc ggctcgcata cttacgtgtg gcaagtcctg	480
aagggaaacg gacctatggc atccgatcct ctgtgcctga cttactccta cctttcccat	540
gtggacctcg tgaaggacct gaacagcggg ctgattggtg cacttctcgt gtgccgcgaa	600
ggttcgctcg ctaaggaaaa gaccagacc ctccataagt tcatcctttt gttcgctgtg	660
ttcgatgaag gaaagtcatg gcattccgaa actaagaact cgctgatgca ggaccgggat	720
gccgcctcag cccgcgcctg gcctaaaatg catacagtca acggatacgt gaatcggta	780
ctgcccgggc tcatcggttg tcacagaaag tccgtgtact ggcacgtcat cggcatgggc	840
actacgcctg aagtgcactc catcttcctg gaagggcaca ccttcctcgt gcgcaaccac	900
cgccaggcct ctctggaaat ctccccgatt acctttctga ccgccagac tctgctcatg	960
gacctggggc agttccttct cttctgccac atctccagcc atcagcacga cggaatggag	1020
gcctacgtga agtgggactc atgcccggaa gaacctcagt tgcggatgaa gaacaacgag	1080
gaggccgagg actatgacga cgatttgact gactccgaga tggacgtcgt gcggttcgat	1140
gacgacaaca gccccagctt catccagatt cgcagcgtgg ccaagaagca ccccaaaacc	1200
tgggtgcact acatcgcggc cgaggaagaa gattgggact acgccccgtt ggtgctggca	1260
cccgatgacc ggtcgtacaa gtcccagtat ctgaacaatg gtccgcagcg gattggcaga	1320
aagtacaaga aagtgcggtt catggcgtag actgacgaaa cgtttaagac ccgggaggcc	1380
attcaacatg agagcggcat tctgggacca ctgctgtacg gagaggtcgg cgataccctg	1440
ctcatcatct tcaaaaacca ggctccccg ccttacaaca tctaccctca cggaatcacc	1500
gacgtgcggc cactctactc gcggcgctg ccgaagggcg tcaagcacct gaaagacttc	1560
cctatcctgc cgggcgaaat cttcaagtat aagtggaccg tcaccgtgga ggacgggccc	1620
accaagagcg atcctaggtg tctgactcgg tactactcca gcttcgtgaa catggaacgg	1680

# ES 2 926 585 T3

gacctggcat	cgggactcat	tggaccgctg	ctgatctgct	acaaagagtc	ggtggatcaa	1740
cgcggaacc	agatcatgtc	cgacaagcgc	aacgtgatcc	tgttctccgt	gtttgatgaa	1800
aacagatcct	ggtacctcac	tgaaaaacatc	cagagggttc	tcccaaacc	cgcaggagtg	1860
caactggagg	accctgagtt	tcaggcctcg	aatatcatgc	actcgattaa	cggttacgtg	1920
ttcgactcgc	tgcagctgag	cgtgtgcctc	catgaagtcg	cttactggta	cattctgtcc	1980
atcgcgccc	agactgactt	cctgagcgtg	ttcttttccg	gttacacctt	taagcacaag	2040
atggtgtacg	aagataccct	gaccctgttc	cctttctccg	gcgaaacggt	gttcatgtcg	2100
atggagaacc	cggtctgtg	gattctggga	tgccacaaca	gcgactttcg	gaaccgcgga	2160
atgactgccc	tgctgaaggt	gtcctcatgc	gacaagaaca	ccggagacta	ctacgaggac	2220
tcctacgagg	atatctcagc	ctacctcctg	tccaagaaca	acgcgatcga	gccgcgcagc	2280
ttcagccaga	acggcgcgcc	aacatcagag	agcgccaccc	ctgaaagtgg	tcccgggagc	2340
gagccagcca	catctgggtc	ggaaacgcca	ggcacaagtg	agtctgcaac	tcccaggtcc	2400
ggacctggct	ccgagcctgc	cactagcggc	tccgagactc	cgggaacttc	cgagagcgct	2460
acaccagaaa	gcggaccctg	aaccagtacc	gaacctagcg	agggtctctg	tccgggcagc	2520
ccagccggct	ctcctacatc	cacggaggag	ggcacttccg	aatccgccac	cccggagtca	2580
gggccaggat	ctgaaccctg	tacctcaggc	agtgagacgc	caggaacgag	cgagtccgct	2640
acaccggaga	gtgggcccag	gagccctgct	ggatctccta	cgtccactga	ggaagggtca	2700
ccagcgggct	cgccaccag	cactgaagaa	ggtgcctcga	gcccgcctgt	gctgaagagg	2760
caccagcgag	aaattaccct	gaccaccctc	caatcggtac	aggaggaaat	cgactacgac	2820
gacaccatct	cggtggaaat	gaagaaggaa	gatttcgata	tctacgacga	ggacgaaaat	2880
cagtcccctc	gtcatttcca	aaagaaaact	agacactact	ttatcgccgc	ggtggaaaga	2940
ctgtgggact	atggaatgtc	atccagccct	cacgtccttc	ggaaccgggc	ccagagcgga	3000
tcggtgcctc	agttcaagaa	agtgggtgtc	caggagttca	ccgacggcag	cttcacccag	3060
ccgctgtacc	ggggagaact	gaacgaacac	ctgggcctgc	tcggtcccta	catccgcgcg	3120
gaagtggagg	ataacatcat	ggtgaccttc	cgtaaccaag	catccagacc	ttactccttc	3180
tattcctccc	tgatctcata	cgaggaggac	cagcgccaag	gcgccgagcc	ccgcaagaac	3240
ttcgtcaagc	ccaacgagac	taagacctac	ttctggaagg	tccaacacca	tatggccccg	3300
accaaggatg	agtttgactg	caaggcctgg	gcctacttct	ccgacgtgga	ccttgagaag	3360
gatgtccatt	ccggcctgat	cgggccgctg	ctcgtgtgtc	acaccaacac	cctgaaccca	3420
gcgcatggac	gccaggtcac	cgtccaggag	tttgctctgt	tcttcaccat	ttttgacgaa	3480
actaagtcc	ggtacttcac	cgagaatatg	gagcgaaact	gtagagcgcc	ctgcaatatc	3540

cagatggaag atccgacttt caaggagaac tatagattcc acgccatcaa cgggtacatc 3600  
 atggatactc tgccggggct ggtcatggcc caggatcaga ggattcgggtg gtacttgctg 3660  
 tcaatgggat cgaacgaaaa cattcactcc attcacttct ccggtcacgt gttcactgtg 3720  
 cgcaagaagg aggagtacaa gatggcgctg tacaatctgt accccgggggt gttcgaaact 3780  
 gtggagatgc tgccgtccaa ggccggcatc tggagagtgg agtgctgat cggagagcac 3840  
 ctccacgcgg ggatgtccac cctcttcctg gtgtactcga ataagtgcc gaccccgctg 3900  
 ggcattggcct cgggccacat cagagacttc cagatcacag caagcggaca atacggccaa 3960  
 tgggcgccga agctggcccc cttgcactac tccggatcga tcaacgcatg gtccaccaag 4020  
 gaaccgttct cgtggattaa ggtggacctc ctggcccta tgattatcca cggaattaag 4080  
 acccagggcg ccaggcagaa gttctcctcc ctgtacatct cgcaattcat catcatgtac 4140  
 agcctggacg ggaagaagtg gcagacttac aggggaaact ccaaccggcac cctgatggtc 4200  
 tttttcggca acgtggattc ctccggcatt aagcacaaca tcttcaacc accgatcata 4260  
 gccagatata ttaggctcca cccactcac tactcaatcc gctcaactct tcggatggaa 4320  
 ctcatggggt gcgacctgaa ctctgctcc atgccgttg ggatggaatc aaaggctatt 4380  
 agcgacgccc agatcacgcg gagctcctac ttcactaaca tgttcgccac ctggagcccc 4440  
 tccaaggcca ggctgcactt gcagggacgg tcaaatgcct ggcgcccgca agtgaacaat 4500  
 ccgaaggaaat ggcttcaagt ggatttccaa aagaccatga aagtgaccgg agtcaccacc 4560  
 caggagtgta agtcccttct gacctcgatg tatgtgaagg agttcctgat tagcagcagc 4620  
 caggacgggc accagtggac cctgttcttc caaaacggaa aggtcaaggt gttccagggg 4680  
 aaccaggact cgttcacacc cgtggtgaac tccctggacc cccactgct gacgcgttac 4740  
 ttgaggattc atcctcagtc ctgggtccat cagattgcat tgccaatgga agtcctgggc 4800  
 tgcgaggccc aggacctgta ctga 4824

&lt;210&gt; 73

&lt;211&gt; 12

5 &lt;212&gt; PRT

&lt;213&gt; Secuencia Artificial

&lt;220&gt;

&lt;223&gt; motivo AD

10

&lt;400&gt; 73

Gly Glu Ser Pro Gly Gly Ser Ser Gly Ser Glu Ser  
 1 5 10

15 &lt;210&gt; 74

&lt;211&gt; 12

&lt;212&gt; PRT

&lt;213&gt; Secuencia Artificial

20 &lt;220&gt;

&lt;223&gt; motivo AD

&lt;400&gt; 74

Gly Ser Glu Gly Ser Ser Gly Pro Gly Glu Ser Ser  
 1 5 10

25

<210> 75  
 <211> 12  
 <212> PRT  
 5 <213> Secuencia Artificial  
  
 <220>  
 <223> motivo AD  
  
 10 <400> 75  
  
**Gly Ser Ser Glu Ser Gly Ser Ser Glu Gly Gly Pro**  
**1 5 10**  
  
 <210> 76  
 <211> 12  
 <212> PRT  
 15 <213> Secuencia Artificial  
  
 <220>  
 20 <223> motivo AD  
  
 <400> 76  
  
**Gly Ser Gly Gly Glu Pro Ser Glu Ser Gly Ser Ser**  
**1 5 10**  
  
 25 <210> 77  
 <211> 12  
 <212> PRT  
 30 <213> Secuencia Artificial  
  
 <220>  
 <223> motivo AE, AM  
  
 <400> 77  
  
 35 **Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu**  
**1 5 10**  
  
 <210> 78  
 <211> 12  
 40 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> motivo AE, AM, AQ  
  
 45 <400> 78  
  
**Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro**  
  
 50 **1 5 10**  
  
 <210> 79  
 <211> 12  
 <212> PRT  
 55 <213> Secuencia Artificial  
  
 <220>  
 <223> motivo AE, AM, AQ  
  
 60 <400> 79  
  
**Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro**  
**1 5 10**



<210> 80  
 <211> 12  
 <212> PRT  
 5 <213> Secuencia Artificial  
  
 <220>  
 <223> motivo AE, AM, AQ  
  
 10 <400> 80  
  
**Gly Thr Ser Thr Glu Pro Ser Glu Gly Ser Ala Pro**  
**1 5 10**  
  
 <210> 81  
 15 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 20 <223> motivo AF, AM  
  
 <400> 81  
  
**Gly Ser Thr Ser Glu Ser Pro Ser Gly Thr Ala Pro**  
**1 5 10**  
  
 25 <210> 82  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
  
 30 <220>  
 <223> motivo AF, AM  
  
 <400> 82  
  
 35 <210> 83  
 <211> 12  
 40 <212> PRT  
 <213> Secuencia Artificial  
  
 <220>  
 <223> motivo AF, AM  
  
 45 <400> 83  
  
**Gly Thr Ser Pro Ser Gly Glu Ser Ser Thr Ala Pro**  
**1 5 10**  
  
 50 <210> 84  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
  
 55 <220>  
 <223> motivo AF, AM  
  
 <400> 84  
  
**Gly Ser Thr Ser Ser Thr Ala Glu Ser Pro Gly Pro**  
 60 **1 5 10**

<210> 85  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 5  
 <220>  
 <223> motivo AG, AM  
 <400> 85  
 10  
 Gly Thr Pro Gly Ser Gly Thr Ala Ser Ser Ser Pro  
 1 5 10  
 <210> 86  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 15  
 <220>  
 <223> motivo AG, AM  
 20  
 <400> 86  
 Gly Ser Ser Thr Pro Ser Gly Ala Thr Gly Ser Pro  
 1 5 10  
 25  
 <210> 87  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 30  
 <220>  
 <223> motivo AG, AM  
 <400> 87  
 35  
 Gly Ser Ser Pro Ser Ala Ser Thr Gly Thr Gly Pro  
 1 5 10  
 <210> 88  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 40  
 <220>  
 <223> motivo AG, AM  
 45  
 <400> 88  
 Gly Ala Ser Pro Gly Thr Ser Ser Thr Gly Ser Pro  
 1 5 10  
 50  
 <210> 89  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 55  
 <220>  
 <223> motivo AQ  
 <400> 89  
 60  
 Gly Glu Pro Ala Gly Ser Pro Thr Ser Thr Ser Glu  
 1 5 10

<210> 90  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 5  
 <220>  
 <223> motivo AQ  
 <400> 90  
 10  
**Gly Thr Gly Glu Pro Ser Ser Thr Pro Ala Ser Glu**  
**1 5 10**  
 <210> 91  
 <211> 12  
 15 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> motivo AQ  
 20 <400> 91  
**Gly Ser Gly Pro Ser Thr Glu Ser Ala Pro Thr Glu**  
**1 5 10**  
 25 <210> 92  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 30 <220>  
 <223> motivo AQ  
 <400> 92  
**Gly Ser Glu Thr Pro Ser Gly Pro Ser Glu Thr Ala**  
 35 **1 5 10**  
 <210> 93  
 <211> 12  
 <212> PRT  
 40 <213> Secuencia Artificial  
 <220>  
 <223> motivo AQ  
 45 <400> 93  
**Gly Pro Ser Glu Thr Ser Thr Ser Glu Pro Gly Ala**  
**1 5 10**  
 <210> 94  
 <211> 12  
 50 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 55 <223> motivo AQ  
 <400> 94  
**Gly Ser Pro Ser Glu Pro Thr Glu Gly Thr Ser Ala**  
**1 5 10**  
 60

<210> 95  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 5  
 <220>  
 <223> motivo BC  
 <400> 95  
 10  
**Gly Ser Gly Ala Ser Glu Pro Thr Ser Thr Glu Pro**  
**1 5 10**  
 <210> 96  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 15  
 <220>  
 <223> motivo BC  
 20  
 <400> 96  
**Gly Ser Glu Pro Ala Thr Ser Gly Thr Glu Pro Ser**  
 25 **1 5 10**  
 <210> 97  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 30  
 <220>  
 <223> motivo BC  
 35  
 <400> 97  
**Gly Thr Ser Glu Pro Ser Thr Ser Glu Pro Gly Ala**  
**1 5 10**  
 <210> 98  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 40  
 <220>  
 <223> motivo BC  
 45  
 <400> 98  
**Gly Thr Ser Thr Glu Pro Ser Glu Pro Gly Ser Ala**  
**1 5 10**  
 50  
 <210> 99  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 55  
 <220>  
 <223> motivo BD  
 <400> 99  
 60  
**Gly Ser Thr Ala Gly Ser Glu Thr Ser Thr Glu Ala**  
**1 5 10**

<210> 100  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 5  
 <220>  
 <223> motivo BD  
 <400> 100  
 10  
**Gly Ser Glu Thr Ala Thr Ser Gly Ser Glu Thr Ala**  
**1 5 10**  
 <210> 101  
 <211> 12  
 15 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> motivo BD  
 20 <400> 101  
**Gly Thr Ser Glu Ser Ala Thr Ser Glu Ser Gly Ala**  
**1 5 10**  
 25 <210> 102  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 30 <220>  
 <223> motivo BD  
 <400> 102  
**Gly Thr Ser Thr Glu Ala Ser Glu Gly Ser Ala Ser**  
 35 **1 5 10**  
 <210> 103  
 <211> 1607  
 <212> PRT  
 40 <213> Secuencia Artificial  
 <220>  
 <223> coFVIII-6-XTEN Secuencia de Proteína  
 45 <400> 103

# ES 2 926 585 T3

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe  
1 5 10 15

Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser  
20 25 30

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg  
35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val  
50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile  
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln  
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser  
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser

# ES 2 926 585 T3

115		120		125
Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp				
130		135		140
Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu				
145		150		155
				160
Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser				
		165		170
				175
Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile				
		180		185
				190
Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr				
		195		200
				205
Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly				
		210		215
				220
Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp				
225		230		235
				240
Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr				
		245		250
				255
Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val				
		260		265
				270
Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile				
		275		280
				285
Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser				
		290		295
				300
Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met				
305		310		315
				320
Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His				
		325		330
				335
Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro				
		340		345
				350
Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp				
		355		360
				365

# ES 2 926 585 T3

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser  
370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr  
385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro  
405 410 415

Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn  
420 425 430

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met  
435 440 445

Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu  
450 455 460

Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu  
465 470 475 480

Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro  
485 490 495

His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys  
500 505 510

Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe  
515 520 525

Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp  
530 535 540

Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg  
545 550 555 560

Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu  
565 570 575

Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val  
580 585 590

Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu  
595 600 605

Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp  
610 615 620



# ES 2 926 585 T3

Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val  
 625 630 635 640  
 Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp  
 645 650 655  
 Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe  
 660 665 670  
 Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr  
 675 680 685  
 Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro  
 690 695 700  
 Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly  
 705 710 715 720  
 Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp  
 725 730 735  
 Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys  
 740 745 750  
 Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Gly Ala Pro Thr  
 755 760 765  
 Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser Glu Pro Ala Thr  
 770 775 780  
 Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser  
 785 790 795 800  
 Gly Pro Gly Ser Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr  
 805 810 815  
 Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Thr Ser Thr Glu Pro  
 820 825 830  
 Ser Glu Gly Ser Ala Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr  
 835 840 845  
 Glu Glu Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser Gly Pro Gly Ser  
 850 855 860  
 Glu Pro Ala Thr Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala  
 865 870 875 880

# ES 2 926 585 T3

Thr Pro Glu Ser Gly Pro Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr  
 885 890 895  
 Glu Glu Gly Ser Pro Ala Gly Ser Pro Thr Ser Thr Glu Glu Gly Ala  
 900 905 910  
 Ser Ser Pro Pro Val Leu Lys Arg His Gln Arg Glu Ile Thr Arg Thr  
 915 920 925  
 Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser  
 930 935 940  
 Val Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn  
 945 950 955 960  
 Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg His Tyr Phe Ile Ala  
 965 970 975  
 Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser Ser Pro His Val  
 980 985 990  
 Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro Gln Phe Lys Lys Val  
 995 1000 1005  
 Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr  
 1010 1015 1020  
 Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly Pro Tyr Ile  
 1025 1030 1035  
 Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Arg Asn Gln  
 1040 1045 1050  
 Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr Glu  
 1055 1060 1065  
 Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe Val Lys  
 1070 1075 1080  
 Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His Met  
 1085 1090 1095  
 Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe  
 1100 1105 1110  
 Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly

# ES 2 926 585 T3

1115	1120	1125
Pro Leu Leu Val Cys His Thr	Asn Thr Leu Asn Pro	Ala His Gly
1130	1135	1140
Arg Gln Val Thr Val Gln Glu	Phe Ala Leu Phe Phe	Thr Ile Phe
1145	1150	1155
Asp Glu Thr Lys Ser Trp Tyr	Phe Thr Glu Asn Met	Glu Arg Asn
1160	1165	1170
Cys Arg Ala Pro Cys Asn Ile	Gln Met Glu Asp Pro	Thr Phe Lys
1175	1180	1185
Glu Asn Tyr Arg Phe His Ala	Ile Asn Gly Tyr Ile	Met Asp Thr
1190	1195	1200
Leu Pro Gly Leu Val Met Ala	Gln Asp Gln Arg Ile	Arg Trp Tyr
1205	1210	1215
Leu Leu Ser Met Gly Ser Asn	Glu Asn Ile His Ser	Ile His Phe
1220	1225	1230
Ser Gly His Val Phe Thr Val	Arg Lys Lys Glu Glu	Tyr Lys Met
1235	1240	1245
Ala Leu Tyr Asn Leu Tyr Pro	Gly Val Phe Glu Thr	Val Glu Met
1250	1255	1260
Leu Pro Ser Lys Ala Gly Ile	Trp Arg Val Glu Cys	Leu Ile Gly
1265	1270	1275
Glu His Leu His Ala Gly Met	Ser Thr Leu Phe Leu	Val Tyr Ser
1280	1285	1290
Asn Lys Cys Gln Thr Pro Leu	Gly Met Ala Ser Gly	His Ile Arg
1295	1300	1305
Asp Phe Gln Ile Thr Ala Ser	Gly Gln Tyr Gly Gln	Trp Ala Pro
1310	1315	1320
Lys Leu Ala Arg Leu His Tyr	Ser Gly Ser Ile Asn	Ala Trp Ser
1325	1330	1335
Thr Lys Glu Pro Phe Ser Trp	Ile Lys Val Asp Leu	Leu Ala Pro
1340	1345	1350

# ES 2 926 585 T3

Met	Ile	Ile	His	Gly	Ile	Lys	Thr	Gln	Gly	Ala	Arg	Gln	Lys	Phe
1355						1360					1365			
Ser	Ser	Leu	Tyr	Ile	Ser	Gln	Phe	Ile	Ile	Met	Tyr	Ser	Leu	Asp
1370						1375					1380			
Gly	Lys	Lys	Trp	Gln	Thr	Tyr	Arg	Gly	Asn	Ser	Thr	Gly	Thr	Leu
1385						1390					1395			
Met	Val	Phe	Phe	Gly	Asn	Val	Asp	Ser	Ser	Gly	Ile	Lys	His	Asn
1400						1405					1410			
Ile	Phe	Asn	Pro	Pro	Ile	Ile	Ala	Arg	Tyr	Ile	Arg	Leu	His	Pro
1415						1420					1425			
Thr	His	Tyr	Ser	Ile	Arg	Ser	Thr	Leu	Arg	Met	Glu	Leu	Met	Gly
1430						1435					1440			
Cys	Asp	Leu	Asn	Ser	Cys	Ser	Met	Pro	Leu	Gly	Met	Glu	Ser	Lys
1445						1450					1455			
Ala	Ile	Ser	Asp	Ala	Gln	Ile	Thr	Ala	Ser	Ser	Tyr	Phe	Thr	Asn
1460						1465					1470			
Met	Phe	Ala	Thr	Trp	Ser	Pro	Ser	Lys	Ala	Arg	Leu	His	Leu	Gln
1475						1480					1485			
Gly	Arg	Ser	Asn	Ala	Trp	Arg	Pro	Gln	Val	Asn	Asn	Pro	Lys	Glu
1490						1495					1500			
Trp	Leu	Gln	Val	Asp	Phe	Gln	Lys	Thr	Met	Lys	Val	Thr	Gly	Val
1505						1510					1515			
Thr	Thr	Gln	Gly	Val	Lys	Ser	Leu	Leu	Thr	Ser	Met	Tyr	Val	Lys
1520						1525					1530			
Glu	Phe	Leu	Ile	Ser	Ser	Ser	Gln	Asp	Gly	His	Gln	Trp	Thr	Leu
1535						1540					1545			
Phe	Phe	Gln	Asn	Gly	Lys	Val	Lys	Val	Phe	Gln	Gly	Asn	Gln	Asp
1550						1555					1560			
Ser	Phe	Thr	Pro	Val	Val	Asn	Ser	Leu	Asp	Pro	Pro	Leu	Leu	Thr
1565						1570					1575			
Arg	Tyr	Leu	Arg	Ile	His	Pro	Gln	Ser	Trp	Val	His	Gln	Ile	Ala
1580						1585					1590			
Leu	Arg	Met	Glu	Val	Leu	Gly	Cys	Glu	Ala	Gln	Asp	Leu	Tyr	
1595						1600					1605			

## REIVINDICACIONES

1. Un ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica un polipéptido con actividad de FVIII, en donde la secuencia de nucleótidos comprende al menos 95 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71.
- 5 2. La molécula de ácido nucleico aislada de la reivindicación 1, en donde la secuencia de nucleótidos comprende al menos 99 % de identidad de secuencia con los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71.
3. La molécula de ácido nucleico aislada de la reivindicación 1, en donde la secuencia de nucleótidos comprende los nucleótidos 58-2277 y 2320-4374 de SEQ ID NO: 71.
- 10 4. La molécula de ácido nucleico aislada de una cualquiera de las reivindicaciones 1 a 3, en donde la secuencia de nucleótidos comprende, además, una secuencia de ácido nucleico que codifica un péptido señal, en donde la secuencia de ácido nucleico que codifica el péptido señal tiene al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % de identidad de secuencia con:
  - (i) los nucleótidos 1 a 57 de SEQ ID NO: 1;
  - (ii) los nucleótidos 1 a 57 de SEQ ID NO: 2;
  - (iii) los nucleótidos 1 a 57 de SEQ ID NO: 3;
  - (iv) los nucleótidos 1 a 57 de SEQ ID NO: 4;
  - (v) los nucleótidos 1 a 57 de SEQ ID NO: 5;
  - (vi) los nucleótidos 1 a 57 de SEQ ID NO: 6;
  - (vii) los nucleótidos 1 a 57 de SEQ ID NO: 70;
  - (viii) los nucleótidos 1 a 57 de SEQ ID NO: 71; o
  - (ix) los nucleótidos 1 a 57 de SEQ ID NO: 68.
- 25 5. La molécula de ácido nucleico aislada de una cualquiera de las reivindicaciones 1 a 4, en donde la molécula de ácido nucleico comprende una o más propiedades seleccionadas del grupo que consiste en:
  - (a) el índice de adaptación de codones humanos de la molécula de ácido nucleico o una porción de la misma se incrementa con respecto a SEQ ID NO: 16;
  - (b) la frecuencia de codones óptimos de la secuencia de nucleótidos o una porción de la misma se incrementa con respecto a la SEQ ID NO: 16;
  - 30 (c) la secuencia de nucleótidos o una porción de la misma contiene un mayor porcentaje de nucleótidos G/C en comparación con el porcentaje de nucleótidos G/C en SEQ ID NO: 16;
  - (d) el uso relativo de codones sinónimos de la secuencia de nucleótidos o una porción de la misma se incrementa con respecto a SEQ ID NO: 16;
  - 35 (e) el número efectivo de codones de la secuencia de nucleótidos o una porción de la misma se reduce con respecto a SEQ ID NO: 16;
  - (f) la secuencia de nucleótidos contiene menos secuencias MARS/ARS (SEQ ID NOs: 21 y 22) con respecto a SEQ ID NO: 16;
  - (g) la secuencia de nucleótidos contiene menos elementos desestabilizantes (SEQ ID NO: 23 y 24) con respecto a SEQ ID NO: 16; y
  - 40 (h) cualquier combinación de los mismos.
6. La molécula de ácido nucleico aislada de una cualquiera de las reivindicaciones 1 a 5, que comprende, además, una secuencia de nucleótidos heteróloga que codifica una secuencia de aminoácidos heteróloga.
- 45 7. La molécula de ácido nucleico aislada de la reivindicación 6, en donde la secuencia heteróloga de aminoácidos es una región constante de inmunoglobulina o una porción de la misma, XTEN, transferrina, albúmina o una secuencia PAS.
8. La molécula de ácido nucleico aislada de la reivindicación 7, en donde la secuencia de nucleótidos comprende SEQ ID NO: 72.
9. La molécula de ácido nucleico aislada de la reivindicación 6 o 7, en donde la secuencia de aminoácidos heteróloga está enlazada al extremo N o al extremo C de la secuencia de aminoácidos codificada por la secuencia de nucleótidos o insertada entre dos aminoácidos en la secuencia de aminoácidos codificada por la secuencia de nucleótidos en uno o más sitios de inserción seleccionados de la Tabla 3.
- 50 10. La molécula de ácido nucleico aislada de una cualquiera de las reivindicaciones 1 a 9, en donde el polipéptido de FVIII es un FVIII de longitud completa o un FVIII con el dominio B suprimido.
11. Un vector que comprende la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 10.

12. El vector de la reivindicación 11, en donde el vector es un vector lentiviral.
13. Una célula huésped que comprende la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 10 o el vector de la reivindicación 11 o 12.
- 5 14. Un método para producir un polipéptido con actividad de FVIII, que comprende: cultivar la célula huésped de la reivindicación 13 bajo condiciones en las que se produce un polipéptido con actividad de FVIII y recuperar el polipéptido con actividad de FVIII.
15. Una molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 10, o el vector de la reivindicación 11 o 12, para uso en un método para tratar un trastorno hemorrágico.

**FIG. 1A: coFVIII-3 – SEQ ID NO: 1**

ATGCAGATCGAACTGAGCACCTGCTTCTTTCTGTGCCTGCTGAGGTTTTGCTTTAGCGCCACCAGGAGATACTATCTGGGCGC  
 CGTGGAAGTGAAGTGGGACTATATGCAGTCTGATCTGGGCGAAGTGGCAGTGGATGCCAGGTTTCCCCCAGAGTGCCCAAAA  
 GCTTTCCCTTTAATACCAGCGTGGTGTATAAGAAAACCTGTTTGTGAATTCACTGATCATCTGTTAATATCGCCAAGCCC  
 AGGCCCCCTGGATGGGCGTGTGGGCCCCACCATCCAGGCCGAAGTGTATGATACCGTGGTCATCACCTGAAAAACATGGC  
 CAGCCATCCAGTGAGCCTGCATGCTGTGGGCGTCTCCTATTGGAAAGCCTCTGAAGGCGCCGAGTATGATGATCAGACCAGCC  
 AGAGGGAAAAAGAAGATGATAAAGTCTTTCTGGGGGAGCCATACCTATGTCTGGCAGGTCCTGAAAGAAAATGGCCCCATG  
 GCCAGCGATCCCCTGTGCCTGACCTATAGCTATCTGAGCCATGTGGACCTGGTGAAGGATCTGAACAGCGGCCCTGATTGGGGC  
 CCTGCTGGTGTGCAGGGAAGGCAGCCTGGCCAAAGAAAAAACCCAGACCTGCATAAGTTTATCCTGCTGTTTGGCCGTGTTTG  
 ATGAAGGCAAAAGCTGGCATTCTGAAACCAAAAACAGCCTGATGCAGGACAGGGATGCCGCCCTCTGCCAGGGCCTGGCCCAAA  
 ATGCATACCGTGAATGGCTATGTGAATAGGAGCCTGCCTGGCCTGATTGGCTGCCACAGGAAAAGCGTGTATTGGCATGTGAT  
 CGGCATGGGACACACCCCGAAGTGCATAGCATCTTTCTGGAAGGCCATACCTTCTGGTCAAGAACACAGGCAGGCCAGCC  
 TGAAATCAGCCCCATCACCTTCTGACGCCCCAGACCTGCTGATGGATCTGGGCCAGTTTCTGCTGTTTTGCCACATCTCC  
 AGCCATCAGCATGATGGCATGGAAGCCTATGTGAAAGTCGATAGCTGCCCGAAGAACCCAGCTGAGGATGAAAAACAATGA  
 AGAAGCCGAAGACTATGATGATGATCTGACTGATTCTGAAATGGATGTGGTCAGGTTTGATGATGATAATAGCCCCAGCTTTA  
 TCCAGATCAGGAGCGTGGCCAAAAACATCCCAAGACCTGGGTGATTATATCGCTGCTGAGGAAGAAGATTGGGACTATGAA  
 CCCCTGGTGTGGCCCCCTGATGATAGGAGCTATAAAAGCCAGTATCTGAACAATGGCCCCAGAGGATTGGCAGGAAGTATAA  
 AAAAGTCAGGTTTATGGCCTACACTGATGAAACCTTCAAGACCAGGGAAGCCATCCAGCATGAGTCTGGCATCCTGGGCCCCC  
 TGCTGTATGGCGAAGTGGGGGACACCCCTGCTGATCATCTTTAAAAATCAGGCCAGCAGGCCCTATAATATCTATCCCCATGGC  
 ATCACTGATGTGAGGCCCCCTGTACAGCAGGAGGCTGCCAAAGGCGTGAAACATCTGAAAGATTTTCCCATCCTGCCTGGCGA  
 AATCTTTAAGTATAAATGGACTGTGACTGTGGAAGATGGCCCCACCAAAAGCGATCCAGGTGCTGACCAGGTATTATTCCA  
 GCTTTGTGAATATGGAACCGCATCTGGCCTCTGGCCTGATTGGCCCCCTGCTGATCTGCTATAAAGAGTCTGTGGACCAGAGG  
 GGCAATCAGATCATGAGCGATAAAAGGAATGTATCCTGTTCTCTGTCTTTGATGAGAATAGGAGCTGGTACCTGACCGAAAA  
 CATCCAGAGGTTTCTGCCCAATCCCGCCGGCGTGACGTGGAAGATCCCGAGTTTCAGGCCAGCAATATCATGCATAGCATCA  
 ATGGCTATGTCTTTGATAGCCTGCAGCTGAGCGTGTGCCTGCATGAGGTGGCCTATTGGTATATCCTGAGCATCGGCGCCAG  
 ACCGATTTTCTGAGCGTGTTTTTCTCTGGCTATACCTTTAAACATAAAATGGTGTATGAGGACACCCCTGACCCTGTTTCCCTT  
 CTCTGGCGAAACCGTGTTTATGAGCATGGAATCCCGCCCTGTGGATCCTGGGCTGCCACAACAGCGATTTCAAGAACAGGG  
 GCATGACTGCCCTGCTGAAAGTCTCCAGCTGCGATAAAACACTGGGACTATTATGAGGACAGCTATGAGGACATCAGCGCC  
 TATCTGCTGAGCAAGAACAATGCCATCGAAGCCAGGAGCTTTAGCCAGAATCCCCAGTGCTGAAAGGCATCAGAGGGAAAT  
 CACCAGGACACCCCTGCAGTCTGATCAGGAAGAAATCGACTATGATGATACCATCAGCGTGGAAATGAAGAAAGAAGATTTTG  
 ATATCTATGATGAAGATGAAATCAGAGCCCCAGGAGCTTTCAGAAGAAAACAGGCATTACTTCATCGCTGCTGTGGAAGG  
 CTGTGGGACTATGGCATGTCCAGCAGCCCCCATGTGCTGAGGAACAGGGCCAGTCTGGCAGCCTGCCCGATTAAAAAAGT  
 CGTGTTCAGGAGTTTACCAGTGGCAGCTTTACCCAGCCCCCTGTATAGGGGCGAACTGAATGAACATCTGGCCTGCTGGGCC  
 CCTACATCAGGGCCGAAGTGAAGATAATATCATGGTGACCTTCAGGAACCAGGCCAGCAGGCCCTACAGCTTTTATTCCAGC  
 CTGATCAGCTATGAGGAAGATCAGAGGCAGGGGGCTGAGCCCAGGAAAACTTTGTGAAACCAATGAAACCAAGACCTACTT  
 TTGGAAGTCCAGCATCATATGGCCCCCACCAGGATGAATTTGATTGCAAGCCCTGGGCCTACTTCTCTGATGTGGACCTGG  
 AAAAGATGTGCATAGCGGCCTGATTGGCCCCCTGCTGGTGTGCCACACCAATACCCTGAACCTGCCATGGCAGGCAGGTG  
 ACTGTGCAGGAGTTTGCCCTGTTCTTTACCATCTTTGATGAAACCAAGCTGGTACTTCACCGAAAACATGGAAGGAAGTGA  
 CAGGGCCCCCTGCAACATCCAGATGGAAGATCCACCTTTAAAGAAAATTATAGGTTCCATGCCATCAATGGCTATATCATGG  
 ATACCTGCTGGCCTGGTGTGATGGCCAGGACAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAATGAAAAACATCCAT  
 AGCATCCATTTCTCTGGCCATGTCTTTACCGTCAGGAAAAAGAAGAGTATAAAATGGCCCTGTATAATCTGTACCCTGGGGT  
 GTTTGAAACCTGGAAATGCTGCCAGCAAAGCCGGCATCTGGAGGTTGGAATGCCTGATTGGCGAACATCTGCATGCTGGCA  
 TGAGCACCTGTTTCTGGTGTATAGCAATAAGTGCCAGACCCCCCTGGGCATGGCCTCTGGCCATATCAGGGATTTTCAGATC  
 ACTGCCCTCTGGCCAGTATGGCCAGTGGGCCCCAACTGGCCAGGCTGCATTATTCGGAAGCATCAATGCCGAGGACCCAA  
 AGAACCTTTAGCTGGATCAAAGTCGATCTGCTGGCCCCCATGATCATCCATGGCATCAAGACCCAGGGGGCCAGGCAGAAAT  
 TTTCCAGCCTGTACATCAGCCAGTTTATCATCATGTATAGCCTGGATGGCAAAAAATGGCAGACCTATAGGGGCAATAGCACC  
 GGCACCTGATGGTGTCTTTGGCAATGTGGACAGCAGCGGCATCAAACATAATATCTTTAATCCCCCATCATCGCCAGGTA  
 TATCAGGCTGCATCCACCCATTATAGCATCAGGAGCACCCTGAGGATGGAACCTGATGGGCTGGCATCTGAACAGCTGCAGCA  
 TGCCCCCTGGGCATGGAAGCAAAGCCATCAGCGATGCCAGATCACTGCCTCCAGCTACTTCACTAATATGTTTGCCACCTGG  
 AGCCCCAGCAAAGCCAGGCTGCATCTGCAGGGCAGGAGCAATGCCTGGAGGCCCCAGGTCAACAATCCCAAGAATGGCTGCA  
 GGTGATTTTCAGAAAACCATGAAAGTCACTGGGGTGACCACCCAGGGGTGAAAGCCTGCTGACCAGCATGTATGTGAAAG  
 AGTTTCTGATCTCCAGCAGCCAGGATGGCCATCAGTGACCTGTCTTTTCAAGATGGCAAAGTCAAAGTCTTTCAAGGGCAAT  
 CAGGACAGCTTTACCCCTGTGGTGAATAGCCTGGATCCCCCCTGCTGACCAGGTATCTGAGGATCCATCCCCAGAGCTGGGT  
 GCATCAGATCGCCTGAGGATGGAAGTGTGGGCTGCGAAGCCAGGACCTGTACTGA

**FIG. 1B: coFVIII-4 – SEQ ID NO: 2**

ATGCAGATCGAGCTGAGCACGTGCTTCTTCTGTGCCTGCTGAGGTTCTGCTTCAGCGCCACCAGGAGGTACTACCTGGGGCGC  
 CGTGGAGCTGAGCTGGGACTACATGCAGAGCGACCTGGGCGAGCTGCCCGTGGACGCCAGGTTCCCCCAGGGTGCCCAAGA  
 GCTTCCCCCTCAACACGAGCGTGGTGTACAAGAAGACCCTGTTCTGAGGTTTACCGACCATCTGTTCAATATCGCCAAGCCC  
 AGGCCCCCTGGATGGGGCTGCTGGGGCCACGATCCAGGCCGAGGTGTACGACACCGTGGTTCATCACCTGAAGAATATGGC  
 CAGCCACCCCGTGGCTGCACGCCGTGGGCGTGAGTACTGGAAGGCCAGCGAGGGCGCCGAGTACGACGACCAGACAGCC  
 AGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCGGACGCCACACCTACGTGTGGCAGGTGCTGAAGGAGAATGGGCCCCATG  
 GCCAGCGACCCCTGTGCCTGACCTACTCTTACCTGAGCCACGTGGATCTGGTGAAGGACCTGAACAGCGGCTGATCGGGCGC  
 CCTGCTGGTGTGCAGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTTCGCCGTGTTTCG  
 ACGAGGGCAAGAGCTGGCAGACGAGACCAAGAACAGCCTGATGCAGGATAGGGACGCCGCCAGCGCCAGGGCCTGGCCCAAG  
 ATGCACACCGTGAACGGCTACGTGAACAGGTCTCTGCCCGCCCTGATCGGCTGCCACAGGAAGAGCGTGTACTGGCACGTGAT  
 CGGCATGGGGACACCCCGAGGTGCACAGCATCTTCTGGAGGGCCACACGTTCTTGGTGAAGAAATCACAGGCAGGCCAGCC  
 TGGAGATCAGCCCGATCACCTTCTGACCGCCAGACCCTGCTGATGGACCTGGGGCAGTTCCTGCTGTTCTGCCATATCAGC  
 TCTCACCAGCAGCAGCGCATGGAGGCCCTACGTGAAGGTGGATAGCTGCCCGAGGAGCCCGAGCTGAGGATGAAGAACAACGA  
 GGAGGCCGAGGACTACGACGACGACCTGACCGACAGCGAGATGGACGTGGTGAAGTTCGACGACGACAATAGCCCGAGCTTCA  
 TCCAGATCAGGAGCGTGGCCAAGAAGCACCCCAAGACCTGGGTGCATTACATCGCCGCCGAGGAGGAGGATTGGGACTACGCC  
 CCCCTGGTGTGGCCCCCGACGACAGGTCTTACAAGAGCCAGTACCTGAACAACGGGGCCCCAGAGGATCGGCAGGAAGTACAA  
 GAAGGTGAGGTTTCATGGCTACACCGACGAGACCTTCAAGACCAGGGAGGGCGATCCAGCACGAGAGCGGGATCCTGGGGCCCC  
 TGCTGTACGGCGAGGTGGGCGACACGCTGCTGATCATCTTCAAGAACAGGCCAGCAGGCCGTACAATATCTACCCCCACGGG  
 ATCACCAGCGTGAGGCCCTGTACTCTAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCTGCCCGGCCGA  
 GATCTTCAAGTACAAGTGGACCGTGACCGTGGAGGACGGGCCACGAAGAGCGACCCAGGTGCCTGACCAGGTACTACAGCT  
 CTTTCTGTAACATGGAGAGGGACCTGGCCAGCGGCCCTGATCGGGCCCCCTGCTGATCTGCTACAAGGAGAGCGTGGATCAGAGG  
 GGCAACCAGATCATGAGCGACAAGAGGAACGTGATCCTGTTCAAGGTGTTTCGACGAGAATAGGTCTTGGTACCTGACCGAGAA  
 TATCCAGAGGTTCTGCCCCAACCCCGCGCGCTGCGAGCTGGAGGATCCCGAGTTCAGGCCAGCAACATCATGCACAGCATCA  
 ACGGCTACGTGTTTCGACAGCCTGCAGCTGAGCGTGTGCCTGCACGAGGTGGCCTACTGGTACATCCTGAGCATCGGCGCCCCAG  
 ACCGACTTCTGAGCGTGTCTTTCAGCGCTACACCTTCAAGCACAAAGATGGTGTACGAGGATACCTTGACCTGTTCCCTT  
 CAGCGGCGAGACCGTGTTCATGAGCATGGAGAACCCGGCCCTGTGGATCCTGGGCTGCCATACTCCGACTTCAGGAATAGGG  
 GCATGACCGCCCTGCTGAAGGTGAGCTCTTGCAGACAGAACACCCGGCGACTACTACGAGGATAGCTACGAGGATATCAGCGCC  
 TACCTGCTGAGCAAGAACAACGCCATCGAGCCAGGTCTTTCAGCCAGAACCCCCCGTGTGAAGAGGCACCAGAGGGAGAT  
 CACCAGGACGACCTGCAGAGCGACCGAGGAGATCGACTACGACGACACGATCAGCGTGGAGATGAAGAAGGAGGATTTCG  
 ACATCTACGACGAGGACGAGAATCAGAGCCCCAGGTCTTCCAGAGAAGAGCCAGGCAATTACTTCATCGCCGCCGTGGAGAGG  
 CTGTGGGACTACGGCATGAGCAGCTCTCCCCACGTGCTGAGGAATAGGGCCCCAGAGCGGCAGCGTGCCTCAGTTCAAGAAGGT  
 GGTGTTCCAGGAGTTCACCGACGGCAGCTTCAACCCAGCCCTGTACAGGGCGGAGCTGAACGAGCACCTGGGCTGCTGGGGC  
 CCTACATCAGGGCCGAGGTGGAGGATAACATCATGGTGACCTTCAGGAATCAGGCCAGCAGGCCCTATAGCTTCTATAGCTCT  
 CTGATCAGCTACGAGGAGGATCAGAGGCAGGGCGCCGAGCCAGGAAGAATTCGTGAAGCCCAACGAGACCAAGACCTACTT  
 CTGGAAGGTGCAGCACCATGCCCCCACGAAGGACGAGTTCGACTGCAAGGCCTGGGCTACTTCAGCGACGTGGATCTGG  
 AGAAGGACGTGCACAGCGGCCCTGATCGGGCCCCCTGCTGGTGTGCCACCAACACCCTGAACCCCGCCACGGCAGGCAGGTG  
 ACCGTGCAGGAGTTCGCCCTGTTCTTCAACATCTTCGACGAGACCAAGAGCTGGTACTTCACCGAGAATATGGAGAGGAATTG  
 CAGGGCCCCCTGCAATATCCAGATGGAGGACCCGACCTTCAAGGAGAATTACAGGTTCCACGCCATCAACGGCTACATCATGG  
 ACACGCTGCCCGGCTGGTTCATGGCCAGGATCAGAGGATCAGGTGGTATCTGCTGAGCATGGGGAGCAACGAGAATATCCAC  
 AGCATCCACTTCAGCGGCCACGTGTTCAACGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAATCTGTACCCCGCGCT  
 GTTCGAGACCGTGGAGATGCTGCCAGCAAGGCCGGGATCTGGAGGCTGGAGTGCCTGATCGGCGAGCACCTGCACGCCGGCA  
 TGAGCACGCTGTTCTGCTGTACTCTAACAAGTGCCAGACCCCTTGGGGATGGCCAGCGGCCACATCAGGGACTTCCAGATC  
 ACCGCCAGCGGCCAGTACGGCCAGTGGGCCCCAAGCTGGCCAGGCTGCATATTCCGGAAGCATCAACGCCCTGGAGCACGAA  
 GGAGCCCTTCAGCTGGATCAAGGTGGATCTGCTGGCCCCATGATCATCCACGGGATCAAGACCCAGGGCGCCAGGCAGAACT  
 TCAGCTCTCTGTATATCAGCCAGTTCATCATATGTAATCTCTGGACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACC  
 GGCACGCTGATGGTGTCTTCGGCAACGTGGACTCTAGCGGGATCAAGCACAAATATCTTCAACCCCCCATCATCGCCAGGTA  
 CATCAGGCTGCACCCACCCATTACTCTATCAGGTCTACCCTGAGGATGGAGCTGATGGGCTGCGACCTGAACAGCTGCAGCA  
 TGCCCCCTGGGATGGAGAGCAAGGCCATCAGCGACGCCAGATCACCGCCAGCTCTTACTTCACCAACATGTTTCGCCACCTGG  
 AGCCCCGAGCAAGGCCAGGTGCACCTGCAGGGCAGGTCTAACGCTGGAGGCCCGAGGTGAACAACCCCAAGGAGTGGCTGCA  
 GGTGGATTTCCAGAAGACCATGAAGGTGACCGGCGTGACACGCGAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGG  
 AGTTCTGATCAGCTCTAGCCAGGACGGCCACAGTGGACCCCTGTTCTTCCAGAACGGCAAGGTGAAGGTGTTCCAGGGCAAC  
 CAGGATAGCTTCAACCCCTGGTGAACAGCCTGGACCCCCCTGCTGACCAGGTATCTGAGGATCCACCCCCAGAGCTGGGT  
 GCACCAGATCGCCCTGAGGATGGAGGTGCTGGGCTGCGAGGCCAGGATCTGTATTGA



**FIG. 1C: coFVIII-5 – SEQ ID NO: 70**

ATGCAAATCGAACTGAGCACCTGTTTCTTCCTCTGCCTGCTGAGATTCTGTTTCTCCGCGACCCGCGGATACTACCTGGGAGCAGTGG  
 AGCTCTCCTGGGATTACATGCAGAGCGACCTTGGGGAGCTGCCGTGGATGCCAGGTTCCCTCCCCGGGTGCCAAAGTCGTTTCCGTT  
 CAACACCTCCGTGGTGTACAAAGAACTCTGTTCTGTGGAGTTACCGACCACCTGTTCAATATCGCCAAGCCCAGACCTCCCTGGATG  
 GGGCTGTTGGGACCTACCATCCAAGCGAGGTGTACGACACTGTGGTCATCACTCTGAAGAATATGGCCTCGCATCCCGTGTCCCTGC  
 ACGCCGTGGGAGTGTCTTACTGGAAAGCGTCCGAGGGGGCCGAATACGACGACCAGACCTCGCAGAGAGAAAAGGAAGATGACAAGGT  
 GTTCCCAGGAGGATCGCACACCTACGTGTGGCAAGTGTGAAGGAGAACGGCCCAATGGCCTCCGACCCGCTGTGCCTGACCTACTCG  
 TACCTGTCCCACGTGGACCTCGTGAAGGACCTCAACTCGGACTGATTGGAGCCCTGCTGGTCTGCAGGGAAGGCTCACTGGCGAAAAG  
 AAAAGACTCAGACCTTGCACAAGTTCATTCTGCTGTTCTGCTGTGTTGACGAGGGGAAGTCGTGGCACAGCGAGACTAAGAATCCCT  
 GATGCAAGATAGAGATGCCGCTCCGCCCCGGCCTGGCCTAAGATGCACACCGTGAACGGTTACGTGAACCGCTCCCTCCCTGGCCTG  
 ATTGGATGCCACCGGAAGTCCGTGTACTGGCACGTGATCGGGATGGGACCACCCCGAGGTGCACAGCATCTTCTTGGAAAGTCA  
 CATTTCTCGTGGCAACACCGGCGAGCCTCCCTGGAAATCAGCCCCATTACCTTCTCACTGCCAGACTCTGCTGATGGACCTGGG  
 ACAGTTCCTGCTGTCTGCCATATCTCCTCCCAACATGACGGAATGGAGGCATACGTGAAGGTCGATTCTGCCCTGAGGAACCC  
 CAGCTCCGCATGAAGAACAATGAGGAAGCCGAGGACTACGACGACGACCTGACGGATAGCGAGATGGATGTGGTCCGGTTCGATGACG  
 ATAACAGCCCTTCTTTCATCCAAATTGCTCGGTGGCAAGAAGCACCCCAAGACCTGGGTGCATTACATTGCGCGGAAGAAGAGGA  
 CTGGGATTATGCCCGCTTGTCTCGTCTGACGACCGGAGCTACAAGAGCCAGTACCTGAACAACGGTCCACAGAGGATCGGTAGA  
 AAGTACAAGAGGTCCGCTTCATGGCTATACGACGAAACCTTCAAACTAGAGAGGCCATCCAACACGAATCCGGCATCTCGGGCC  
 CGCTCTGTACGGAGAAGTCGGCGACACCTTCTCATTATCTTCAAGAACCAGGCTTCCCGGCGGTACAACATCTATCCGCATGGGAT  
 CACTGACGTGGCCCACTGTACTCGCGGCGCTGCCAAGGGTGTCAAACACCTGAAGGATTTTCCGATCTTCCGGGAGAAATCTTC  
 AAGTACAAGTGGACCGTGACCGTGAAGATGGCCCACTAAGTCTGACCCTAGATGCCCTACCCGCTACTACTCATCTTCTCGTCAACA  
 TGGAGCGCGACCTGGCCAGCGGACTGATCGGCCGCTGCTGATTGCTACAAGGAATCAGTGGACCAACGGGGAAACAGATCATGTC  
 GGATAAGAGGAACGTCATCTCTTCTCCGTGTTTGACGAAAACCGGTGCTGGTACCTGACTGAAAACATCCAGCGGTTCTCCCAAC  
 CCCGCGGGCGTGCAGCTGGAAGATCTGAGTTTCAGGCATCAAACATCATGCACTCCATTAACGGCTACGTGTTCCGATTCGCTGCAGC  
 TGAGCGTGTGCTGCACGAAGTGGCCTACTGGTACATCTGTCCATTGGTGCCAGACTGACTTCTGTCCGTGTTTTCTCCGGCTA  
 CAGTTCAAGCACAAAGATGGTGTACGAGGACACCTGACCTCTTCCCTTTTCCGGCGAACTGTGTTTATGAGCATGGAGAATCCC  
 GGCTGTGGATCTTGGGCTGCCACAACAGCGACTTCCGTAAACAGAGGAATGACTGCGCTGCTCAAGGTGTCCAGCTGCGACAGAACA  
 CCGGAGACTATTATGAGGACTCATACGAGGACATCTCCGCTACCTCCTGTCCAAGAATAACGCCATTGAACCTCGGAGCTTCAGCCA  
 GAACCCACCCGTGCTTAAGAGACATCAACGGGAGATCACTAGGACCACCTGCAGTCAGACCAGGAGGAAATCGACTACGATGACACC  
 ATCTCGGTGAGATGAAGAAGGAGGACTTTGACATCTACGACGAAGATGAAAACAGAGCCCGAGGTGCTTCCAAAAGAAAACCCGCC  
 ACTACTTTATTGCTGCTGTGAGCGGCTGTGGGACTACGGAATGTCGTCTCGCCGACGTCGTCGCCAACCAGCCAGCCAGAGCGGCTC  
 GGTGCCGCAATTCAAGAAGTCTGTGTTCCAGGAGTTCACTGACGGGAGCTTCACTCAGCCTTTGTACCGGGGAGAACTCAATGAACAT  
 CTCGGCCTCCTCGGACCTTACATCAGAGCAGAAGTGGAAAGATAACATCATGGTCACTTTCGTAACCAAGCCAGCCGCGCTACTCGT  
 TCTACTCTCCCTCATTTCTTACGAAGAGGACCAGCGGCGAGGCGCAGAACCAGCGCAAGAACTTCGTGAAGCCCAACGAAACCAAGAC  
 CTACTTCTGGAAGTGCAGCATCATATGGCCCCGACTAAGGACGAGTTTACTGCAAAGCCTGGGCCTACTTCTCCGATGTGGACTTG  
 GAGAAGGACGTCCACTCCGGCCTCATCGGTCCCTGCTCGTGTGCCATACCAATACCTGAACCCCGCACAGGTGCGCAGGTCAACG  
 TGAGGAGTTGCTCTGTTCTTCACTATCTTCGACGAACTAAGTCTGGTACTTCAACGAGAACATGGAGAGGAACTGCAGAGCCCC  
 CTGTAACATCCAGATGGAGGACCCGACGTTCAGGGAACCTACCGGTTCCACGCCATTAAACGGATACATCATGGATACGCTGCCGGGT  
 CTTGTGATGGCCAGGATCAACGGATCAGATGGTACTTATTGTGATGGGAGCAACGAGAATCCACTCTATTCACTTCTCCGGTC  
 ATGTGTTCACTGTGCGGAAGAAGGAAGGTACAAGATGGCCCTGTACAACCTTTATCCCGGAGTGTTCGAAACTGTGGAATGCTGCC  
 GTCGAAGGCCGGCATTGCGCGGTGGAGTGTGATTGGAGAATCTCCATGCGGGGATGTCAACCTGTTCTTGGTGTATAGCAAC  
 AAGTGCCAGACTCCGCTTGGGATGGCGTCAGGACACATTAGGGATTTCCAGATCACTGCGTCCGGCCAGTACGGCCAAATGGGCCCTA  
 AGCTGGCCCGCCTGCATTACTCCGATCCATTAAACGCTGGTCAACCAAGGAGCCATTCTCCTGGATCAAGGTGGACCTTCTGGCCCC  
 CATGATTATCCACGGAATTAAGACCCAGGGGGCCCGGCAGAACTTCTCCTCACTGTACATCAGCCAGTTTCAATCATGTACTCCCTG  
 GACGGAAGAAGTGGCAAACCTACAGGGGGAACAGCACCGGCACACTGATGGTCTTTTTCGGAATGTGGACTCCTCCGGGATTAAGC  
 ATAACATCTTCAACCTCCGATTATCGCTCGGTACATTAGACTTCACCTACCCACTACAGCATTCGCTCCACCTGCGGATGGAAC  
 GATGGGCTGCGATCTGAACCTGTCAGCATGCCGTGGGAATGGAGTCCAAGCAATTTCCGACGCGCAGATCACCGCTCTGCTTAC  
 TTTACCAACATGTTGCGCACGTGGTCAACGTCAGGCGCGGCTGCACCTCCAGGGAAGATCCAACGCATGGCGGCCACAGGTCAACA  
 ACCCTAAGGAGTGGCTCCAGGTGGACTTCAGAAAACCATGAAGTCAACCGGAGTCAACACCGAGGAGTGAAGTCTGCTGCTGACTTC  
 TATGTACGTCAAGGAGTTCCTGATCTCCAGCAGCCAGGACGGCCAGGTGGACCTGTTCTTCCAAATGGAAGGTCAAGGTGTTT  
 CAGGGCAATCAGGATTCATTACCCCGGTGGTGAATCCCTGATCCACCCCTCCTGACCCGCTACCTTCGCATCCACCCACAGTCTT  
 GGGTGCACAGATCGCGCTGAGGATGGAGGTCTGGGATGCGAAGCCAGGACCTGTACTGA

**FIG. 1D : coFVIII-6 – SEQ ID NO: 71**

ATGCAGATTGAGCTGTCCACTTGTCTTCTTCTGTGCTCCTGCGCTTCTGTTTCTCCGCCACTCGCCGGTACTACCTTGGAGCCGTGG  
 AGCTTTCATGGGACTACATGCAGAGCGACCTGGGCGAACTCCCCGTGGATGCCAGATTCCCCCCCCCGCTGCCAAAGTCCCTTCCCTT  
 TAACACCTCCGTGGTGTACAAGAAAACCTCTTTGTGCGATTCACTGACCACCTGTTCAACATCGCCAAGCCGCGCCACCTTGGATG  
 GGCCCTCCTGGGACCGACCATTCAGCTGAAGTGTACGACACCGTGGTGATCACCTGAAGAACATGGCGTCCACCCCGTGTCCCTGC  
 ATGCGGTGCGGAGTGTCTACTGGAAGGCCCTCCGAAGGAGCTGAGTACGACGACCAGACTAGCCAGCGGGAAAGGAGGACGATAAAGT  
 GTTCCCGGGCGGCTCGCATACTTACGTGTGGCAAGTCTGAAGGAAAACGGACCTATGGCATCCGATCCTCTGTGCTGACTTACTCC  
 TACCTTTCCTATGTGGACCTCGTGAAGGACCTGAACAGCGGGCTGATTGGTGCACTTCTCGTGTGCGCGAAGGTTCCGCTCGCTAAGG  
 AAAAGACCCAGACCTCCATAAGTTCATCTTTTGTGCTGTGTTCTGATGAAGGAAAGTCATGGCATTCCGAAACTAAGAACTCGCT  
 GATGCAGGACCGGGATGCCGCTCAGCCGCGCCTGGCCTAAATGCATACAGTCAACGGATACGTGAATCGGTCACTGCCCGGGCTC  
 ATCGGTTGTACAGAAAGTCCGTGTACTGGCAGCTCATCGGCATGGGCACTACGCTGAAGTGCATCCATCTTCTGGAAGGGCACA  
 CTTCTCTGTGCGCAACACCGCCAGGCCCTCTCTGGAATCTCCCGATTACCTTCTGACCGCCAGACTCTGCTCATGGACCTGGG  
 GCAGTTCCTTCTCTTCTGCCACATCTCCAGCCATCAGCAGCAGCGAATGGAGGCCCTACGTGAAGGTGGACTCATGCCCGGAAGAACCT  
 CAGTTGCGGATGAAGAACAACGAGGAGGCCGAGGACTATGACGACGATTTGACTGACTCCGAGATGGACGTCGTGCGGTTTCGATGACG  
 ACAACAGCCCCAGCTTCATCCAGATTGCGAGCGTGGCCAAGAAGCACCCAAACCTGGGTGCACTACATCGCGCGCGAGGAAGAAGA  
 TTGGGACTACGCCCCGTTGGTGTGGCACCCGATGACCGGTCTGACAAAGTCCAGTATCTGAACAATGGTCCGAGCGGATTGGCAGA  
 AAGTACAAGAAAGTGGCGTTTCATGGCGTACACTGACGAAACGTTTAAAGACCCGGGAGGCCATTCAACATGAGAGCGGCATTCTGGGAC  
 CACTGCTGTACGGAGAGGTGGCGGATACCTGCTCATCATCTTCAAAAACAGGCCCTCCCGGCTTACAACATCTACCTCACGGAAT  
 CACCGACGTGCGGCCACTCTACTCGCGCGCCTGCCGAAGGGCGTCAAGCACCTGAAAGACTTCCCTATCTTCCGGGGCGAAATCTTC  
 AAGTATAAGTGGACCGTCACCGTGGAGGACGGGCCACCAAGAGCGATCCTAGGTGTCTGACTCGGTACTACTCCAGCTTCGTGAACA  
 TGGAACGGGACCTGGCATCGGACTCATTTGACCGCTGCTGATCTGCTACAAAGAGTGGGTGGATCAACGCGGCAACAGATCATGTC  
 CGACAAGCGCAACGTGATCCTGTTCTCCGTGTTTGTGATAAAGAGATCCTGGTACCTCACTGAAAACATCCAGAGGTTCTCCCAAAC  
 CCGCAGGAGTGAACCTGGAGGACCTGAGTTTACGGCCTCGAATATCATGCACTCGATTAACGGTTACGTGTGACTCGCTGCAGC  
 TGAGCGTGTGCTCCATGAAGTCTGCTTACTGGTACATTCTGTCCATCGCGCGCCAGACTGACTTCTGAGCGTGTCTTTTCCGGTTA  
 CACCTTTAAGCACAAGATGGTGTACGAAGATACCTGACCTGTTCCCTTCTCCGGCGAAACGGTGTTCATGTGATGGAGAACCCG  
 GGTCTGTGGATTCTGGGATGCCACAACAGCGACTTTCGGAACCGCGGAATGACTGCCCTGCTGAAGGTGTCTCATGCGACAAGAACA  
 CCGGAGACTACTACGAGGACTCTACGAGGATATCTCAGCCTACCTCTGTCCAAGAACAACGCGATCGAGCGCGCAGCTTCAGCCA  
 GAACCCGCTGTGCTGAAGAGGACCCAGCGAGAAATTACCCGGACCCCTCCAATCGGATCAGGAGGAAATCGACTACGACGACACC  
 ATCTCGGTGGAATGAAGAAGGAGATTTCGATATCTACGACGAGGACGAAATCAGTCCCTCGCTCATTCCAAAAGAAAACAGAC  
 ACTACTTTATCGCGCGGTGGAAGACTGTGGGACTATGGAATGTATCCAGCCCTCACGTCTTCGGAACCGGGGCCAGAGCGGATC  
 GGTGCCCTCAGTTCAAGAAAGTGGTGTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCGCTGTACCGGGGAGAACTGAACGAACAC  
 CTGGGCTGCTCGGTCTTACATCCGCGCGGAAGTGGAGGATAACATCATGGTGACCTTCGTAACCAAGCATCCAGACCTTACTCCT  
 TCTATTCTCCTGATCTCATACGAGGAGGACAGCGCCAAGGCGCCGAGCCCCGCAAGAACTTCGTCAAGCCCAACGAGACTAAGAC  
 CTACTTCTGGAAGGTCCAACACCATATGGCCCCGACCAAGGATGAGTTTACTGCAAGGCTGGGCTACTTCTCCGACGTGGACCTT  
 GAGAAGGATGTCCATTCCGGCTGATCGGGCCGCTGCTCGTGTGTACACCAACACCTGAACCCAGCGCATGGACGCCAGGTACCG  
 TCCAGGAGTTTGTCTGTTCTTACCATTCTTACGAAACTAAGTCTTGGTACTTCACCGAGAATATGGAGCGAACTGTAGAGCGCC  
 CTGCAATATCCAGATGGAAGATCCGACTTTCAGGAGAACTATAGATTCCACGCCATCAACGGGTACATCATGGATACTTCTCCGGGG  
 CTGGTCATGGCCAGGATCAGAGGATTCGGTGGTACTTGTGTCAATGGGATCGAACGAAAACATTCACTCCATTCACTTCTCCGGTC  
 ACGTGTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCGCTGTACAATCTGTACCCCGGGGTGTTCGAAACTGTGGAGATGCTGCC  
 GTCCAAGGCCGGCATCTGGAGAGTGGAGTGCCTGATCGGAGAGCACCTCCACGCGGGGATGTCCACCTCTTCTGGTGTACTCGAAT  
 AAGTGCAGACCCCGTGGGCATGGCCTCGGGCCACATCAGAGACTTCAGATCACAGCAAGCGGACAATACGGCCAATGGGCGCCGA  
 AGCTGGCCCGCTTGCATACTCCGGATCGATCAACGCATGGTCCACCAAGGAACCGTCTCGTGGATTAAAGTGGACCTCCTGGCCCC  
 TATGATTATCCACGGAATTAAGACCCAGGGCGCCAGGCAGAGTTCTCTCCCTGTACATCTCGCAATTTCATCATGTACAGCCTG  
 GACGGGAAGAAGTGGCAGACTTACAGGGGAACTCCACCGGCACCTGATGGTCTTTTTCGGCAACGTGGATTCTCCGGCATTAAGC  
 ACAACATCTTCAACCCACCGATCATAGCCAGATATATTAGGCTCCACCCCACTCACTACTCAATCCGCTCACTCTTCGGATGGAAC  
 CATGGGGTGGACCTGAACTCTGCTCCATGCCGTTGGGGATGGAATCAAAGGCTATTAGCGACGCCCAGATCACCGCGAGCTCCTAC  
 TTCACTAACATGTTTCGCCACCTGGAGCCCTCCAAGGCCAGGCTGCACTTGCAGGGACGGTCAAATGCCTGGCGGCCGCAAGTGAACA  
 ATCCGAAGGAATGGCTTCAAGTGGATTTCAAAAGACCATGAAAGTGACCGGAGTCAACACCCAGGGAGTGAAGTCCCTTCTGACCTC  
 GATGTATGTGAAGGAGTCTGATTAGCAGCAGCCAGGACGGGCACCAAGTGGACCTGTTCTTCCAAAACGGAAAGGTCAAGGTGTTT  
 CAGGGAAACCGGACTGTTTCAACCCGTTGGTGAACCTCTGGACCCCACTGCTGACGCGGTACTTGAGGATTTCATCTCAGTCTT  
 GGGTCCATCAGATTGCATTCGAATGGAAGTCTGGGCTGCGAGGCCAGGACCTGTACTGA

**FIG. 1E: coFVIII-52 – SEQ ID NO: 3**

ATGCAAATCGAACTGAGCACCTGTTTCTTCTCTGCTGCTGAGATTCTGTTTCTCCGCCACCCGCCGATACTACCTGGGAGC  
 AGTGGAGCTCTCTGGGATTACATGCAGAGCGACCTTGGGGAGCTGCCCGTGGATGCCAGGTTCCCTCCCCGGGTGCCAAAGT  
 CGTTTCCGTTCAACACCTCCGTGGTGTACAGAAACTCTGTTTCGTGGAGTTACCCGACCACCTGTTCAATATCGCCAAGCCC  
 AGACCTCCCTGGATGGGGCTGTGGGACCTACCATCCAAGCGGAGGTGTACGACACTGTGGTCATCACTCTGAAGAACATGGC  
 CTCGCATCCCGTGTCCCTGCACGCCGTGGGAGTGTCTTACTGGAAAGCGTCCGAGGGGGCCGAATACGACGACACGACCTCGC  
 AGAGAGAAAAGGAAGATGACAAGGTGTTCCAGGAGGATCGCACACCTACGTGTGGCAAGTGTGAAGGAGAACGGCCCAATG  
 GCCTCCGACCCGCTGTGCCTGACCTACTCGTACCTGTCCACGTGGACCTCGTGAAGGACCTCAACTCGGGACTGATTGGAGC  
 CCTGCTGGTCTGCAGGGAAGGCTCACTGGCGAAAGAAAGACTCAGACCTTGACAAGTTCATTCTGCTGTTGCTGTGTTCCG  
 ACGAGGGGAAGTCTGGCACAGCGAGACTAAGAACTCCCTGATGCAAGATAGAGATGCCGCTCCGCCCGGGCTGGCCTAAG  
 ATGCACACCGTGAACGGTTAGCTGAACCGCTCCCTCCCTGGCCTGATTGGATGCCACCGGAAGTCCGTGTACTGGCACGTGAT  
 CGGGATGGGGACCAACCCCGAGGTGCACAGCATCTTCTGGAAGGTACACATTTCTCGTGCGCAACCAACCGGCAGGCCTCCC  
 TGGAAATCAGCCCCATTACCTTCTCACTGCCAGACTCTGCTGATGGACCTGGGACAGTTCCTGCTGTTCTGCCATATCTCC  
 TCCCACCAACATGACGGAATGGAGGCATACGTGAAGGTGATTCTGCCCTGAGGAACCCAGCTCCGCATGAAGAACATGA  
 GGAAGCCGAGGACTACGACGACGACCTGACGGATAGCGAGATGGATGTGGTCCGGTTCGATGACGATAACAGCCCTTCTCTCA  
 TCCAAATTCGCTCGGTGGCAAGAAGCACCCCAAGACCTGGGTGCATTACATTGCGGCGGAAGAAGAGGACTGGGATTATGCC  
 CCGCTTGTCTCGCTCCTGACGACCGGAGCTACAAGAGCCAGTACCTGAACAACGGTCCACAGAGGATCGGTAGAAAGTACAA  
 GAAGGTCCGCTTCATGGCCTATACCGACGAAACCTTCAAACTAGAGAGGCCATCAACACGAATCCGGCATCTGGGCCCCG  
 TCTGTACGGAGAAGTCCGGCAGACCCCTTCTATTATCTCAAGAACCAGGCTTCCCGCCGTACAACATCTATCCGCATGGG  
 ATCACTGACGTGCGCCCACTGTACTCGCGCGCCTGCCCAAGGTTGTCAAACACCTGAAGGATTTCCGATCCTTCCGGGAGA  
 AATCTTCAAGTACAAGTGGACCGTGACCGTGAAGATGGCCCAACTAAGTCTGACCTAGATGCCCTACCCGCTACTACTCAT  
 CCTTCGTCAACATGGAGCCGACCTGGCCAGCGGACTGATCGGCCCGCTGCTGATTGCTACAAGGAATCAGTGGACCAACGG  
 GGAACCCAGATCATGTCCGATAAGAGGAACGTATCCTCTTCTCCGTGTTTGACGAAAACCGGTCTGGGTACCTGACCGAGAA  
 CATCCAGAGGTTTCTGCCCAACCTGCTGGGGTGCAGCTGGAGGACCCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCA  
 ATGGCTACGTGTTTCGACAGCTTCGAGCTGAGCGTGTGCTTCGACGAGGTGGCCTACTGGTACATCTTGAGCATCGGCGCCAG  
 ACCGACTTCTGAGCGTGTCTTCTCTGGCTACACCTCAAGCACAAAGTGGTGTATGAGGACACCTGACCTGTTTCCCCTT  
 CAGCGGGGAGACTGTCTTTCATGAGCATGGAGAACCCTGGCCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGG  
 GCATGACTGCCCTGCTGAAAGTCTCCAGCTGTGACAAGAACACCGGGGACTACTACGAGGACAGCTACGAGGACATCAGCGCC  
 TACTGCTGAGCAAGAACAATGCCATCGAGCCAGGAGCTTCTCTCAGAACCCCAAGTGTCTGAAGAGGCACCGAGGGAGAT  
 CACGAGGACACCTGCACTGTGACGAGGAGGATCGACTATGATGACACCATCAGCGTGGAGATGAAGAAGGAGGACTTCG  
 ACATCTACGACGAGGACGAGAACAGAGCCCGAGGAGCTTCCAGAAGAGAGGACGAGCACTACTTATTGCTGCTGTGGAGAGG  
 CTGTGGGACTATGGCATGTCCAGCAGCCCCATGTGCTGAGGAACAGGGCCAGTCTGGCAGCGTGGCCAGTTCAGGAAGT  
 CGTGTTCAGGAGTTACCGACGGCAGCTTACCCAGCCCTGTACAGAGGGGAGCTGAACGAGCACCTGGGCTGTCTGGGCC  
 CTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGAACCTTACGGAACAGGGCCAGCAGGCCCACAGCTTCTACAGCAGC  
 CTGATCAGCTACGAGGAGGACAGAGGACGGGGGCTGAGCCAGGAAGAACTTTGTGAAGCCCAATGAAGCCAGACCTACTT  
 CTGGAAGGTGCAGCACCATGCCCCCAAGGACGAGTTCGACTGCAAGGCTGGGCCCTACTTCTCTGACGTGGACCTGG  
 AGAAGGACGTGCACTCTGGCCTGATTGGCCCTGCTGGTGTGTCACACCAACACCTGAACCTGCCATGGCAGGAGGAGT  
 ACTGTGAGGAGTTCGCCCTGTTCTTTCACCATCTTCGATGAAGCAAGAGCTGGTACTTCACTGAGAACATGGAGAGGAAGT  
 CAGGGCCCCCTGCAACATCCAGATGGAGGACCCCACTTCAAGGAGAACTACAGGTCCATGCCATCAATGGCTACATCATGG  
 ACACCTGCTGCGCTGCTGATGGCCAGGACAGAGGATCAGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAACATCCAC  
 AGCATCCACTTCTTGGCCAGCTGTTCACTGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAACCTGTACCTGGGGT  
 GTTCGAAACCGTGGAGATGTGCCAGCAAGGCCGGCATCTGGAGGTTGGAGTGCCTGATTGGGGAGCACCTGCACGCCGGCA  
 TGAGCACCTGTCTCTGGTGTACAGCAACAAGTGCCAGACCCCTGGGCATGGCCTCTGGCCACATCAGGGACTTCCAGATC  
 ACTGCCTCTGGCCAGTACGGCCAGTGGGCCCAAGCTGGCCAGGCTGCACTACTCCGGAAGCATCAATGCCCTGGAGCACCA  
 GGAGCCCTTCAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCACGGCATCAAGACCCAGGGGGCCAGGCAGAAGT  
 TCTCCAGCCTGTACATCAGCCAGTTTATCATCATGTACAGCCTGGACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACC  
 GGCACCTGTATGGTGTCTTCCGCAACGTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCATCATCGCCAGATA  
 CATCAGGCTGCACCCACCCACTACAGCATCAGGAGCACCTGAGGATGGAGCTGATGGGCTGTGACCTGAACAGCTGCAGCA  
 TGCCCTTGGGCATGGAGAGCAAGGCCATCTCTGACGCCAGATCACTGCCTCCAGTACTTACCACATGTTTGGCCACTGG  
 AGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGCAATGCCTGGAGGGCCCCAGGTCAACAACCCCAAGGAGTGGCTGCA  
 GGTGGACTTCCAGAAGACCATGAAGGTGACTGGGGTGACCAACCCAGGGGGTGAAGAGCCTGCTGACCAAGCATGTACGTGAAGG  
 AGTTCCTGATCTCCAGCAGCCAGGACGGCCACCACTGGACCTGTTCTTCCAGAATGGCAAGGTGAAGGTGTTCCAGGGCAAC  
 CAGGACAGCTTCAACCCCTGAGGTCAACAGCCTGGACCCCTGCTGACCATACCTGAGGATCCACCCAGAGCTGGGT  
 GCACCAAGTCCGCTGAGGATGGAGGTGTGGGCTGTGAGGCCAGGACCTGTACTGA

Subrayado continuo = coFVIII-52-NT58 (SEQ ID NO: 7)

Subrayado discontinuo = coFVIII-52-CT (SEQ ID NO: 8)



**FIG. 1F: coFVIII-62 – SEQ ID NO: 4**

ATGCAGATTGAGCTGTCCACTTGTTCCTTCTGCTCCTGCGCTTCTGTTTCTCCGCCACTCGCCGGTACTACCTTGGAGC  
 CGTGGAGCTTTCATGGGACTACATGCAGAGCGACCTGGGCGAATCCCCGTGGATGCCAGATTCCCCCCCCCGCTGCCAAAGT  
 CCTTCCCCCTTAACACCTCCGTGGTGTACAAGAAAACCTCTTTGTGAGTTCCTGACCACCTGTTCAACATCGCCAAGCCG  
 CGCCACCTTGGATGGGCTCCTGGGACCGACCATCAAGCTGAAGTGTACGACACCGTGGTGTACCCCTGAAGAACATGGC  
 GTCCACCCCGTGTCCCTGCATGCGGTGCGAGTGTCTACTGGAAGGCTCCGAAGGAGCTGAGTACGACGACGACAGCTAGCC  
 AGCGGGAAAAGGAGGACGATAAAGTGTTCGGGGCGGCTCGCATACTTACGTGTGGCAAGTCTGAAGGAAAACGGACCTATG  
 GCATCCGATCCTCTGTGCTGACTTACTCTACCTTTCCCATGTGACCTCGTGAAGGACCTGAACAGCGGGCTGATTGTTGC  
 ACTTCTCGTGTGCGCGAAGGTTGCTCGCTAAGGAAAAGACCCAGACCCTCCATAAGTTCATCCTTTTGTTCGCTGTGTTCCG  
 ATGAAGGAAAGTCATGGCATTCCGAACCTAAGAACTCGCTGATGCAGGACCGGGATGCCGCTCAGCCCGCGCTGGCTAAA  
 ATGCATACAGTCAACGGATACGTGAATCGGTCACTGCCCCGGCTCATCGGTTGTACAGAAAGTCCGTGTACTGGCACGTAT  
 CGGCATGGGCACTACGCTGAAGTGCATCCATCTTCTGGAAGGGCACACCTTCTCGTGCACAACACCGCCAGGCCCTCTC  
 TGGAAATCTCCCCGATTACCTTTCTGACCGCCAGACTCTGCTCATGGACCTGGGGCAGTTCCTTCTTCTGCCACATCTCC  
 AGCCATCAGCAGCAGGAATGGAGGCTACGTGAAGTGGACTCATGCCGGAAGAACCTCAGTTGCGGATGAAGAACAACGA  
 GGAGGCCGAGGACTATGACGACGATTTGACTGACTCCGAGATGGACCTCGTGCCTTTCGATGACGACAACAGCCCGAGCTTCA  
 TCCAGATTGCGAGCGTGGCCAAGAAGCACCCCAAAACCTGGGTGCACTACATCGCGGCCGAGGAAGAAGATTGGGACTACGCC  
 CCGTTGGTGTGGCACCCGATGACCGGTGCTACAAGTCCAGTATCTGAACAATGGTCCGACGCGGATTGGCAGAAAGTACAA  
 GAAAGTGGGTTTCATGGCTACACTGACGAAACGTTTAAAGCCCGGAGGCCATTCAACATGAGAGCGGCATTCTGGGACCAC  
 TGCTGTACGGAGAGGTGCGCGATACCTGCTCATCATCTTCAAAAACAGGCCCTCCGCGCTTACAACATCTACCCTCACGGA  
 ATCACCGACGTGCGGCCACTCTACTCGCGCGCCTGCCGAAGGGCGTCAAGCACCTGAAAGACTTCCCTATCTGCGGGCGGA  
 AATCTTCAAGTATAAGTGGACCGTCACCGTGGAGGACGGGCCACCAAGAGCGATCTAGGTGTCTGACTCGGTACTACTCCA  
 GCTTCTGTAACATGGAACGGGACCTGGCATCGGACTCATTTGACCGCTGCTGATCTGCTACAAAGATCGGTGGATCAACGC  
 GGCAACCGATCATGTCCGACAAGCGCAACGTGATCTGTTCTCCGTGTTTGGATGAAAACAGATCCTGGTACCTGACCGAGAA  
 CATCCAGAGGTTCTGCCCCAACCTGCTGGGGTGCACTGGAGGACCCCGAGTTCAGGCCAGCAACATCATGCACAGCATCA  
 ATGGCTACGTGTTTGACAGCTGACGCTGAGCGTGTGCTGACAGAGGTGGCTACTGGTACATCTTGAGCATCGGCGCCAG  
 ACCGACTTCTGAGCGTGTCTTCTCTGCTGCTACACCTTCAAGCACAGATGGTGTATGAGGACACCTGACCTGTTCCCTTT  
 CAGCGGGGAGACTGTCTTTCATGAGCATGGAGAACCTTGGCTGTGGATCCTGGGCTGCCACAACAGCGACTTCAGGAACAGGG  
 GCATGACTGCTGCTGAAAGTCTCCAGCTGTGACAGAAACACCGGGGACTACTACGAGGACAGCTACGAGGACATCAGCGCC  
 TACCTGCTGAGCAAGAACAATGCCATCGAGCCAGGAGCTTCTCTCAGAACCCCAAGTGTGAGAGGACACGAGGGAGAT  
 CACGAGGACACCTTGCAGTCTGACGAGGAGGATGCTGATGATGACACCATCAGCGTGGAGATGAAGAAGGAGGACTTCG  
 ACATCTACGACGAGGACGAGAACAGAGCCCCAGGAGCTTCCAGAGAAGACAGGCACTACTTTCATTGCTGCTGTGGAGAGG  
 CTGTGGGACTATGGCATGTCCAGCAGCCCCATGTCTGAGGAACAGGGCCAGTCTGGCAGCGTGGCCAGTTCAAGAAAGT  
 CGTGTTCAGGAGTTACCGACGGCAGCTTACCCAGCCCCCTGTACAGAGGGGAGCTGAACGAGCACCTGGGCTGCTGGGCC  
 CCTACATCAGGGCCGAGGTGGAGGACAACATCATGTTGACCTTCAGGAACAGGCGAGCAGGCCCTACAGCTTCTACAGCAGC  
 CTGATCAGCTACGAGGAGGACAGAGGCGAGGGGGCTGAGCCAGGAAGAACTTTGTGAAGCCCAATGAACCAAGACCTACTT  
 CTGGAAGGTGACGACACATGGCCCCACCAAGGACGAGTTCGACTGCAAGGCTGGGCTTACTTCTCTGACGTGGACCTGG  
 AGAAGGACGTGCACTCTGGCTGATTGGCCCCCTGTGCTGTTGTCACACCAACACCTGAACCTGCCATGGCAGGCAAGTG  
 ACTGTGACGAGGTTGCGCTGTTCTTCCATCTTCCATGAAACCAAGAGCTGGTACTTCACTGAGAACATGGAGAGGAACTG  
 CAGGGCCCCCTGCAACATCCAGATGGAGGACCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGG  
 ACACCTGCTGCGCTGGCTGCTATGGCCCCAGGACAGAGGATCAGGTGGTATCTGCTGAGCATGGGCGACCAACGAGAATCCAC  
 AGCATCCACTTCTCTGGCCACGTGTTCACTGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAACCTGTACCCTGGGGT  
 GTTCGAAACCGTGGAGATGCTGCCAGCAAGGCCGGCATCTGGAGGGTGGAGTGCCTGATTGGGGAGCACCTGCACGCCGGCA  
 TGAGCACCTGTTCTGCTGTACAGCAACAAGTGCCAGACCCCCCTGGGCATGGCTCTGGCCACATCAGGAGCTTCCAGATC  
 ACTGCTCTGCGCAGTACGGCAGTGGGCCCCAAGCTGGCCAGGCTGCACTACTCCGGAAGCATCAATGCTTGGAGCACCAA  
 GGAGCCCTTCACTGGATCAAGTGGACCTGCTGGCCCCCATGATCATCCACGGCATCAAGACCCAGGGGGCAGGCAAGT  
 TCTCCAGCTGTACATCAGCCAGTTTCATCATGTCATGTCAGCCTGGACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACC  
 GGCACCTGATGGTGTCTTCCGGCAACGTGGACAGCAGCGGCATCAAGCACAACATCTTCAACCCCCCATCATCGCCAGATA  
 CATCAGGCTGCACCCACCCACTACAGCATCAGGAGCACCCTGAGGATGGAGCTGATGGGCTGTGACCTGAACAGCTGCAGCA  
 TGCCCCCTGGGATGGAGAGCAAGGCCATCTCTGACGCCAGATCACTGCCTCCAGCTACTTCAACAACATGTTTGCCACCTGG  
 AGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGCAATGCCCTGGAGGGCCAGGTCAACAACCCCAAGGAGTGGCTGCA  
 GGTGGACTTCCAGAAGACCATGAAGGTGACTGGGGTGACACCCAGGGGGTGAAGAGCCTGCTGACGAGCATGTACGTGAAGG  
 AGTTCCTGATCTCCAGCAGCCAGGACGGCCACAGTGGACCCCTGTTCTTCCAGAAAGTGAAGGTGTTCCAGGGCAAC  
 CAGGACAGCTGACCCCTGCTGCTCAACAGCTGGACCCCTGCTGACCCAGATACCTGAGGATCCACCCCAAGAGCTGGGT  
 GCACCAAGATCCCTCTGAGGATGGAGGTGCTGGGCTGTGAGGCCAGGACCTGTACTGA

Subrayado continuo = coFVIII-62-NT58 (SEQ ID NO: 9)

Subrayado discontinuo = coFVIII-62-CT (SEQ ID NO: 10)



**FIG. 1G: coFVIII-25 – SEQ ID NO: 5**

ATGCAGATTGAGCTGAGCACCTGCTTCTTCTGTGCTGCTGAGGTTCTGCTTCTTCTGCCACCAGGAGATACTACCTGGGCGC  
 CGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCAGTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGA  
 GCTTCCCCCTTCAACACCAGCGTGGTGTACAAGAAGACCTGTTCTGTGGAGTTTCACTGACCACCTGTTCAACATCGCCAAGCCC  
 AGGCCCCCTGGATGGGCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTCACTACCCTGAAGAATATGGC  
 CAGCCACCCCTGCTCCCTGCACGCCGTGGGGGTGAGCTACTGGAAGGCTCTGAGGGCGCCGAGTACGACGACCAGACCAGCC  
 AGAGGGAGAAGGAGGACCAAGGTCTTCCCTGGGGGCGAGCCACACCTACGTGTGGCAGGTCCTGAAGGAGAAGCGCCCCATG  
 GCCTCTGACCCCTGTGCTGACCTACAGCTACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATTGGGGC  
 CCTGCTGGTGTGACGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCTGCACAAGTTTATCTCTGCTGTTTCCGCGTGTTCG  
 ACGAGGGCAAGAGCTGGCACTCTGAAACCAAGAAGCCTGATGCAGGACAGGGACGCCGCTCTGCCAGGGCTGGCCCAAG  
 ATGCACACCGTCAACGGCTACCTCAACAGGAGCCTGCCGTGGCCTGATTGGCTGCCACAGGAAGAGCGTGTACTGGCATGTGAT  
 CGGCATGGGCACCACCCCTGAGGTGCACAGCATCTTCTGGAGGGCCACACCTTCTGGTCAAGAACACAGGCAGGCCAGCC  
 TGGAGATCAGCCCCATCACCTTCTGACCGCCAGACCCTGCTGATGGACCTGGGCCAGTTTCTGCTGTTTCTGCCACATCTCC  
 AGCCACCAGCAGCAGCGCATGGAGGCTACGTGAAAGTGGACAGCTGCCCTGAGGAGCCCCAGCTGAGGATGAAGAACAACGA  
 GGAGGCCGAGGACTATGATGACGACCTGACCGACAGCGAGATGGACGTGGTCAAGTTTCGACGACGACAACAGCCCCAGCTTCA  
 TCCAGATCAGGAGCGTGGCCAAGAAGCACCCCAAGACCTGGGTGCACTACATCGCTGCTGAGGAGGAGGACTGGGACTATGCC  
 CCCCTGGTGTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTGAACAATGGCCCCAGAGGATTGGCAGGAAGTACAA  
 GAAAGTCAGGTTTATGGCTTACCTGATGAAACCTTCAAGACCAGGGAGGCCATCCAGCATGAGTCTGGCATCTTGGGCCCC  
 TGCTGTACGGGAGGTGGGGGACACCCCTGCTGATCATCTTCAAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCATGGC  
 ATCACCAGCGTGGGCCCCCTGTACAGCAGGAGGCTGCCTAAGGGGGTGAAGCACCTGAAAGACTTCCCCATCTGCTGGGGG  
 GATCTTCAAGTACAAGTGGACTGTGACTGTGGAGGACGGCCCCACCAAGAGCGACCCAGGTGCCGTGACCAGATACTACAGCA  
 GCTTCGTCAACATGGAGAGGACCTGGCCTGTGGCTGATTGGCCCCCTGCTGATCTGCTACAAGGAGTCTGTGGACCCAGAGG  
 GGCACACGATCATGAGCGACAAGAGGAAGCTGATCTTCTCTGTCTTTCGACGAGAACAGGAGCTGGTACCTGACTGAAAA  
CATCCAGCGGTTCTTCCCCAACCCCGGGCGTGCAGCTGGAAGATCTGAGTTTCAGGCATCAACATCATGCACTCCATTA  
ACGGCTACGTGTTCGATTTCGTGTCAGCTGAGCGTGTGTCTGCACGAAGTGGCCCTACTGGTACATCTGTCCATTGGTGGCCAG  
ACTGACTTCTGTCCGTGTTTTTCTCCGGTACACGTTCAAGCACAAGATGGTGTACGAGGACACCTGACCTCTTCCCTTT  
TTCCGGCGAACTGTGTTTTATGAGCATGGAGAATCCCGGCTGTGGATCTTGGGCTGCCACAACAGCGACTTCCGTAACAGAG  
GAATGACTGCGCTGCTCAAGGTGTCCAGCTGCGACAAGAACACCGGAGACTATTATGAGGACTCATACGAGGACATCTCCGCC  
TACTCTCTGTCCAGAATAACGCCATTGAACCTCGGAGCTTCAGCCGAGAACCCACCCGTGCTTAAGAGACATCAACGGGAGAT  
CCTAGGACCACCTGCAGTACAGACAGGAGGAAATCGACTACGATGACACCATCTCGGTTCGAGATGAAGAAGGAGGACTTTG  
ACATCTACGACGAAGATGAAACCAGAGCCCGAGGTCGTTCCAAAAGAAAACCCGCCACTACTTTATTGCTGCTGTTCGAGCGG  
CTGTGGGACTACGGAATGTCTCTCGCCGACGTGTCCGCAACCGAGCCAGAGCGGCTCGGTGCGCGAATTCAAGAAGGT  
CGTGTTCAGGAGTTCACTGACGGGAGCTTCACTCAGCCTTTGTACCGGGGAGAACTCAATGAACATCTCGGCTCTCTCGGAC  
CTTACATCAGAGCAGAAGTGAAGATAACATCATGGTCACTTTCCGTAAACAGCCAGCCGCCCTACTCGTTCTACTCTCTCC  
CTCATTTCTTACGAAGAGGACAGCGGCGAGGCGCAGAACCGCGCAAGAACTTCGTGAAGCCCAACGAACCAAGACCTACTT  
CTGGAAGTGCAGCATCATATGGCCCGACTAAGGACGAGTTGACTGCAAGCCCTGGGCTACTTCTCCGATGTGGACTTGG  
AGAAGGACGTCCACTCCGGCTCATCGGTCCCTGCTCGTGTGCGCATACCAATACCCTGAACCCGACACGGTGCAGGTC  
ACCGTGCAGGAGTTCTGCTCTGTTCTTCACTATCTTCGACGAACTAAGTCTGCTTCTACCGAGAACATGGAGAGGAAGT  
CAGAGCCCCCTGTAACATCCAGATGGAGGACCCGACGTCAAGGAAAACCTACCGGTTCCACGCCATTAACGGATACATCATGG  
ATACGCTGCCGGTCTTGTGATGGCCAGGATCAACGGATCAGATGGTACTTATTTGTCGATGGGCAGCAACGAGAACATCCAC  
TCTATTCACTTCTCCGGTTCATGTGTTCACTGTGCGGAAGAAGGAAGATACAAGATGGCCCTGTACAACCTTTATCCCGGAGT  
GTTGCAAACTGTGGAATGTGCGCGTCGAAGGCCGGCATTTGGCGCGTGGAGTGTGATTGGAGAACATCTCCATGCGGGGA  
TGTCAACCCTGTCTCTGTTGATAGCAACAAGTGCCAGACTCCGCTTGGGATGGCGTCAGGACACATTAGGGATTTCCAGATC  
ACTGCGTCCGGCCAGTACGGCCAATGGGCCCCAAGCTGGCCCGCTGCATTACTCCGGATCCATTAAAGCCCTGGTCAACCAA  
GGAGCCATTCTCTGGATCAAGGTGGACCTTCTGGCCCCCATGATTATCCACGGAATTAAGACCCAGGGGGCCCGGAGAAAT  
TCTCTCACTGTACATCAGCCAGTTTATATCATGTACTCCCTGGACGGAAGAAGTGGCAAACTTACAGGGGGAAACAGCACC  
GGCACACTGATGTTCTTTTTCGGAATGTGGACTCTTCCGGGATTAAGCATAACATCTTCAACCTTCCGATTATCGCTCGGTA  
CATTAGACTTACCTTACCCACTACAGCATTCGCTCCACCCTGCGGATGGAAGTGTGGGCTGCGATCTGAACCTGTCGAGCA  
TGCGGTTGGGAATGGAGTCCAAAGCAATTTCCGACGCGAGATCACCGCCCTCGTCTACTTTACCAACATGTTCCGCACGTGG  
TCACCGTCCAAAGCCCGGCTGCACCTCCAGGGAAGATCCAACGCATGGCGGCCACAGGTCAACACCCCTAAGGAGTGGCTCCA  
GGTGGACTTCCAGAAAACCATGAAGGTACCCGAGTCAACCCAGGGAGTGAAGTCGCTGCTGACTTCTATGTACGTCAAGG  
AGTTCCTGATCTCCAGCAGCAGGACGGGACCACTGTTCTTCTTCCAAAATGGAAGGTCAAGGTGTTTCAGGGCAAT  
CAGGATTCACTTCCACCGGTGTAACCTTGTATCCACCCCTCTGACCCGCTACCTTCGCATCCACCACAGTCTTGGGT  
GCACCAGATCGCGCTGAGGATGGAGGTCTGGGATGCGAAGCCAGGACCTGTACTGA

Subrayado continuo = coFVIII-25-NT58 (SEQ ID NO: 11)

Subrayado discontinuo = coFVIII-25-CT (SEQ ID NO: 12)



**FIG. 1H: coFVIII-26 – SEQ ID NO: 6**

ATGCAGATTGAGCTGAGCACCTGCTTCTTCTGTGCTGCTGAGGTTCTGCTTCTCTGCCACCAGGAGATACTACCTGGGCGC  
 CGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGCGAGCTGCCAGTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGA  
 GCTTCCCCCTCAACACCAGCGTGGTGTACAAGAAGACCCTGTTCTGTGGAGTTCACTGACCACCTGTTCAACATCGCCAAGCCC  
 AGGCCCCCTGGATGGGCTGCTGGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTCACTACCCCTGAAGAACATGGC  
 CAGCCACCCCTCTCCCTGCACGCCGTGGGGGTGAGCTACTGGAAGGCCTCTGAGGGCGCCGAGTACGACGACCAGACGAGCC  
 AGAGGGAGAAGGAGGACGACAAGGTGTTCCCTGGGGGACGCCACACCTACGTGTGGCAGGTCTGAAGGAGAACGGCCCCATG  
 GCCTCTGACCCCTGTGCTGACCTACAGCTACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACCTCTGGCCTGATTGGGGC  
 CCTGCTGGTGTGCAGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCTGCACAAGTTCATCTGCTGTTTCGCCGTGTTCCG  
 ACGAGGGCAAGAGCTGGCACTCTGAAACCAAGAAGACCTGATGCAAGGACAGGGACGCCGCTCTGCCAGGGCCTGGCCCAAG  
 ATGCACACCGTCAACGGCTACGTCAACAGGAGCCTGCCTGGCCTGATTGGCTGCCACAGGAAGAGCGTGTACTGGCATGTGAT  
 CGGCATGGGCACCCCTGAGGTGCACAGCATCTTCTGGAGGGCCACACCTTCTGGTCAAGAACACAGGCAGGCCAGCC  
 TGGAGATCAGCCCCATACCTTCTGACCGCCAGACCCTGCTGATGGACCTGGGCCAGTTCCTGCTGTTCTGCCACATCTCC  
 AGCCACCAGCAGCAGCGCATGGAGGCTACGTGAAAGTGGACAGCTGCCCTGAGGAGCCCCAGCTGAGGATGAAGAACAACGA  
 GGAGGCCGAGGACTATGATGACGACCTGACCGACAGCGAGATGGACGTGGTCAAGTTCGACGACGACAACAGCCCCAGCTTCA  
 TCCAGATCAGGAGCGTGGCCAAGAAGCACCCCAAGACCTGGGTGCACTACATCGCTGCTGAGGAGGAGGACTGGGACTATGCC  
 CCCCTGGTGTGCTGGCCCTGATGACAGGAGCTACAGAGCCAGTACCTGAACATGGCCCCAGAGGATGGCAGGAAGTACAA  
 GAAAGTCAGGTTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATCCAGCATGAGTCTGGCATCCTGGGCCCCC  
 TGCTGTACGGGGAGGTGGGGGACACCCTGCTGATCATCTTCAAGAACCAGGCCAGCAGGCCCTACAACATCTACCCCCATGGC  
 ATCACCAGCTGAGGCCCTGTACAGCAGGAGGCTGCCTAAGGGGTGAAGCACCTGAAAGACTTCCCCATCTGCTGGGGG  
 GATCTTCAAGTACAAGTGGACTGTGACTGTGGAGGACGGCCCCACCAAGAGCGACCCAGGTGCCTGACCAGATACTACAGCA  
 GCTTCTGCTCAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCCCTGCTGATCTGCTACAAGGAGTCTGTGGACCAGAGG  
 GGCAACCAGATCATGAGCGACAAGAGGAACGTGATCTGTCTCTGCTTTCGACGAGAACAGGAGCTGGTACCTCACTGAAAA  
 CATCCAGAGGTTCTCCCAAACCCCGAGGAGTGAACCTGGAGGACCTGAGTTTCAGGCCCTCGAATATCATGCACTGATTA  
 ACGGTTACGTGTTTCGACTCGCTGCAGCTGAGCGTGTGCTTCCATGAAGTTCGTTACTGGTACATTCTGTCCATCGGCGCCAG  
 ACTGACTTCTGAGCGTGTCTTTTCCGGTTACACCTTTAAGCACAAGATGGTGTACGAAGATACCCCTGACCCCTGTTCCCTTT  
 CTCCGGCGAAACGGTGTTCATGTGATGGAGAACCCTGGTCTGTGGATTCTGGGATGCCACAACAGCGACTTTCGGAACCGCG  
 GAATGACTGCCCTGCTGAAGGTGTCTCATGCGACAAGAACACCGGAGACTACTACGAGGACTCCTACGAGGATATCTCAGCC  
 TACCTCTGTCCAAGACAACCGCATCGAGCCGCGCAGCTTTCAGCCAGAACCCTGCTGTGAAGAGGACCAGCGAGAAAT  
 TACCCGGACACCCCTCCAATCGGATCAGGAGGAAATCGACTACGACGACACCATCTCGGTGGAAATGAAGAAGGAAGATTTCG  
 ATATCTACGACGAGGACGAAATCAGTCCCTCGCTTCATTTCAAAAAGAAAACCTAGACACTACTTTATCGCCGCGGTGGAAAGA  
 TGTGGGACTATGGAAATGTCATCCAGCCCTCAGCTCCTTCGGAACCCGGGCCAGAGCGGATCGGTGCTCAGTTCAAGAAAGT  
 GGTGTTCCAGGAGTTACCGGACGGCAGCTTACCCAGCCGCTGTACCGGGGAGAACTGAACGAACACCTGGGCCTGCTCGGTC  
 CCTACATCCGCGCGGAAGTGGAGGATAACATCATGGTGACCTTCCGTAACCAAGCATCCAGACCTTACTCTTCTATCTCTCC  
 CTGATCTCATACGAGGAGGACAGCGCCAAGGCGCCGAGCCCCGAAGAATCTCGTCAAGCCCAACGAGACTAAGACCTACTT  
 CTGGAAGGTCCAACACCATATGGCCCCGACCAAGGATGAGTTTGACTGCAAGGCTGGGCCTACTTCTCCGACGTGGACCTTG  
 AGAAGGATGTCCATTCCGGCCTGATCGGGCCGCTGCTCGTGTGTACACCAACACCCCTGAACCCAGCGCATGGACGCCAGGTC  
 ACCGTCCAGGAGTTTCTCTGTTCTTCAACATTTTGACGAACATAAGTCTGGTACTTCAACCGAGAATATGGAGCGAACTG  
 TAGAGCGCCCTCGAATATCCAGATGGAAGATCGACTTTCAAGGAGAATATAGATTCCAGCCATCAACGGGTACATCATGG  
 ATACTCTGCCGGGGCTGGTCATGGCCAGGATCAGAGGATTCTGGTGGTACTTGCTGTCAATGGGATCGAACGAAAACATTCAC  
 TCCATTCACTTCTCCGTCACGTGTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCGCTGTACAATCTGTACCCGGGGT  
 GTTCGAACCTGTGGAGATGCTGCCGTCCAAGGCCGGCATCTGGAGAGTGGAGTGCCTGATCGGAGAGCCTCCACGCGGGGA  
 TGTCCACCTCTTCTGGTGTACTCGAATAAGTGCCAGACCCGCTGGGCATGGCCTCGGGCCACATCAGAGACTTCCAGATC  
 ACAGCAAGCGGACAATACGGCCAATGGGCGCCGAGCTGGCCGCTTGCATCTACTCCGATCGATCAACGCATGGTCCACCAA  
 GGAACCGTTCTCGTGGATTAAGGTGGACCTCTGGCCCCCTATGATTATCCACGGAATTAAGACCCAGGGCGCCAGGCAGAAGT  
 TCTCTCCCTGTACATCTCGCAATTCATCATGTACAGCTGGACGGGAGAGAGTGGCAGACTTACAGGGGAACTCCACC  
 GGCACCCCTGATGGTCTTTTTCGGCAACGTGGATTCTCCGGCATTAAGCACACATCTTCAACCCACCGATCATAGCCAGATA  
 TATTAGGCTCCACCCCACTCACTACTCAATCCGCTCAACTCTTCGGATGGAACCTCATGGGTGCGACCTGAACCTCTGCTCCA  
 TGCCGTTGGGGATGGAATCAAAGGCTATTAGCGACGCCCAGATCACCGCGAGCTCCTACTTCACTAATATGTTTCGCCACCTGG  
 AGCCCCFCCAAGGCCAGGCTGCATTTGCAGGGACGGTCAAATGCTGGCGCCGCAAGTGAACAATCCGAAGGAATGGCTTCA  
 AGTGGATTTCAAAAGACCATGAAAGTGACCGGAGTCAACACCCAGGGAGTGAAGTCCCTTCTGACCTCGATGTATGTGAAGG  
 AGTTCTGATTAGCAGCAGCCAGGACGGGCACCGTGGACCTGTCTTCTCCAAAACGGAAGGTCAAGGTGTTCCAGGGGAAC  
 CAGGACTCGTTACACCCGTGGTGAACCTCCCTGGACCCCCCTGCTGACGCGGTACTTGAGGATTATCTCTCAGTCTGGGT  
 CCATCAGATTGCATTGCGAATGGAAGTCTGGGCTGCGAGGCCAGGACCTGTACTGA

Subrayado continuo = coFVIII-26-NT58 (SEQ ID NO: 13)

Subrayado discontinuo = coFVIII-26-CT (SEQ ID NO: 14)



**FIG. 11: BDD-FVIII (no optimizado; "parental"),**  
**Secuencia de Nucleótidos (SEQ ID NO: 16)**

ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCTTTTGCAGATTCTGCTTTAGTGCCACCAGAAAGATACTACCTGGGTGCAGTG  
 GAACTGTCATGGGACTATATGCAAAGTGATCTCGGTGAGCTGCTGTGGACGCAAGATTTCTCCTAGAGTGCCAAAAATCTTTTCC  
 ATTC AACACCTCAGTCGTGTACAAAAAGACTCTGTTTGTAGAATTCACGGATCACCTTTTCAACATCGCTAAGCCAAGGCCACCT  
 GGATGGGTCTGCTAGGTCTACCATCCAGGCTGAGGTTTATGATACAGTGGTCATTACACTTAAGAACATGGCTTCCCATCCTGTC  
 AGTCTTCATGCTGTTGGTGATCTCTACTGGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAAAGGGAGAAAGAA  
 GATGATAAAGTCTTCCCTGGTGGAAAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAATGGTCCAATGGCCTCTGACCCACTGT  
 GCCTTACCTACTCATATCTTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTATGTAGAGAA  
 GGGAGTCTGGCCAAGGAAAAGACACAGACCTTGACAAATTTATACTACTTTTGTCTGATTGATGAAGGGAAAAGTTGGCACT  
 CAGAAACAAAGAACTCCTTGATGCAGGATAGGGATGCTGCATCTGCTCGGGCCTGGCCTAAAATGCACACAGTCAATGGTTATGT  
 AAACAGGTCTCTGCCAGGTCTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTCTGAAAGTG  
 CACTCAATATTCTCGAAGGTCACACATTTCTGTGAGGAACCATCGCCAGGCGTCTTGGAAATCTCGCCAATAACTTTCTTACT  
 GCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTTTGTCTATCTCTTCCCACCAACATGATGGCATGGAAGCTTATGTC  
 AAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAATAATGAAGAAGCGGAAGACTATGATGATGATCTTACTGAT  
 TCTGAAATGGATGTGGTCAGGTTTGTATGATGACAACTCTCCTTCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCCTAAAAC  
 TGGGTACATTACATTGCTGCTGAAGAGGAGGACTGGGACTATGCTCCCTAGTCTCGCCCCGATGACAGAAGTTATAAAAGTC  
 AATATTTGAACAATGGCCCTCAGCGGATTGGTAGGAAGTACAAAAAGTCCGATTATGGCATACACAGATGAAACCTTTAAGAC  
 TCGTGAAGCTATTACAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAAGTTGGAGACACACTGTTGATTATATTAAAG  
 ATCAAGCAAGCAGACCATATAACATCTACCCTCACGGAATCACTGATGTCCTGCTTTGTATTCAAGGAGATTACAAAAAGGTGTA  
 AAACATTTGAAGGATTTTCAATTCTGCCAGGAGAAATATTCAAATATAAATGGACAGTGACTGTAGAAGATGGGCCAACTAAAT  
 CAGATCCTCGGTGCTGACCCGCTATTACTAGTTCGTTAATATGGAGAGAGATCTAGCTTCAGGACTCATTGGCCCTCTCTCA  
 TCTGCTACAAAGAATCTGTAGATCAAAGAGGAAACAGATAATGTCAGACAAGAGGAATGTCATCCTGTTTCTGTATTTGATGA  
 GAACCGAAGCTGGTACCTCACAGAGAATATACAACGCTTCTCCCAATCCAGCTGGAGTGCAGCTTGAGGATCCAGAGTTCCAA  
 GCCTCCAACATCATGCACAGCATCAATGGCTATGTTTTGATAGTTTGAGTTGTCAGTTTGTGATGAGGTGGCATACTGGTA  
 CATTCTAAGCATTGGAGCACAGACTGACTTCTTCTGTCTTCTCTGGATATACCTTCAAACACAAAAATGGTCTATGAAGACAC  
 ACTCACCTATTCCCATCTCAGGAGAACTGTCTTCATGTCGATGGAAAACCCAGGTCTATGGATTCTGGGGTGCCACAACCTCAG  
 ACTTTCCGAACAGAGGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAGAACTGGTGATTATTACGAGGACAGTTATGA  
 AGATATTTCACTACTTGTGAGTAAAAACAATGCCATTGAACCAAGAAGCTTCTCTCAAAACCCAGCTTGAACGCCATC  
 AACGGGAAATAACTCGTACTACTCTTCACTCAGTCAAGAGGAAATTGACTATGATGATACCATATCAGTTGAAATGAAGAAGGA  
 AGATTTTGACATTATGATGAGGATGAAATCAGAGCCCCGAGCTTCAAAGAAAAACGACACTATTTTATTGCTGCAGTGG  
 AGAGGCTCTGGGATTATGGGATGAGTAGCTCCCATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGA  
 AAGTTGTTTTCCAGGAATTTACTGATGGCTCCTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGGGG  
 CCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACCTTCAGAAATCAGGCCTCTCGTCCCTATTCTTCTATTCTAGCCTT  
 ATTTCTTATGAGGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAATTTGTCAGCCTAATGAAACCAAACTTACTTTTGA  
 AAGTGCAACATCATATGGCACCCTAAAGATGAGTTTGACTGCAAGGCTGGGCTTATTTCTGATGTTGACCTGGAAAAAGAT  
 GTGCACTCAGGCCTGATTGGACCCCTTCTGCTGTCACACTAACACACTGAACCTGCTCATGGGAGACAAGTGACAGTACAGG  
 AATTTGCTCTGTTTTTACCATCTTTGATGAGACCAAAAGCTGGTACTTCACTGAAAATATGGAAAGAACTGCAGGGCTCCCTGC  
 AATATCCAGATGGAAGATCCACTTTTAAAGAGAATTATCGCTCCATGCAATCAATGGCTACATAATGGATACACTACCTGGCTT  
 AGTAATGGCTCAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAACATCCATTCTATTCTTTAGTGGA  
 CATGTGTTCACTGTACGAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCTATCCAGGTGTTTTGAGACAGTGGAAATGT  
 TACCATCAAAGCTGGAATTTGGCGGGTGGAAATGCCTTATTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTTCTGGGTGA  
 CAGCAATAAGTGTGAGACTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAATTACAGCTTCAGGACAAATATGGACAG  
 TGGGCCCCAAAGCTGGCCAGACTTCATTATCCGGATCAATCAATGCCTGGAGCACAAGGAGCCCTTTTCTGGATCAAGGTGG  
 ATCTGTTGGCACCATGATTATTCACGGCATCAAGACCCAGGGTGCCGTCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCA  
 TCATGTATAGTCTTATGGGAAGAGTGGCAGACTTATCGAGGAAATCCACTGGAACCTTAAATGGTCTTCTTTGGCAATGTGGAT  
 TCATCTGGGATAAAACACAATATTTTAAACCTCCAATTATTGCTCGATACATCCGTTTGACCCAACTCATTATAGCATTTCGAGC  
 ACTCTTCGATGGAGTTGATGGGCTGTGATTTAAATAGTTGAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCAC  
 AGATTACTGCTTCATCTACTTTACCAATATGTTTGCCACCTGGTCTCCTTCAAAGCTCGACTTCACCTCCAAGGGAGGAGTAATG  
 CCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATGAAAGTCACAGGAGTAATACTC  
 AGGGAGTAAATCTCTGCTTACCAGCATGTATGTGAAGGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTT  
 TTTCAGAATGGCAAAGTAAAGGTTTTTCAGGGAAATCAAGACTCCTTCACACCTGTGGTGAAGTCTCTAGACCCACCGTTACTGAC  
 TCGTACCTTCGAATTCACCCCAAGAGTTGGGTGCACAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAGGCACAGGACCTC  
 TAC

**\*BDD-FVIII = FVIII Suprimido Dominio B**

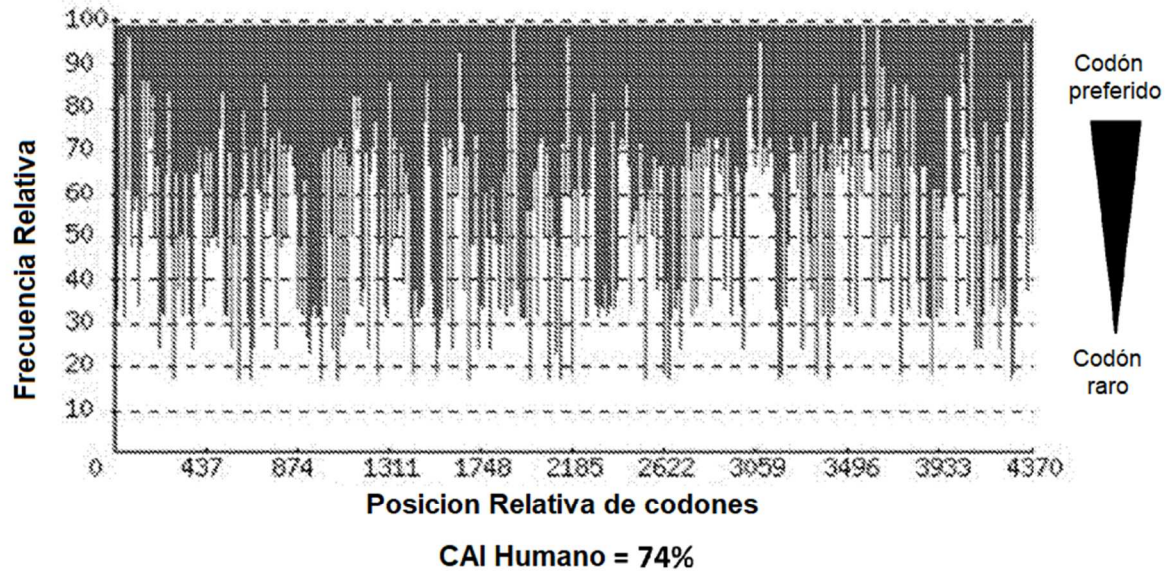
**FIG. 1J: BDD-FVIII (no optimizado; "parental"),**  
**Secuencia de Aminoácidos (SEQ ID NO: 17)**

ATRRYYLGAVELSWDYMQSDLGELPVDARFPFPRVPKSFPEFNTSVVYKTLFVEFTDHLFNIAPRPPWMGLLGPTIQAEVYDVTVITLK  
 NMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGA  
 LLVCREGLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRSVYWHVIGMGT  
 TPEVHSIFLEGHTFLVRNHRQASLEISPIFLTAQTLLMDLGQFLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLT  
 DSEMDVVRFDDDNPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLPDDRYSYKSYLNNGPQRIGRKYKKVRFMAYTDETFKTR  
 EAIQHESGILGPLLYGEVGDTHLIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFILPGEIFKYKWTVTVEDGPTKSDPRCLTRY  
 SSFVNMERDLASGLIGPLLYCYKESVDQGRNQIMSDKRNVLFSVFDENRSWYLTENIQRFNPAGVQLEDPEFQASNIMHSINGYVF  
 DSLQLSVCLHEVAYWYLSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMENPGLWILGCHNSDFRNRGMTALLKVSSC  
 DKNTGDYEDSYEDISAYLLSKNNAIEPRSFQNPVLRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTR  
 HYFIAAVERLWDYGMSSSPHVLNRNRAQSGSVPOFKKVVFEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSF  
 YSSLISYEEDQRQGAEPKRFVKNETKTYFWKVQHMAPTKDEFDCKAWAYFSDVDLEKDVHSGILGPLLVCHTNTLNPAHGRQVT  
 VQEFALFFTIFDETKSWYFTENMERNCRAPCNQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSG  
 HVFTVRKKEEYKMALYNLYPGVFETVEMLPKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYQWA  
 PKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGKIQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGLMVFFGNVDSSGIKH  
 NIFNPPIIARYIRLHPHYSIRSLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNP  
 KEWLQVDFQKTMKVTVGTQGVKSLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW  
 VHQIALRMEVLGCEAQDLY

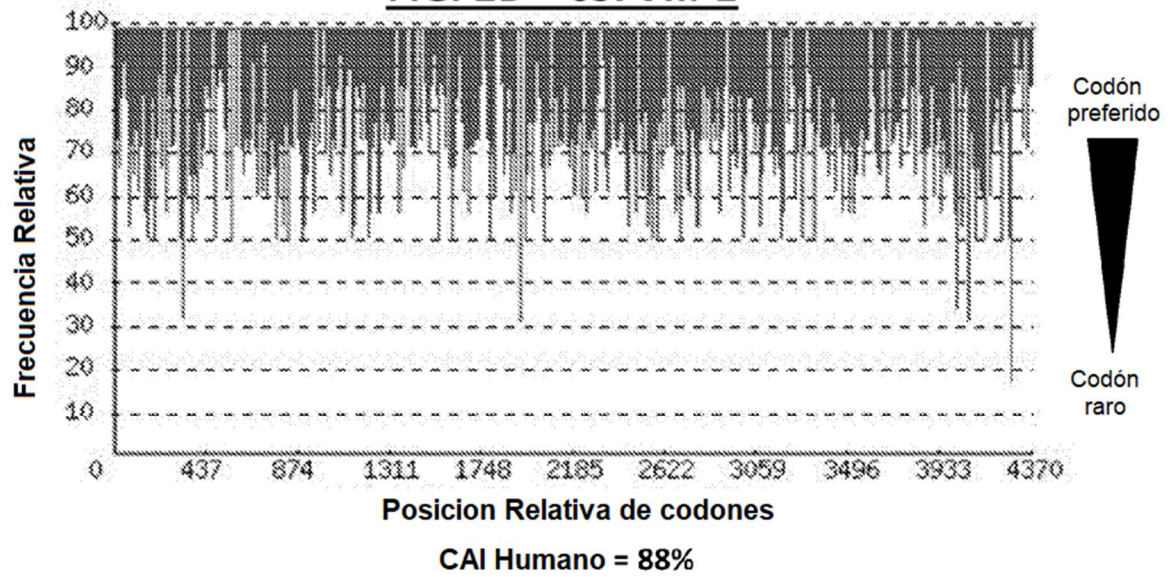
**\*BDD-FVIII = FVIII Suprimido Dominio B**



**FIG. 2A – BDD FVIII No Optimizado**

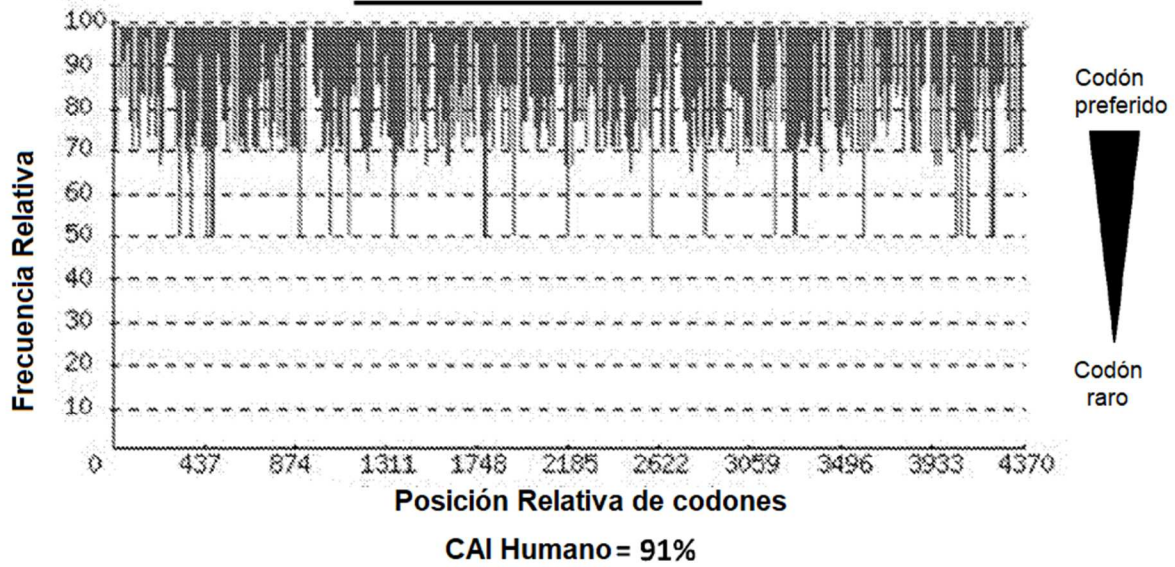


**FIG. 2B – coFVIII-1**

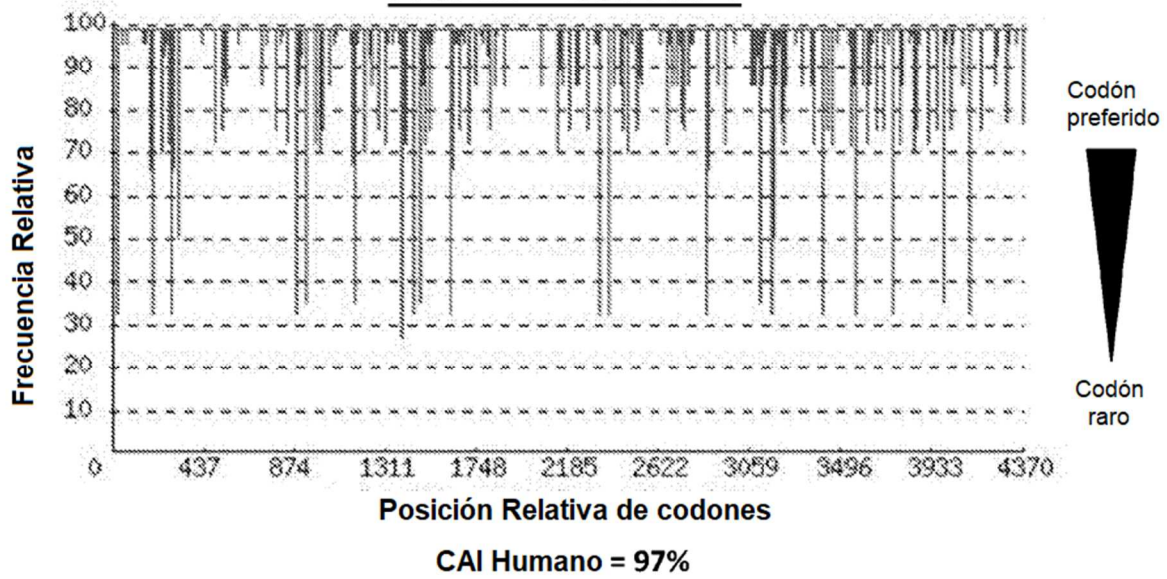


**BDD FVIII = FVIII Suprimido Dominio B**

**FIG. 2C – coFVIII-3**

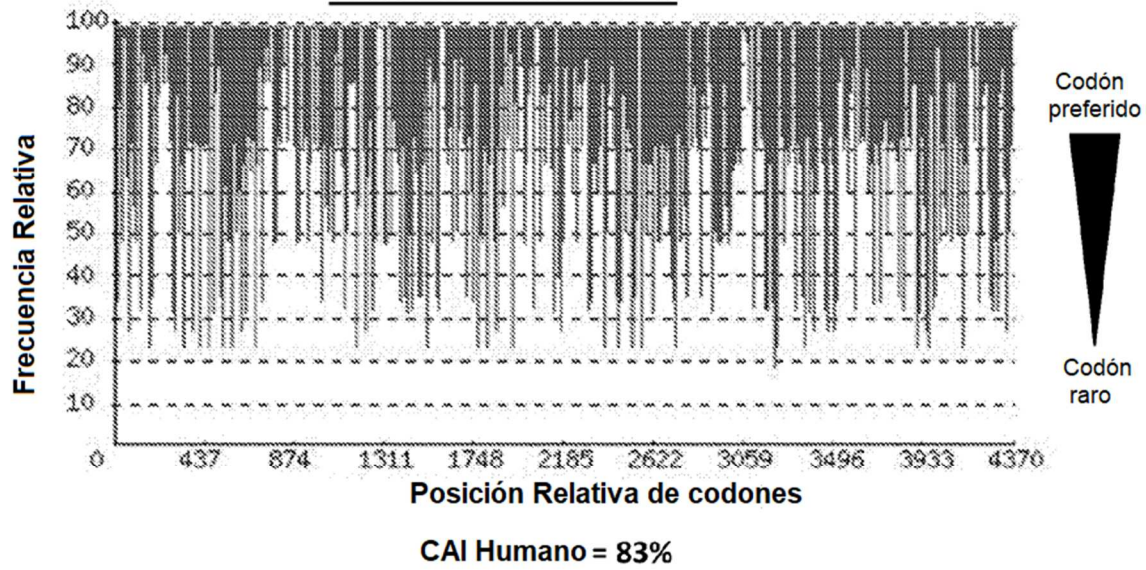


**FIG. 2D – coFVIII-4**

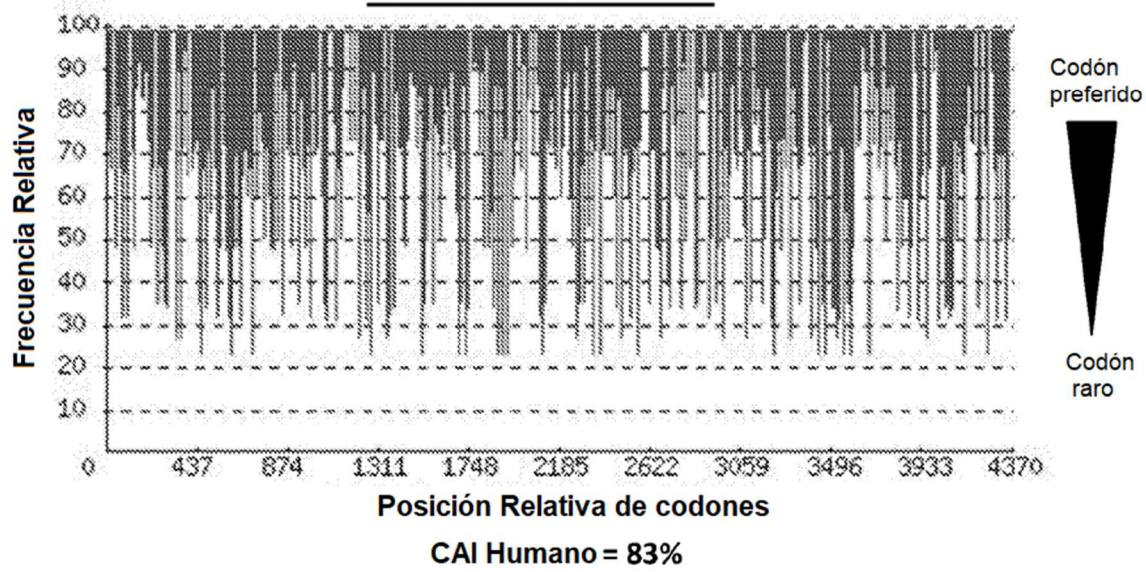




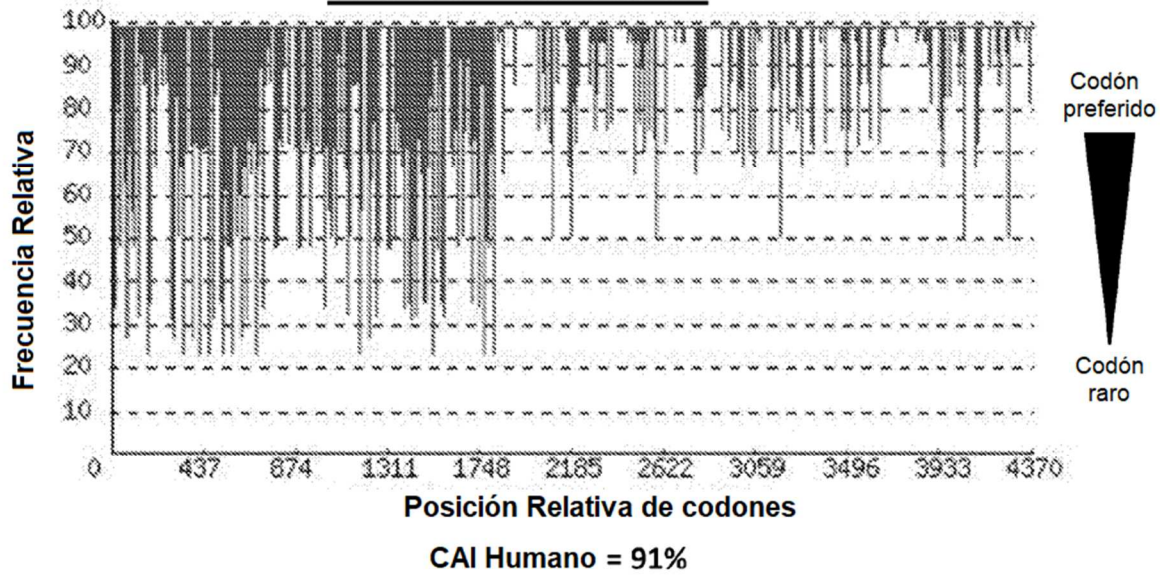
**FIG. 2E – coFVIII-5**



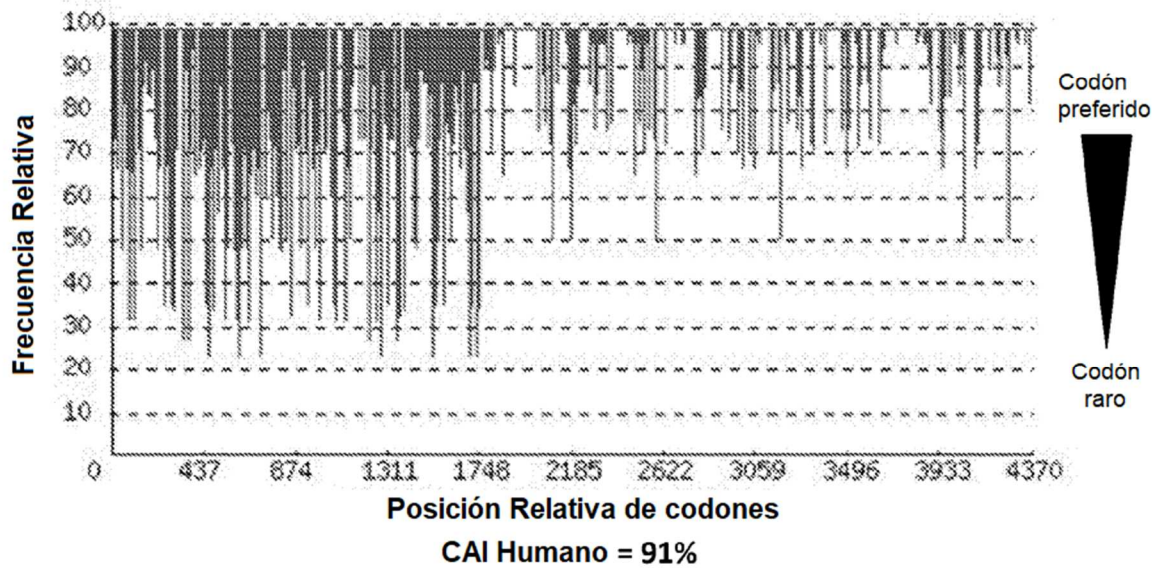
**FIG. 2F – coFVIII-6**



**FIG. 2G – coFVIII-52**

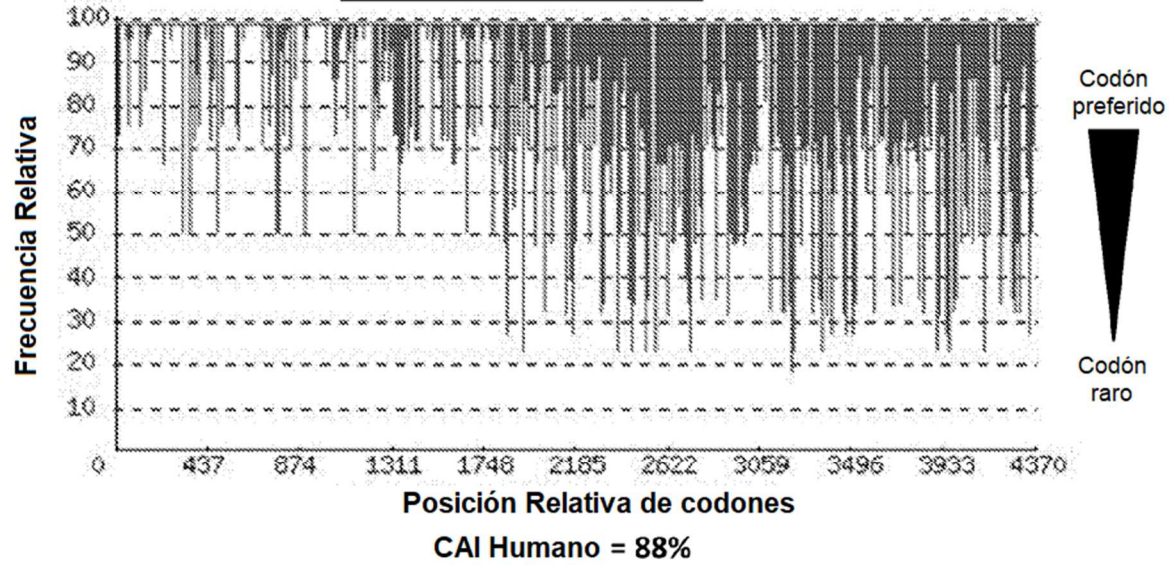


**FIG. 2H – coFVIII-62**

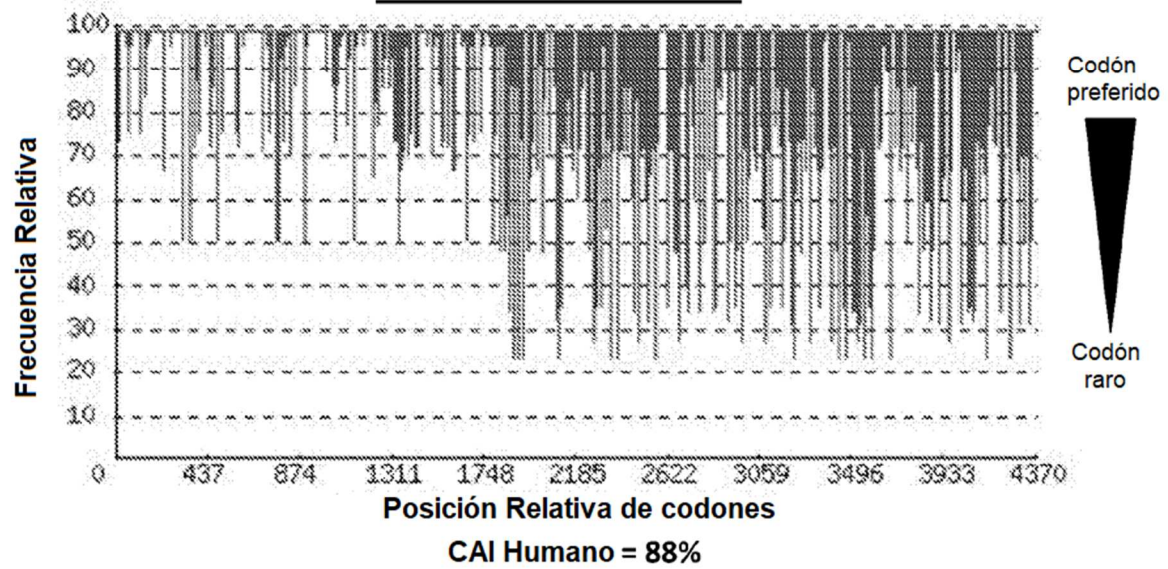




**FIG. 2I – coFVIII-25**



**FIG. 2J – coFVIII-26**



**FIG. 3: Mapas de Plásmidos de FVIII Optimizado por Codones  
Plásmidos de Expresión**

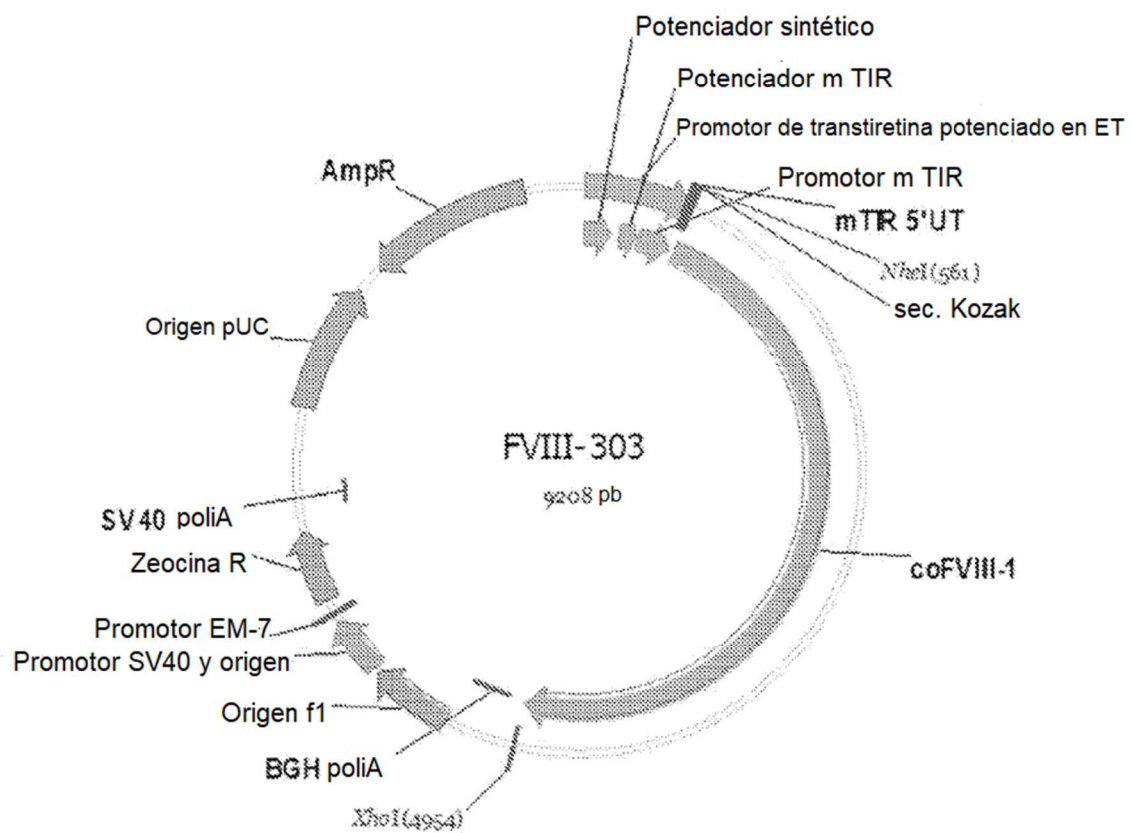
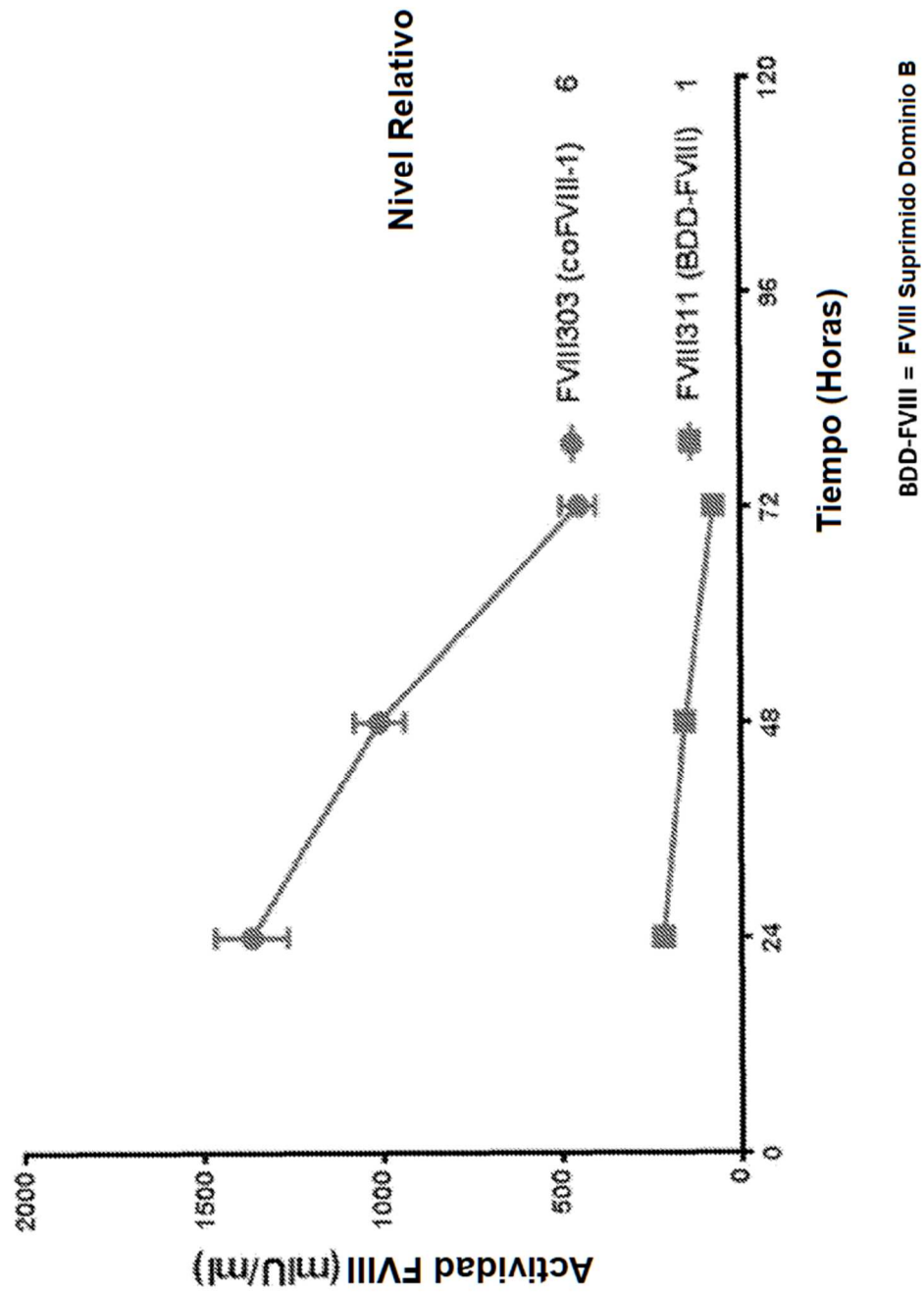
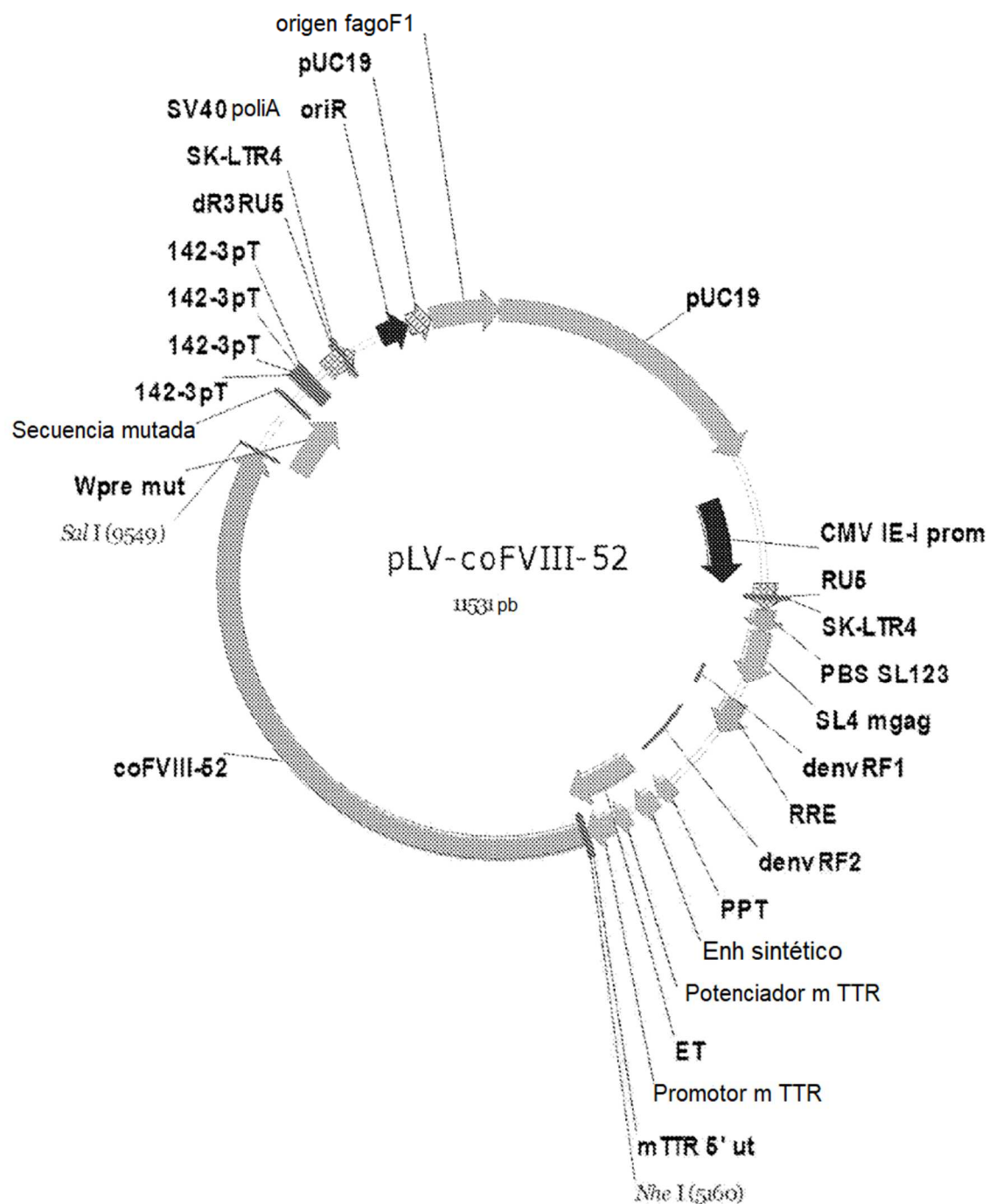


FIG. 4: Actividad de FVIII en Ratones HemA



**FIG. 5: Mapas de Plásmidos de FVIII Optimizado por Codones  
Plásmidos de Expresión**





**FIG. 6A: coFVIII-3, coFVIII-4, coFVIII-5 y coFVIII-6  
con respecto a coFVIII-1**

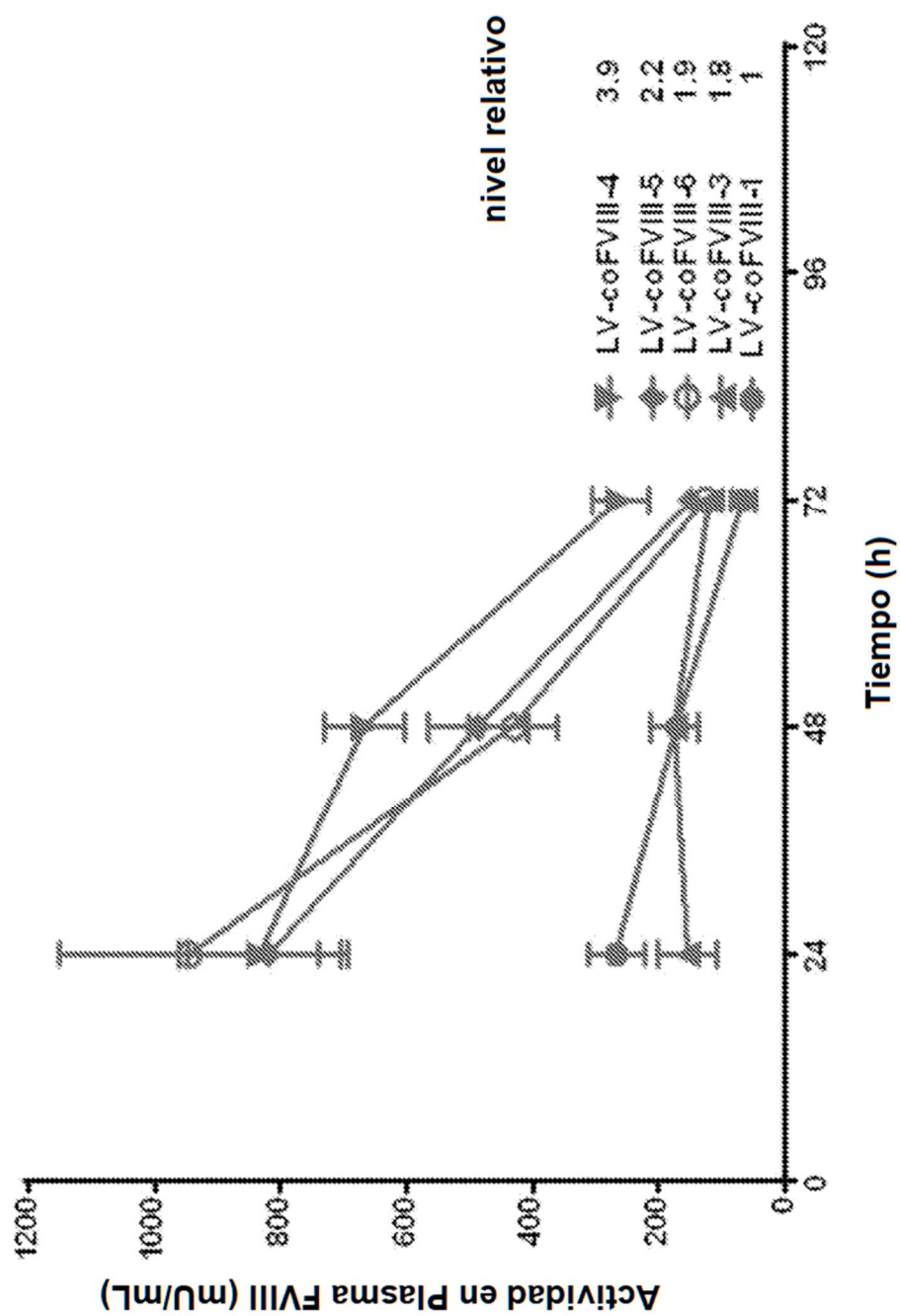


FIG. 6B: coFVIII-25 y coFVIII-26 con respecto a coFVIII-1

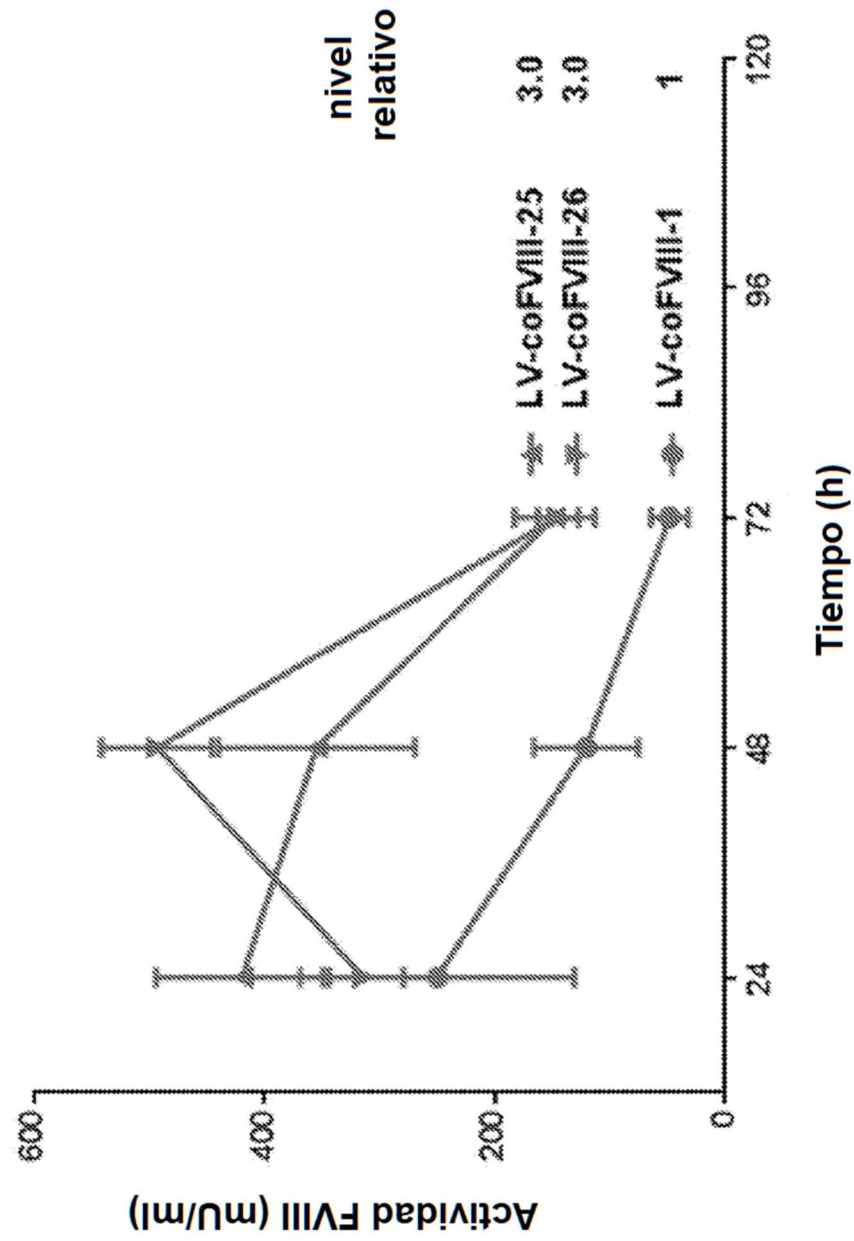
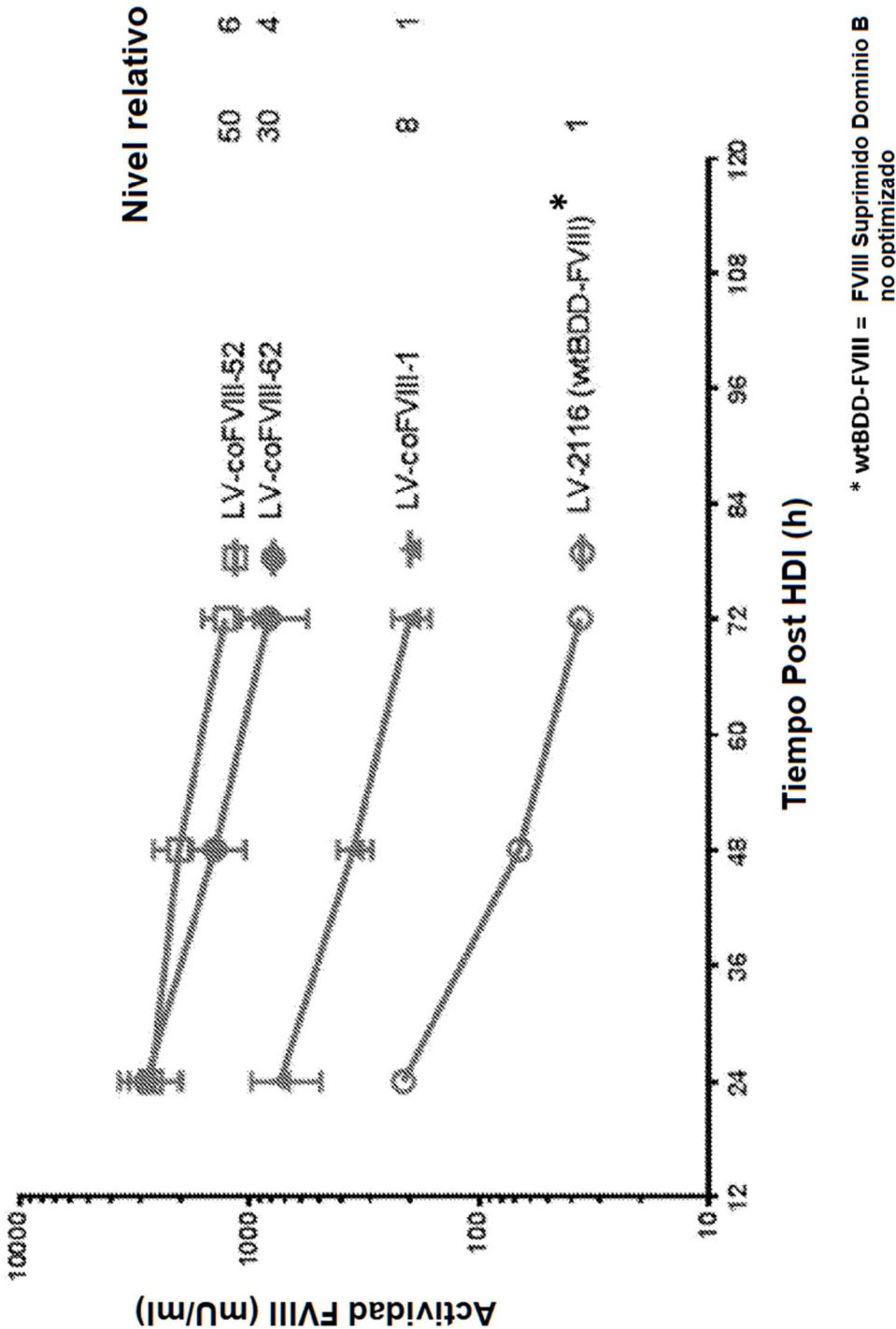
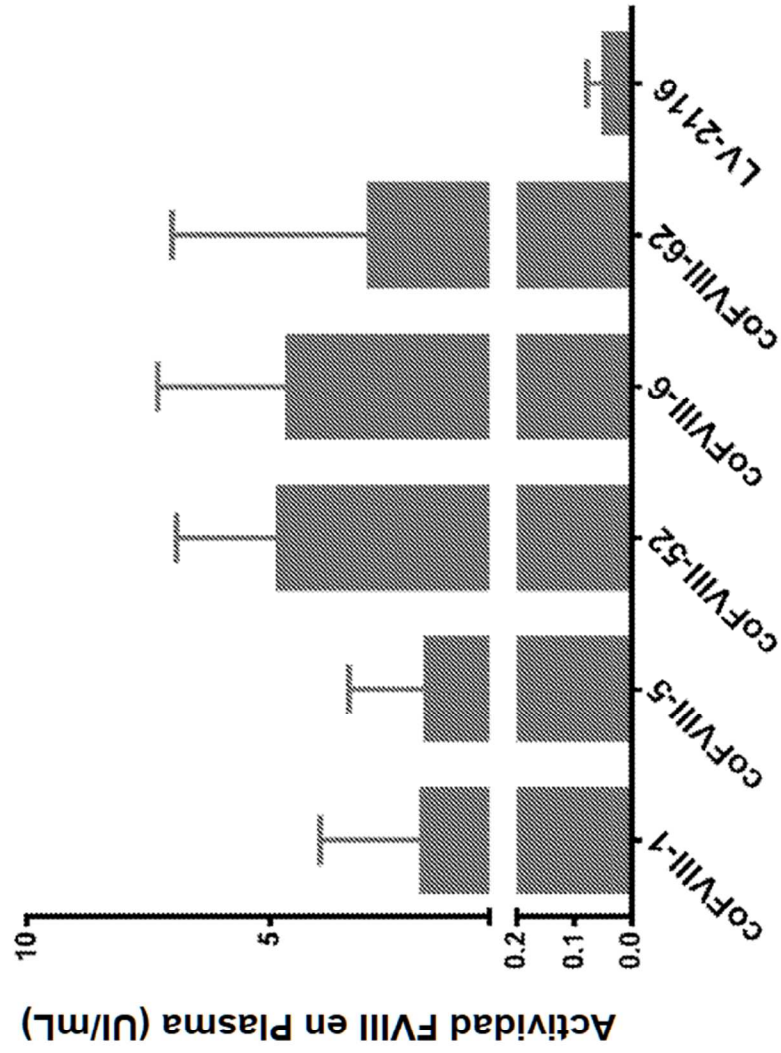


FIG. 6C: coFVIII-52 y coFVIII-62 con respecto a BDD-FVIII y coFVIII-1



**FIG. 7: Actividad de FVIII Mediada por Vector Lentiviral en Ratones HemA**

\*LV-2116 = FVIII Suprimido Dominio B  
bajo promotor ET en plásmido lentiviral



**FIG. 8A: coFVIII-52-XTEN – SEQ ID NO: 19**

ATGCAAAATCGAACTGAGCACCTGTTTCTTCTCTGCCTGCTGAGATTCTGTTTCTCCGCGACCCGCCGATACCTGGGAGCAGTGG  
 AGCTCTCCTGGGATTACATGCAGAGCGACCTTGGGGAGCTGCCCCTGGATGCCAGGTTCCCTCCCGGGTGCCAAAGTCGTTTCCGTT  
 CAACACCTCCGTTGGTGTACAAGAAAACCTGTTCTGTTGGAGTTACCCGACCACTGTTCAATATCGCCAAGCCGAGACCTCCCTGGATG  
 GGGCTGTTGGGACCTACCATCCAAGCGGAGGTGTACGACACTGTGGTTCATCACTCTGAAGAACATGGCCTCGCATCCCGTGTCCCTGC  
 ACGCCGTGGGAGTGTCTTACTGGAAGCGTCCGAGGGGGCCGAATACGACGACAGACCTCGCAGAGAGAAAGGAAGATGACAAGGT  
 GTTCCAGGAGGATCGCACACCTACGTGTGGCAAGTGTGAAGGAGAACGGCCCAATGGCCTCCGACCCGCTGTGCCTGACCTACTCG  
 TACCTGTCCCAGTGGACCTCGTGAAGGACCTCAACTCGGACTGATTGGAGCCCTGCTGGTCTGCAGGGAAGGCTCACTGGCGAAAG  
 AAAAGACTCAGACCTTGCACAAGTTCATTCTGCTGTGCTGCTGTTTCGACGAGGGGAAGTCTGGGCACAGCGAGACTAAGAAGTCCCT  
 GATGCAAGATAGAGATGCCGCTCCGCCGGGCTGGCCTAAGATGCACACCGTGAACGGTTACGTGAACCGCTCCCTCCCTGGCCTG  
 ATTGGATGCCACCGGAAGTCCGTGTACTGGCAGTGTATCGGGATGGGACCACCCCGAGGTGCACAGCATCTTCTGGAAGGTGACA  
 CATTCTCGTGCACAACACCGGCAGGCTCCCTGGAATACGCCCCATTACCTTCCTCACTGCCAGACTCTGCTGATGGACCTGGG  
 ACAGTTCCTGCTGTTCTGCCATATCTCTCCACCAACATGACGGAATGGAGGCATACGTGAAGGTGCGATTCTGCCCTGAGGAACCC  
 CAGCTCCGCATGAAGAACAATGAGGAAGCCGAGGACTACGACGACGACCTGACGGATAGCGAGATGGATGTGGTCCGGTTCGATGACG  
 ATAACAGCCCTTCTTCAATCCAAATTCGCTCGGTGGCAAAGAGACCCCAAGACCTGGGTGCATTACATTGCGGCGGAAGAAGAGGA  
 CTGGGATTATGCCCCGCTTGTCTCGCTCTGACGACCGGAGCTACAAGAGCCAGTACCTGAACAACGGTCCACAGAGGATCGGTAGA  
 AAGTACAAGAAGTCCGCTTCATGGCCTATACCGACGAAACCTTCAAACTAGAGAGGCCATCCAACACGAATCCGGCATCTTGGGCC  
 CGCTCTGTACGAGAAGTCGGCGACACCTTCTCATTATCTTCAAGAACAGGCTTCCCGGCCGTACAACATCTATCCGCATGGGAT  
 CACTGACGTGCGCCACTGTACTCGCGCGCCTGCCCAAGGGTGTCAAAACCTGAAGGATTTTCCGATCTTCCGGGAGAAATCTTC  
 AAGTACAAGTGGACCGTGACCGTGAAGATGGCCCACTAAGTCTGACCTAGATGCCTCAGCCGCTACTACTCATCTTCTGTCACA  
 TGGAGCGCAGCTGGCCAGCGGACTGATCGGCCCGCTGCTGATTGCTTACAGGAATCAGTGGACCAACGGGGAACAGATCATGTC  
 GGATAAGAGGAACGTATCTCTTCTCCGTGTTTACGAAAACCGGTGCTGGTACCTGACCGAGAACATCCAGAGGTTCTGCCCAAC  
 CCTGCTGGGTGACGTGGAGGACCCGAGTTCAGGCCAGCAACATCATGCACAGCATCAATGGCTACGTGTTGACAGCCTGCAGC  
 TGAGCGTGTGCCTGCACGAGGTGGCTACTGGTACATCTGAGCATCGCGCGCCAGACCGACTTCTGAGCGTGTCTTCTCTGGCTA  
 CACCTTCAAGCACAAGATGGTGTATGAGGACACCTGACCTGTTCCCTTTCAGCGGGGAGACTGTCTTATGAGCATGGAGAACCCT  
 GGCCTGTGGATCTTGGGTGCCACAACAGCGACTTCAGGAACAGGGGCATGACTGCCCTGCTGAAAGTCTCCAGCTGTGACAAGAACA  
 CCGGGGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTTACCTGCTGAGCAAGAACATGCCATCGAGCCAGGAGCTTCTCTCA  
 GAACGGCGCGCAACATCAGAGAGCGCCACCCCTGAAAGTGGTCCCGGGAGCGAGCCAGCCACATCTGGGTGCGAAACGCCAGGCACA  
 AGTGAGTCTGCAACTCCCGAGTCCGGACCTGGCTCCGAGCCTGCCACTAGCGGGTCCGAGACTCCGGGAACCTCCGAGAGCGTACAC  
 CAGAAAGCGGACCCGGAACAGTACCGAACCTAGCGAGGGCTCTGCTCCGGGCGCCAGCCGGCTCTCTTACATCCACGGAGGAGGG  
 CACTTCCGAATCCGCCACCCCGAGTCAAGGCCAGGATCTGAACCCGCTACCTCAGGCAGTGAGACGCCAGGAACGAGCGAGTCCGCT  
 ACACCGGAGAGTGGGCCAGGGAGCCTGCTGGATCTCTACGTCCACTGAGGAAGGGTCAACAGCGGGCTCGCCCAACGAGCATGAAG  
 AAGGTGCTCGAGCCCCCAGTGTGAAGAGGCACAGAGGGAGATCACCAGGACCACTTGCAGTCTGACAGGAGGAGATCGACTA  
 TGATGACACCATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAGGACGAGAACCAGAGCCCCAGGAGCTTCCAGAAG  
 AAGACAGGCACTACTTCTTCTGCTGCTGTGGAGAGGCTGTGGGACTATGGCATGTCCAGCAGCCCCATGTGCTGAGGAACAGGGCCC  
 AGTCTGGCAGCGTGCCCCAGTTCAAGAAAGTCTGTGTTCCAGGAGTTCACCGACGGCAGCTTCAACAGCCCCGTGTACAGAGGGGAGCT  
 GAACGAGCACCTGGGCCCTGCTGGGCCCTTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGGAACAGGCCAGCAGG  
 CCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACAGGAGGAGGCTGAGCCAGGAAGAAGCTTGTGAAGCCCAATG  
 AAACCAAGACCTACTTCTGGAAGGTGCAGCACCATGCCCCCACCAGGACGAGTTCGACTGCAAGGCTGGGCCACTTCTCTGA  
 CGTGGACCTGGAGAAGGACGTGCACTCTGGCCTGATTGGCCCCCTGCTGGTGTGCCACCAACACCCTGAACCCCTGCCCATGGCAGG  
 CAGGTGACTGTGAGGAGTTCGCCCTGTTCTTCAACATCTTCGATGAAACCAAGAGCTGGTACTTCACTGAGAACATGGAGAGGAACT  
 GCAGGGCCCCCTGCAACATCCAGATGGAGGACCCACCTTCAAGGAGAACTACAGGTTCATGCCATCAATGGCTACATCATGGACAC  
 CCTGCTGGCCTGGTTCATGGCCAGGACAGAGGATCAGGTGGTATCTGCTGAGCATGGCAGCAACGAGAACATCCACAGCATCCAC  
 TTCTCTGGCCACGTGTTTCACTGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAACCTGTACCCTGGGGTGTTCGAAAACCGTGG  
 AGATGCTGCCAGCAAGGCCGCGCATCTGGAGGGTGGAGTGCCTGATTGGGGAGCACCTGCACGCCGGCATGAGCACCTGTTCTGCT  
 GTACAGCAACAAGTGCCAGACCCCTGGGCATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGCCAGTACGGCCAG  
 TGGGCCCCCAAGCTGGCCAGGCTGCACTACTCCGGAAGCATCAATGCCGTGGAGCACCAAGGAGCCCTTCACTGAGATCAAAGTGGACC  
 TGCTGGCCCCCATGATCATCCACGGCATCAAGACCAGGGGGCCAGGCGAGAAGTCTCCAGCCTGTACATCAGCCAGTTTCATCATCAT  
 GTACAGCCTGGACGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACCGGCACCCCTGATGGTGTCTTTCGGCAACGTGGACAGCAGC  
 GGCATCAAGCACAACATCTTCAACCCCCCATCATCGCCAGATACATCAGGCTGCACCCACCCACTACAGCATCAGGAGCACCCCTGA  
 GGATGGAGCTGATGGGCTGTGACCTGAACAGCTGCAGCATGCCCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCAGATCACTGC  
 CTCCAGTACTTACCAACATGTTTGGCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGCAATGCTGGAGGCC  
 CAGGTCAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAGACCATGAAGGTGACTGGGGTGACCAACAGGGGGTGAAGAGCC  
 TGCTGACCAGCATGTACGTGAAGGAGTTCCTGATCTCCAGCAGCCAGGACGGCCACAGTGGACCCCTGTTCTTCCAGAATGGCAAGGT  
 GAAGGTGTTCCAGGGCAACAGGACAGCTTCAACCCCTGTGGTCAACAGCCTGGACCCCCCTGCTGACAGATACCTGAGGATCCAC  
 CCCCAGAGCTGGGTGCACAGATCGCCCTGAGGATGGAGGTGCTGGGCTGTGAGGCCAGGACCTGTACTGA

**Subrayado = secuencia de nucleótidos de XTEN (SEQ ID NO: 18)**

**FIG. 8B: coFVIII-1-XTEN – SEQ ID NO: 20**

ATGCAGATTGAGCTGTCTACTTGCTTTTCTGTCGCTGCTGAGGTTTGTCTTTCCGCTACACGAAGGTATTATCTGGGGGCTGTGG  
 AACTGTCTTGGGATTACATGCAGAGTGACCTGGGAGAGCTGCCAGTGGAACGCAAGGTTTCCCCCTAGAGTCCCTAAGTCATTCCTT  
 CAACACTAGCGTGGTCTACAAGAAAACACTGTTCTGTTGAGTTTACTGATCACCTGTTCAACATCGCAAAGCCTAGGCCACCTGGATG  
 GGACTGCTGGGGCCAACAATCCAGGCCGAGGTGTACGACACCGTGGTCTATTACACTTAAGAATCGGCTCACACCCCGTGGCTGC  
 ATGCTGTGGGCGTCAGCTACTGGAAGGCTTCCGAAGGAGCAGAGTATGACGATCAGACTTCCAGAGAGAAAAAGGAGACGATAAGGT  
 GTTTCTGGCGGATCTCATACCTACGTGTGGCAGGTCTGAAAGAGAATGGCCCTATGGCCTCCGACCCCTGTGTGCTGACCTACTCT  
 TATCTGAGTACGTGGACCTGGTCAAGGATCTGAACAGCGGCTGATCGGAGCCCTGCTGGTGTGAGGGAAGGAGCCTGGCTAAGG  
 AGAAAACCCAGACACTGCATAAGTTCTTCTGCTGTTGCGCGTGTGACGAAGGGAATCATGGCACAGCGAGACAAAGAATAGTCT  
 GATGCAGGACAGGGATGCCGCTTCAGCCAGAGCTTGGCCCAAATGCACACTGTGAACGGCTACGTCAATCGCTCACTGCTGGGCTG  
 ATCGCTGCCACCGAAAGAGCGTGTATTGGCATGTCATCGGGATGGGCACACACCTGAAAGTCACTCCATTTTCTGGAGGGACATA  
 CTTTCTGTGTCGCAACCCAGCAGGCTTCCCTGGAGATCTCTCAATTACCTTCTGACAGCACAGACTCTGCTGATGGACCTGGG  
 GCAGTCTCTGCTGTTTGGCCATCAGTCCCACAGCATGATGGCATGGAGGCTTACGTGAAAGTGGACTCTTGTCCCGAGGAACCT  
 CAGCTGCGGATGAAGAACAATGAGGAAGCAGAAGACTATGACGATGACCTGACCGACTCCGAGATGGATGTGGTCCGATTCTGATGACG  
 ATAACAGCCCTCTTTATCCAGATTAGATCTGTGGCCAAGAAACACCTTAAGACATGGGTCCATTACATCGCAGCCGAGGAAGAGGA  
 CTGGGATTATGCACCACTGGTGTGGCACCAGACGATCGCTCTCAAAATCTCAGTATCTGAACAATGGGCCACAGAGGATTGGCAGA  
 AAGTACAAGAAAGTGGGTTTCATGGCATATACCGATGAGACCTTCAAGACTCGCGAAGCCATCCAGCACGAGAGCGGCATCTGGGAC  
 CACTGCTGTACGGAGAAGTGGGAGACACCTGCTGATCATTTTCAAGAACCAGGCCAGCCGGCTTACAATATCTATCCACATGGGAT  
 TACAGATGTGGCCCTCTGTACAGCAGGAGACTGCCAAGGGCGTCAACACCTGAAGGACTTCCCAATCTGCCCGGAGAAATCTTC  
 AAGTACAAGTGGACTGTCAACCGTCGAGGATGGCCCACTAAGAGCGACCTCGGTGCTGACCCGCTACTATTCTAGTTTCTGTGAATA  
 TGGAAAGAGATCTGGCAAGCGGACTGATCGGACCACTGCTGATTGTTTACAAAGAGAGCGTGGATCAGAGAGGCAACAGATCATGTC  
 CGACAAGCGGAATGTGATTCTGTTGAGTGTCTTACGAAAACAGGTGATGTTACCTGACCGAGAACATCCAGAGATTCTGCCTAAT  
 CCAGCTGGGCTGACGTGGAAGATCTGAGTTTTCAGGCATCTAACATCATGCATAGTATTAATGGCTACGTGTTGACAGTTTGCAGC  
 TGAGCGTGTGCTGACGAGGTCGCTTACTGGTATATCTGAGCATTGGGGCACAGACAGATTCTGAGCGTGTCTTTTCCGGCTA  
 CACTTTTAAAGCATAAAATGGTCTATGAGGACACACTGACTCTGTTCCCTTCAGCGGCGAAACCGTGTATTAGCATGGAGAATCCC  
 GGACTGTGGATTCTGGGGTGCCACAACAGCGATTTCAGAAATCGCGGAATGACTGCCCTGCTGAAAGTGTCAAGCTGTGACAAGAA  
 CCGGGGACTACTATGAAGATTACACGAGGACATCAGCGCATATCTGCTGTCCAAAACATGCCATTGAACCCCGGTCTTTAGTCA  
 GAATGGCGCGCAACATCAGAGAGCGCCACCCCTGAAAGTGGTCCCGGGAGCGAGCCAGCCACATCTGGGTGCGAAACGCCAGGCACA  
 AGTGAAGTCTGCAACTCCCGAGTCCGGACCTGGCTCCGAGCCTGCCACTAGCGGCTCCGAGACTCCGGGAACCTCCGAGAGCGCTACAC  
 CAGAAAGCGGACCCGGAACAGTACCGAACCTAGCGAGGCTCTGCTCCGGGCGAGCCAGCCGGCTCTCTACATCCAGGAGGAGGG  
 CACTTCCGAATCCGCCACCCCGAGTCAAGGCGCAGGATCTGAACCCGCTACCTCAGGCAGTGAGACGCCAGGAACGAGCGAGTCCGCT  
 ACACCGGAGAGTGGGCGAGGAGCCCTGCTGGATCTCTACGTCCACTGAGGAAGGGTCACCGCGGGCTCGCCACAGCACTGAAG  
 AAGTGCCTCGAGCCCTCCAGTGTGAAGCGGCACCGCGGAGATCACCCGCACTACCTGCAGAGTGATCAGGAAGAGATCGACTA  
 CGACGATACATTTCTGTGGAATGAAGAAAGAGGACTTCGATATCTATGACGAAGATGAGAACCAGAGTCCCTCGATCATTTCCAGAAG  
 AAAACAGGCATTACTTTATTGCCGAGTGGAGCGGCTGTGGGATTATGGCATGTCTCTAGTCCCTACGCTGCTGCGAAATAGGGCCC  
 AGTCAGGAAGCGTCCACAGTTCAAGAAAGTGGTCTTCCAGGAGTTTACAGACGGGCTCTTACTCAGCCACTGTACAGGGGCGAAT  
 GAACGAGCCTGGGCTGCTGGGCGCTATATCAGAGCAGAAGTGGAGGATAACATATGGTCACTTCAAGAAATCAGGCTCTCGG  
 CCTTACAGTTTTATTCAAGCCTGATCTCTTACGAAGAGGACAGCGACAGGAGCTGAACACGAAAAAACTTCGTGAAGCCTAATG  
 AGACCAAAACATACTTTTGAAGGTGACGACCATATGGCCCCAACAAAAGACGAGTTCGATTGCAAGGCATGGGCTATTTTCTGA  
 CGTGGATCTGGAGAAGGACGTGCACAGTGGCCTGATTGGCCCACTGCTGGTGTGCCATACTAACACCCCTGAATCCAGCCCGGCGG  
 CAGGTCAGTGTCCAGGAGTTCGCTCTGTTCTTTACCATCTTTGATGAGACAAAGAGCTGGTACTTACCAGAAAACATGGAGCGAAAT  
 GCAGGGCTCCATGTAACATTAGATGGAAGACCCACATCAAGGAGAACTACCGCTTTCATGCTATCAATGGATACATCATGGATAC  
 TCTGCCCGGGCTGGTATGGCAGAGGACAGAGAATCCGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAATATCCACTCAATTCAT  
 TTCAGCGGGCAGCTGTTTACTGTGAGGAAGAAAGAGTACAAGATGGCCCTGTACAACCTGTATCCCGGCGTGTTCGAAACCGTGC  
 AGATGCTGCCTAGCAAGGCCGAATCTGGAGAGTGAATGCCTGATTGGAGAGCACCTGCATGCTGGGATGTCTACCTGTTTCTGGT  
 GTACAGTAATAAGTGTGACACCCCTGGGAATGGCATCCGGGCATATCAGGGATTTCCAGATTACCGCATCTGGACAGTACGGACAG  
 TGGGCACCTAAGCTGGCTAGACTGCACTATTCCGGATCTATCAACGCTTGGTCCACAAAAGAGCCTTTCTCTTGGATTAAAGTGGACC  
 TGCTGGCCCCAATGATCATTCATGGCATCAAACTCAGGAGCTCGGCAGAGTTCTCTCTGTACATCTCACAGTTTATCATCAT  
 GTACAGCCTGGATGGGAAGAAATGGGAGACATACCGCGCAATAGCACAGGAACCTGATGGTGTCTTTGGCAACGTGGACAGCAGC  
 GGAATCAAGCACAACATTTTCAATCCCCCTATCATTGCTAGATACATCCGGCTGCACCAACCCATTATTCTATTGCAAGTACACTGA  
 GGATGGAAGTATGGGATGCGATCTGAACAGTTGTTCAATGCCCCGGGGATGGAGTCAAGGCAATCTCTGACGCCAGATTACCGC  
 CAGCTCCTACTTCAATATATGTTGTACCTGGAGCCCTTCCAAAGCAAGACTGCACCTGCAAGGCCGAGCAACGCATGGCGACCA  
 CAGGTGAACAAATCCCAAGGAGTGGTGTGAGGTCGATTTCAGAAAACATGAAGGTGACCGGGGTCAAACTCAGGGCGTGAAAGTC  
 TGCTGACCTCAATGACGTCAAGGAGTCTGATCTCTAGTTCAAGACGGACATCAGTGGACACTGTTCTTTCAAGACGGGAGGT  
 GAAAGTCTTCCAGGCAATCAGGATTCTTTACACCTGTGTTACAGCAGTCTAGACCTCCACTGCTGACCCAGTACCTGAGAATCCAC  
 CCTCAGTCTGGGTGCACAGATTGCCCTGAGAATGGAAGTGTGGGATGCGAGGCCAGGATCTGTACTGA

**Subrayado = secuencia de nucleótidos de XTEN (SEQ ID NO: 18)**

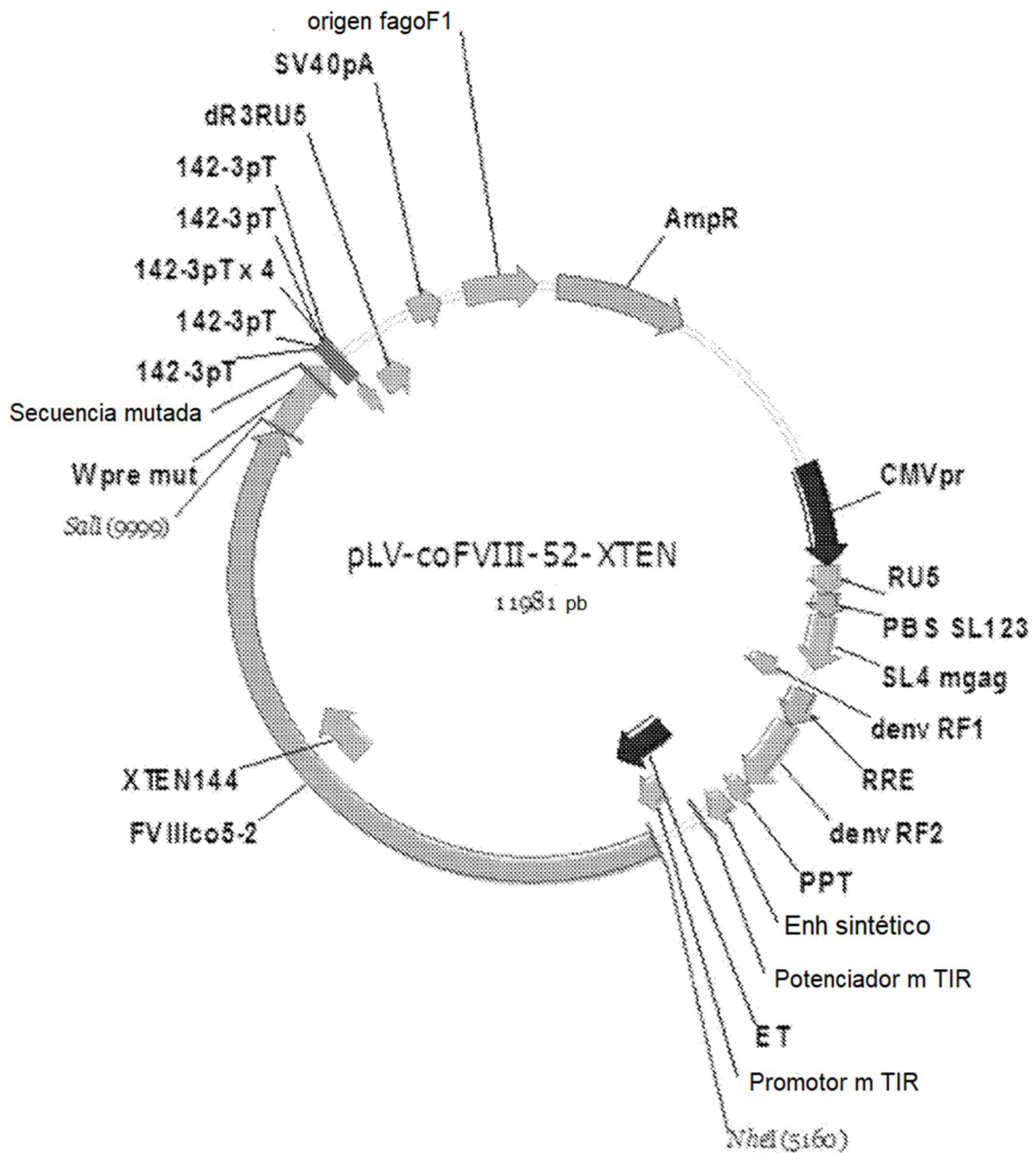


**FIG. 8C: coFVIII-6-XTEN – SEQ ID NO: 72**

ATGCAGATTGAGCTGTCCACTTGTCTTCTCTGTGCCTCCTGCGCTTCTGTTTCTCCGCCACTCGCCGGTACTACCTTGGAGCCGTGG  
 AGCTTTTCATGGGACTACATGCAGAGCGACCTGGGCGAACTCCCGTGGATGCCAGATTCCCCCCCCCGGTGCCAAAGTCTTCCCTT  
 TAACACCTCCGTGGTGTACAAGAAAACCTCTTTGTGAGTTTCACTGACCACCTGTTCAACATCGCCAAGCCGCGCCACCTTGGATG  
 GCCTCCTGGGACCGACCTTCAAGCTGAAGTGTACGACACCGTGGTATCACCTGAAGAACATGGCGTCCACCCCGTGTCCCTGC  
 ATGCGGTGCGAGTGTCTACTGGAAGGCTCCGAAGGAGTGTAGTACGACGACGAGTACGACGCGGAAAGGAGGACGATAAAGT  
 GTTCCCGGGCGGCTCGCATACTTACGTGTGGCAAGTCTGAAGGAAAACGACCTATGGCATCCGATCCTCTGTGCTGACTTACTCC  
 TACCTTTCCCATGTGGACCTCGTGAAGGACCTGAACAGCGGGCTGATTGGTGCACCTTCTCGTGTGCCGCGAAGGTTGCTCGCTAAGG  
 AAAAGACCCAGACCTCCATAAGTTTCATCTTTTGTTCGTGTGTTTCGATGAAGGAAAGTATGGCATTCCGAAACTAAGAACTCGCT  
 GATGCAGGACCGGGATGCCGCTCAGCCCGCGCTGGCTTAAATGCATACAGTCAACGGATACGTGAATCGGTCACTGCCCGGGCTC  
 ATCGGTTGTACAGAAAGTCCGTGTACTGGCAGTGTATCGGCATGGGCACTACGCTGAAGTGCATCCATCTTCTGGAAGGACACA  
 CCTTCTCGTGTGCGCAACCCAGGCGCTCTCTGGAATCTCCCGATTACCTTCTGACCGCCAGACTCTGCTCATGGACCTGGG  
 GCAGTTCTTCTCTTCTGCCACATCTCCAGCCATCAGCACGACGGAATGGAGGCTACGTGAAGGTGGACTCATGCCCGGAAGAACCT  
 CAGTTGCGGATGAAGAACACGAGGAGGCGGAGGACTATGACGACGATTGACTGACTCCGAGATGGACGTCGTGCGGTTCTGATGACG  
 ACAACAGCCCCAGCTTTCATCCAGATTGCGACGCTGGCAGAAAGCACCACCAACCTGGGTGCATACATCGCGCGGAGGAAGAAGA  
 TTGGGACTACGCCCCGTTGGTGTGCTGGCAGCGATGACCGGCTCTCAAGTCCGAGTATCTGAACAATGGTCCGCGAGGATTGGCAGA  
 AAGTACAAGAAAGTGGGTTTCTGCGGTACTGACGAAACGTTTAAAGACCGGGAGGCGCATCAACATGAGAGCGGCATTCTGGGAC  
 CACTGCTGTACGAGAGGTCGCGGATACCTGCTCATCATCTTCAAAACAGGCTCCCGGCTTACAACATCTACCTCACGGAAT  
 CACCGAGTGCAGGCACTTACTCGCGGCGCTGCGGAGGGGCTCAAGCACCTGAAAGACTTCCCTATCTGCGCGGCGAAATCTTC  
 AAGTATAAGTGGACCGTCAACGTTGGAGGACGGGCGCCCAAGAGCGATCTAGGTGTCTGACTCGGTACTACTCCAGCTTCTGTGAACA  
 TGGAACGGGACCTGGCATCGGACTCATTGGACCGTGTCTGATCTGCTACAAAGAGTCCGTGGATCAACGCGGCAACGAGATCATGTC  
 CGAGAGCTACGCCCCGTTGGTGTGCTGGCAGGATATCTAGCCTTACCTCTGTCCAGAACACGCGATCGAGCCGCGCAGCTTCAGCCA  
 GAACGCGCGCCAAACATCAGAGAGCGCCACCCCTGAAAGTGGTCCGCGGAGCGAGCCAGCCACATCTGGGTGCGAAACGCCAGGCACA  
 AGTGAAGTGTCAACTCCCGAGTCCGACCTGGCTCCGAGCCTGCCACTAGCGGCTCCGAGACTCCGGGAACTCCGAGAGCGCTACAC  
 CAGAAAGCGGACCCGGAACAGTACCGAACCTAGCGAGGGCTGCTCCGCGGAGCCAGCCGCGCTCTCTACATCCAGGAGGAGGG  
 CACTTCCGAATCCGCCACCCCGGAGTCAAGGCGAGGATCTGAACCCGCTACCTCAGGCAGTGAGACGCCAGGAACGAGCGAGTCCGCT  
 ACACCGGAGAGTGGGCGAGGAGCCCTGCTGGATCTCTACGTCCACTGAGGAAGGGTCAACAGCGGGCTCGCCACCAGCACTGAAG  
 AAGGTGCCCTCGAGCCCCGCTGTGCTGAAGAGGCAACGCGAGAAATACCCGAGACCCCTCCAATCGGATCAGGAGGAAATCGACTA  
 CGACGACACCATCTCGGTGGAAATGAAGAAGGAAGATTTGATATCTACGACGAGGACGAAATCAGTCCCTCGCTCATTCCAAAG  
 AAACTAGACACTACTTTATCGCCGCGGTGAAAGACTGTGGGACTATGGAATGTATCCAGCCCTCACGTCTTCCGAACCGGGCCC  
 AGAGCGGATCGGTGCTCAGTTCAAGAAAGTGGTGTCCAGGAGTTCACCGACGCGAGCTTCAACCGAGCGCTGTACCGGGGAGAACT  
 GAACGAACACCTGGGCTGCTCGGTCCCTACATCCGCGCGGAAGTGGAGGATAACATCATGGTGACCTTCCGTAACCAAGCATCCAGA  
 CCTTACTCCTTCTATTCTCCCTGATCTCATACGAGGAGGACGCGCAAGGCGCGAGCCCGCAAGAACTTCGTCAAGCCCAACG  
 AGACTAAGACCTACTTCTGGAAGGTCCAACACCATATAGGCCCCGACCAAGGATGAGTTTGAAGGCTGAGGCGCTTCTCTCCGA  
 CGTGGACCTTGAGAAGGATGTCCATTCCGGCTGATCGGGCGCTGCTCGTGTGTACACCAACACCCCTGAACCCAGCGCATGGACGC  
 CAGGTACCGTCCAGGAGTTGCTCTGTTCTTCAACATTTTTGACGAACTAAGTCTGTTACTTACCGAGAATATGGAGCGAACT  
 GTAGAGCGCCCTGCAATATCCAGATGGAAGATCCGACTTTCAAGGAGAACTATAGATTCCACGCCATCAACGGGTACATCATGGATAC  
 TCTGCCGGGGCTGGTTCATGGCCAGGATCAGAGGATTCGGTGGTACTTGTGTCAATGGGATCGAACGAAACATTCACTCCATTAC  
 TTCTCCGCTCACGTGTTTACTGTGCGCAAGAAGGAGGAGTACAAGATGGCGCTGTACAATCTGTACCCCGGGGTGTTGCAAACTGTGG  
 AGATGCTGCCGTCCAAGGCGGCATCTGGAGAGTGGAGTGCCTGATCGGAGAGCCTCCACGCGGGGATGTCCACCTCTTCTGTT  
 GTACTCGAATAAGTGCCAGACCCGCTGGGATGGCTCGGGCCACATCAGAGACTTCCAGATCAGCAAGCGGACAATACGGCCAA  
 TGGGCGCGCAAGCTGGCCCGCTTGCACTACTCCGATCGATCAACGATGGTCCACCAAGGAACCGTTCTCGTGATTAAAGTGGACC  
 TCCTGGCCCTATGATTATCCACGGAATTAAGACCCAGGCGCCAGGCGAAGTTCTCTCCCTGTACATCTCGCAATTATCATCAT  
 GTACAGCTGGACGGGAAGTGGCAGACTTACAGGGGAACTCCACCGGCACCTGATGGTCTTTTTCGGCAACGTGGATTCTCC  
 GGCATTAAAGCACACATCTTCAACCCACCGATCATAGCCAGATATATAGGCTCCACCCACTCACTACTCAATCCGCTCAACTCTTC  
 GGATGGAATCATGGGTTGCGACCTGAACCTCTGCTCCATGCGGTGGGGATGGAATCAAAGGCTATTAGCGACGCCAGATCACC  
 GAGCTCTTACTTCACTAACATGTTCCGCCACCTGGAGCCCTCCAAGGCCAGGCTGCACTTGCAGGGACGGTCAAATGCCTGGCGGCC  
 CAAGTGAAACATCCGAAGGAATGGCTTCAAGTGGATTTCCAAAGACCATGAAAGTGACCGGAGTCAACCCAGGGAGTGAAGTCCC  
 TTCTGACCTCGATGTATGTGAAGGAGTTCTGATTAGCAGACGCCAGGACGGGACAGTGGACCTGTTCTTCCAAACGGAAAGGT  
 CAAGGTGTTCCAGGGGAACAGGACTCGTTACACCCGTGGTGAACCTCCCTGGACCCCCACTGCTGACGCGGTACTTGAGGATTTCAT  
 CCTCAGTCTGGTCCATCAGATTGCATGGAATGGAAGTCTGGGCTGCGAGGCCAGGACCTGTACTGA

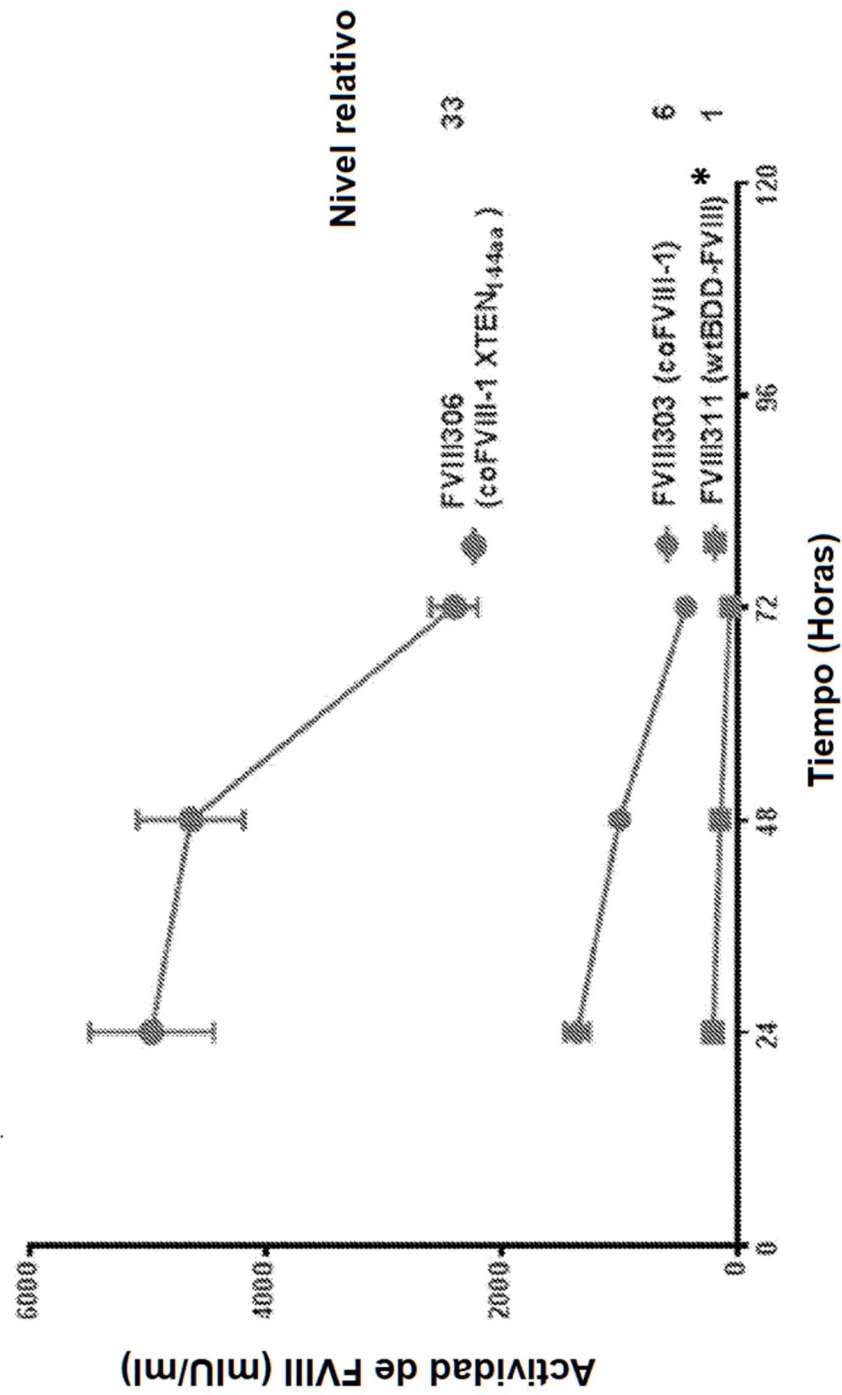
**Subrayado = secuencia de nucleótidos de XTEN (SEQ ID NO: 18)**

**FIG.9: Mapa de Plásmidos de pLV-coFVIII-52-XTEN**

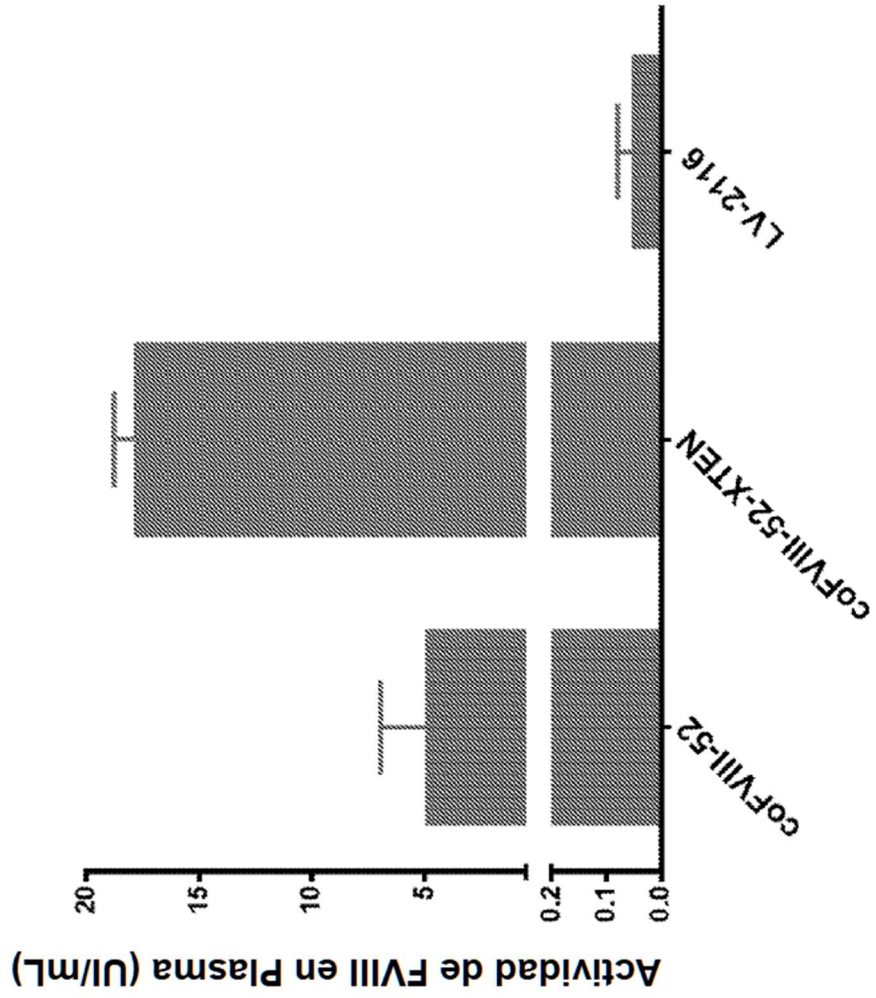




**FIG. 10A: Actividad de FVIII en Ratones Hema Post Inyección  
Hidrodinámica de ADN de Plásmido**



**FIG. 10B: Actividad de FVIII en Ratones Hema 21 días Post Inyección Lentiviral**



\*LV-2116 = FVIII Suprimido Dominio B  
bajo promotor ET en plásmido lentiviral

**FIG. 11A: Factor FVIII Humano Maduro de Longitud Completa, Secuencia de Aminoácidos (SEQ ID NO: 15)**

ATRRYYLGAVELSWDYMQSDLGELPVDARFPFPRVPKSPFNTSVVYKTLFVEFTDHLFNIAPRPPWMGLLGPTIQAEVYDTVITLK  
 NMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGA  
 LLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSYYWHVIGMG  
 TPEVHSIFLEGHTFLVRNHRQASLEISPTFLTAQTLMDLGQFLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDLT  
 DSEMDVVRFDDDNPSFIQIRSAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTR  
 EAIQHESGILGPLYGEVGDTLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFILPGEIFKYKWTVTVEDGPTKSDPRCLTRY  
 SSFVNMERDLASGLIGLLICYKESVDQGRNQIMSDKRNILFVSVFENRSWYLTENIQRFLPNPAGVQLDEPEFQASNIMHSINGYVF  
 DSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTFPFSGETVFMSEMPGLWILGCHNSDFRNRGMTALLKVSSC  
 DKNTGDYEDSYEDISAYLLSKNNAIEPRFSQNSRHPSTRQKQFNATTIPENDIEKTPWFARHTPMPKIQNVSDDLMLLRQSPTPH  
 GLSLDLQEAKEYETSDDPSPGAIDSNNSLSEMTFRPQLHHSQDMVFTPESEGLQLRLNEKLGTTAATELKKLDFKVSSTSNLISTIPSD  
 NLAAGTDNTSSLGPPSMPVHYDSQLDITLFGKKSSPLTESGGPLSLSEENNDKLLSEGLMNSQESSWGKNVSTESGRLFKGKRAHGP  
 ALLTKDNALFKVSISSLKTNKTSNNSATNRKTHIDGSPSLIENSPPSVWQNILESDETEFKKVTPLIHDRMLMDKNATALRLNHSNKTSS  
 KNMEMVQKKKEGPIPPDAQNPDMSSFFKMLFLPESARWIQRTHGKNSLNSGQGPSKQLVSLGPEKSEVGQNFLEKKNVVGKGEF  
 TKDVGLKEMVFPSSRNFLTNLDNLHENNTHNQEKKEIEKKEKTLIQENVVLPQIHTVTGTKNFMKNLFLSTRQNVESYDGAYAP  
 VLQDFRSLNDSTNRKHTAHFSKKGEEENLEGLNQTKQIVEKYACTTRISPNTSQQNFVTQSKRALKQFRLPLEETELEKRIIVDDTS  
 TQWSKNMKHLTPSTLTQIDYNEKEGKAITQSPLSDCLTRSHSIPQANRSLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRKDSDGVQ  
 ESSHFLQGAKKNNLSLAILTLEMTGDQREVGLGTSATNSVTYKKVENTVLPKPDLPKTSKGVELLPKVHIYQKDLFPTETSNNGSPGHLDL  
 VEGSLLQGTEGAIKWNEANRPGKVPFLRVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWSQEKSPKTAFAKKKDTILSNACESNHAI  
 AAINEGQNKPEIEVTWAKQGRTERLCSQNPPVLKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEENQSPRSFQKKTRHYFI  
 AAVERLWDYGMSSSPHVLNRNAQSGSVPPQFKVVFQETDGSFTQPLYRGELNEHLGLLPYIRAEVEDNIMVTFNRNQASRPYSFYSSL  
 ISYEEDQRQGAEPKRNFKPNETKYFWKVQHMAPTKDEFCKAWAYFSDVDLEKDVHSGLIGPLLCHTNTLNPAHGRQVTVQE  
 FALFFTFIDETKSWYFTENMERNCRAPCNQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQIRIRWYLLSMGNSNENIHSIHSGHVF  
 TVRKKEEYKMALYNLYPGVFETVEMLPKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPGLMASGHIRDFQITASGQYQWAPKLA  
 RLHYSGSINAWSTKEPFSWIKVDLLAPMIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHINFN  
 PPIIARYIRLHPTHSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEW  
 LQVDFQKTMKVTGVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQ  
 IALRMEVLGCEAQDLY

**FIG. 11B: Factor de von Willebrand de Longitud Completa,  
Secuencia de Aminoácidos (SEQ ID NO: 44)**

ATRRYYLGAVELSWDYMQSDLGELPVDARFPVRPKSFPFNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDVTVITLK  
 NMASHPVSLHAVGVSYWKASEGAIEDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGA  
 LLVCREGLAKEKTQTLHKFILLFAVDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSYYVHVIGMGT  
 TPEVHSIFLEGHTFLVRNHRQASLEISPTFLTAQTLLMDLGQFLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDLT  
 DSEMDVVRFDNNSPSFIQIRSVAKKHPKTWVHYIAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKVRFMAYTDEFKTR  
 EAIQHESGILGPLLYGEVGDITLLIFKNQASRPYNIYPHGIDTVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRY  
 SSFVNMRDLASGLIGPLLYCYKESVDQGRNQIMSDKRNVLFSVFDENRSWYLTEINIQRFPLNPAGVQLEDPEFQASNIMHSINGYVF  
 DSLQLSVCLHEVAYWYLSIGAQTDFLSVFFSGYTFKHKMYVEDTLTLFPFSGETVFMSEMPGLWILGCHNSDFRNRGMTALLKVSSC  
 DKNTGDYEDSYEDISAYLLSKNNAIEPRSFQNSRHPSTRQKQFNATTIPENDIEKTDPWFAHRTMPMKIQNVSSDMLLRQSPTPH  
 GLSLDLQEAKEYTFSDPSPGAIDSNNSLSEMTFRPQLHHSQDMVFTPEGLQLRLNEKLGTTAATELKKLDFKVSSTSNNLISTPSD  
 NLAAGTDNTSSLGPPSPMPVHYDSQLDITLFGKSSPLTESGGPLSLSEENNDKLLSGLMNSQESSWGKNVSTESGRLFKGKRAHGP  
 ALLTKDNALFKVISLTKNTKSNNSATNRKTHIDGPSLLIENSPPSVWQNILESDETFKKVTPLIHDMRLMDKNATALRLNHSNKTSS  
 KNMEMVQQKKEGPIPPDAQNPDMSFFKMLFLPESARWIQRTHGKNSLNSGQGPSKQLVSLGPEKSVEGQNFLEKNKVVVGKGEF  
 TKDVGLEKEMVFPSSRNFLTNLDNLHENNTNHNQEKIIEKKETLIQENVVLPQIHTVTGKTNFMKNLFLSTRQNVESYDGAYAP  
 VLQDFRSLNDSTNRKHTAHFSKKGEEENLEGLNQTKQIVEKYACTTRISPNTSQQNFVTQRSKRALKQFRPLEETELEKRIIVDDTS  
 TQWSKNMKHLTPSTLTQIDYNEKEGAITQSPLSDCLTRSHSIPQANRSPLPIAKVSSFPSIRPIYLTRVLFDQDNSSHLPAASYRKKDSGVQ  
 ESSHFLQGAKKNNLSAILTLEMTGDQREVGLSATSNTSVYKKVENTVLPKPDLPKTSKGKVELLPKVHIYQKDLFPTETSNGSPGHLDL  
 VEGSLLQGTEGAIKWNEANRPGKVPFLRVATESAKTPSKLLDPLAWDNHYGTQIPKEEWSQEKSPKTAFAKKKDTILSNACSNHAI  
 AAINEGQNKPEIEVTWAKQGRTERLCSQNPVLRHQRERITRTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKTRHYFI  
 AAVERLWDYGMSSSPHVLNRNAQSGSVPPQFKVVFQEFDTGDSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSL  
 ISYEEDQRQGAEPKRFVKNETKTYFWKVQHMAPTKDEFDCAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAGHRQVTVQE  
 FALFFTIFDETKSWYFTENMERNCRAPCNQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSENIHSHFSGHVF  
 TVRKKEEYKMALYNLYPGVFETVEMLPKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMA SGHIRDFQITASGQYQWAPKLA  
 RLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNI  
 PPIIARYIRLHPHTYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEW  
 LQVDFQKTMKVTGVTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQ  
 IALRMEVLGCEAQDLY

(X es cualquier aminoácido natural)



**FIG. 11C: XTEN AE42-4, secuencia proteínas (SEQ ID NO: 46)**

GAPGSPAGSPTSTFEFGTSEFATPESGPGSEFATSGSETPASS

**FIG. 11D: XTEN AE42-4, secuencia ADN (SEQ ID NO: 47)**

GCGCGCCAGTTCTCTCTGCTGCTCCCCACCTCAACAGAAGAGGGGACAAGCGAAAGCGCTACGCCTGAGAGTGGCCCTGGCT  
CTGAGCCAGCCACCTCCGGCTCTGAAACCCCTGCTCTGAGC

**FIG. 11E: XTEN AE144-2A, secuencia proteínas (SEQ ID NO: 48)**

TSSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSE  
TPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPG

**FIG. 11F: XTEN AE144-2A, secuencia ADN (SEQ ID NO: 49)**

GGCGCGCCAACCAAGTACGGAGCCGTCGAGGGGAGCGCACCAGGAAGCCCGGCTGGGAGCCCGACTTCTACCGAAGAGGGTACA  
TCTACCGAACCAAGTGAAGGTTTCAGCACCAGGCACCTCAACAGAACCCTCTGAGGGCTCGGCGCTGGTACAAGTGAGTCCGCCAC  
CCCAGAATCCGGGCTGGGACAAGCACAGAACCTTCGGAAGGGAGTGCCCCTGGAACATCCGAATCGGCAACCCAGAATCAGGG  
CCAGGATCTGAGCCCGCGACTTCGGGCTCCGAGACGCTGGGACATCCACCGAGCCCTCCGAAGGATCAGCCCCAGGCACCAGCA  
CGGAGCCCTCTGAGGGGAAGCGCACCCTGTACCAGCGAAAGCGCAACTCCCGAATCAGGTCCCGGTACGAGCGAGTCGGCGACCCC  
GGAGAGCGTCCGACGGTGCACGACG

**FIG. 11G: XTEN AE144-3B, secuencia proteínas (SEQ ID NO: 50)**

SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS  
APGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG

**FIG. 11H: XTEN AE144-3B, secuencia ADN (SEQ ID NO: 51)**

GGCGCGCCAAGTCCCCTGGAAGCCCACTAGCACCGAAGAGGGGACCTCAGAGTCCGCCACCCCGAGTCCGGCCCTGGCTCTG  
AGCCTGCCACTAGCGGCTCCGAGACTCTGGCACATCCGAAAGCGCTACACCCGAGAGTGGACCCGGCACCTCTACCGAGCCCACT  
GAGGGCTCCGCCCCTGGAACAAGACCGAGCCAGCGAAGGCAGCGCCCAAGGGACCTCCACAGAGCCCACTGAAGGCAGTGCT  
CCTGGCACCAAGCACCGAACCAAGCAGAGGGCTCTGCACCCGGGACCTCCACCGAGCCAAAGCGAAGGCTCTGCCCTGGCACTTCCA  
CCGAGCCCAAGCAGAGGCAGCGCCCTGGAGCCCCGCTGGCTCTCCACCAAGCATGAGGAGGGGCACATCTACCGAACCAAGTGA  
AGGCTCTGCACCAAGTGCTCCGAGC

**FIG. 11I: XTEN AE144-4A, secuencia proteínas (SEQ ID NO: 52)**

TSSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTST  
EEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG

**FIG. 11J: XTEN AE144-4A, secuencia ADN (SEQ ID NO: 53)**

GGCGCGCAACGTCCGAAAGTGCTACCCCTGAGTCAGGCCCTGGTAGTGAGCTGCCACAAGCGGAAGCGAAACTCCGGGGACCT  
CAGAGTCTGCCACTCCGAATCGGGGCCAGGCTCTGAACCGGCCACTTCAGGGAGCGAAACACAGGAACATCGGAGAGCGCTAC  
CCCGGAGAGCGGGCCAGGAACTAGTACTGAGCTAGCGAGGGAAGTGCACCTGGTACAAGCGAGTCCGCCACCCGAGTCTGG  
CCCTGGCTCTCCAGCGGGCTACCCACGAGCACTGAAGAGGGCTCTCCCGCTGGCAGCCCAACGTGACAGAAGAAGGATCACCA  
CGAGGCTCCCCACATCAACGAGGAGGGGTACATCAGAATCTGCTACTCCCGAGAGTGGACCCGGTACCTCCACTGAGCCCAGCG  
AGGGGAGTGCACAGGTGCCTCGAGC

**FIG. 11K: XTEN AE144-5A, secuencia proteínas (SEQ ID NO:54)**

TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGP  
GSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEG

**FIG. 11L: XTEN AE144-5A, secuencia ADN (SEQ ID NO:55)**

GGCGCGCCAAACATCAGAGAGCGCCACCCCTGAAAGTGGTCCCGGGAGCGAGCCACATCTGGGTGCGAAACGCCAGGCACAA  
GTGAGTCTGCAACTCCCGAGTCCCGACCTGGCTCCGAGCCTGCCACTAGCGGCTCCGAGACTCCGGGAACCTCCGAGAGCGCTACAC  
CAGAAAGCGGACCCGGAACAGTACCGAACCTAGCGAGGGCTCTGCTCCGGGAGCCAGCCGGCTCTCTACATCCACGAGGAG  
GGCACTTCCGAATCCGCCACCCCGAGTACGGGCCAGGATCTGAACCCGCTACCTCAGGCAGTGAGACGCCAGGAACGAGCGAGTC  
CGCTACACCGGAGAGTGGGCCAGGGAGCCCTGCTGGATCTCTACGTCCACTGAGGAAGGGTCACCAGCGGGCTCGCCACACAGCA  
CTGAAGAAGGTGCCTCGAGC

**FIG. 11M: XTEN AE144-6B, secuencia proteínas (SEQ ID NO:56)**

TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAP  
GTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG

**FIG. 11N: XTEN AE144-6B, secuencia ADN (SEQ ID NO:57)**

GGCGCGCCAAACATCTACCGAGCCTTCCGAAGGCTCTGCCCTGGGACCTCAGAATCTGCAACCCCTGAAAGCGGCCCTGGAACCTCC  
GAAAGTGCCACTCCCGAGAGCGGCCAGGGACAAGCGAGTCAGCAACCCCTGAGTCTGGACCCGGCAGCGAGCCTGCAACCTCTGG  
CTCAGAGACTCCCGGCTCAGAACCCTGCTACCTCAGGCTCCGAGACACCCGGCTCTCTGCTGGGAGTCCCACTCCACCGAGGAAGG  
AACATCCACTGAGCCTAGTGAGGGCTCTGCCCTGGAACAGCACAGAGCCAAGTGAGGGCAGTGACACAGGATCCGAGCCAGCAA  
CCAGCGGGTCCGAGACTCCCGGACCTCTGAGTCTGCCACCCAGAGAGCGGACCCGGCACTTCAACCGAGCCCTCGAAGGATCA  
GCACCAAGGTGCCTCGAGC

**FIG. 11O: XTEN AG144-1, secuencia proteínas (SEQ ID NO:58)**

PGSSPSASTGTGPGSSPSASTGTGPGTSGTASSSPGSSPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSPSGA  
TGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSS

**FIG. 11P: XTEN AG144-1, secuencia ADN (SEQ ID NO:59)**

GGCGCGCCACCCGGGTCGTCCCGTCGGCGTCCACCGGAACAGGGCCAGGGTCATCCCGTCAGCGTCGACTGGGACGGGACCCGG  
GACACCCGGTTCGGGGAGTCATCTCTCGCCTGGTTCGTCCACCCCGTCAGGAGCCACGGGTCGCCGGGAAGCAGCCCAAGCGC  
ATCCACTGGTACAGGGCCTGGGGCTTACCCGGGTACTTCATCCACGGGGTCACCGGGAACGCCGGATCGGGGACGGCTTCTCATC  
ACCAGGATCGTCAACACCCTCGGGCGCAACGGGCAGCCCCGGAACCCCTGGTTCGGGTACGGCGTCGTCGAGCCCCGGTGCGAGCC  
CGGGAACAAGCTCGACAGGATCGCCTGGGGCGTACCCGGCACGTGAGCACAGGCAGCCCCGGAACCCCTGGATCGGGAACCGC  
GTCGTCAAGCGCCTCGAGC

**FIG. 11Q: XTEN AG144-A, secuencia proteínas (SEQ ID NO:60)**

GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTSGTASSSPGSSPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTAS  
SSPGSSPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSP

**FIG. 11R: XTEN AG144-A, secuencia ADN (SEQ ID NO:61)**

GGCGCGCCAGGTGCCTCGCCGGGAACATCATCAACTGGTTACCCGGGTATCCCTCGGCCTCAACCGGGACGGGTCCCGGCTCA  
TCCCCAGCGCCAGCACTGGAACAGGTCCTGGCACTCTGGTTCGGTACGGCATCGTCATCCCGGGAAGCTCAACACCGTCCGGA  
GCGACAGGATCACCTGGCTCGTCACCTTGGCGTCAACTGGAACGGGGCCAGGGGCTCACCCGGAACGTCCTGACTGGGTGCGCT  
GGTACGCCGGGATCAGGAACGGCTCATCTCGCCTGGGTCTCAACGCCCTCGGGTGCGACTGGTTCGCCGGGAACCTCTGGCTCG  
GGGACGGCCTCGTCGCTGGGCGTACCCGGGACGAGTCCACGGGGTCCCTGGAGCGTACCCGGGACCTCTCGACAGG  
TAGCCCGGCTCGAGC



**FIG. 11S: XTEN AG144-B, secuencia proteínas (SEQ ID NO: 62)**

GTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGAT  
GSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSP

**FIG. 11T: XTEN AG144-B, secuencia ADN (SEQ ID NO: 63)**

GGCGCGCCAGGTACACCGGGCAGCGGCACGGCTTCGTCGTACCCGGCTCGTCCACACCGTCGGGAGCTACGGGAAGCCAGGAGC  
GTACCGGGAAACGTCGTCACCGGGGTACCGGGTACGCCAGGTAGCGGCACGCCAGCAGCTCGCCAGGTTTCATGACCCCGTCGG  
GAGCGACTGGGTCGCCCCGATCAAGCCCGTCAGCTTCCACTGGAACAGGACCCGGGTCGTCGCCGTCAGCCTCAACGGGGACAGGA  
CCTGGTTTCATGACCGCGTCAGGGGCGACAGGCTCGCCCGATCGTCAACACCTCGGGGGCAACGGGGAGCCCTGGTGCCTCGCC  
TGGAACCTCATCCACCGAAGCCCGGGGCTCGCCGGTACGAGCTCCACGGGATCGCCCGAGCGTCCCCGGAACCTCAAGCA  
CAGGGAGCCCTGCCTCGAGC

**FIG. 11U: XTEN AG144-C, secuencia proteínas (SEQ ID NO: 64)**

GTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSST  
GSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP

**FIG. 11V: XTEN AG144-C, secuencia ADN (SEQ ID NO: 65)**

GGCGCGCCAGGTACACCGGATCGGGTACAGCGTCATCGAGCCCGGTGCGTCACCTGGTACGTCGAGCACGGGGTCGCCAGGGGC  
GTCCCTGGGACGTCCTCAACAGGCTCGCCGGTGCCTACCCGGCAGCTCGTCCACGGGTTACCTGGTAGCTCCCTTCCGCGTCC  
ACTGGCACCGGGCTGGAACCTCGGGGAGCGGCACAGCGAGCTCGTCGCCGGGAGCATCGCCTGGGACATCGAGCACCGGGTCGC  
CAGGAGCATCGCCGGAACATCCAGCACAGGAAGCCCGCGCGTCGCCCGGACATCAAGCACAGGTTCCCCGGGATCGAGCACG  
CCGTCCGGAGCCACTGGATCACAGGGAGCTCGACACCTCCGGCGCAACGGGATCGCCCGAGCCAGCCCGGTACGTCAAGCAC  
TGGTCCCTGCCTCGAGC

**FIG. 11W: XTEN AG144-F, secuencia proteínas (SEQ ID NO: 66)**

GSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTG  
TGPPTGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP

**FIG. 11X: XTEN AG144-F, secuencia ADN (SEQ ID NO: 67)**

GGCGCGCCAGGCTCCAGCCCTCCGCGAGCACGGGAACCGGACAGGTTTCGTACCCCTAGCATCAACGGGGACGGGACCGGGGG  
CGTCACAGGAACGTCCTCCACGGGTCGCCGGGTGCATACCCGGAACGTCATCGACCGGATCGCCAGGGAGCTCGACGCCATCAG  
GCGCAACAGGATCACCTGGCTCAAGCCCTAGCGCGTCAACCGGCACGGGTCCGGGTGCCTCCCTGGCACGTCCAGCACCGGATCAC  
CCGGATCGAGCCATCCGCTCAACCGGAACCGGACCGGTACACAGGTCGGGAACAGCCTCCTCGTACACAGCTCCTCAACCC  
CCTCGGAGCCACGGGTTGCCCGGTTCTGTAACGCCTCCGGAGCAACTGGTAGCCCGGAGCATCGCCAGGAACCTTCGAGCACG  
GGGTCGCCCCGCTCGAGC

**FIG. 11Y: Promotor ET, secuencia proteínas (SEQ ID NO: 69)**

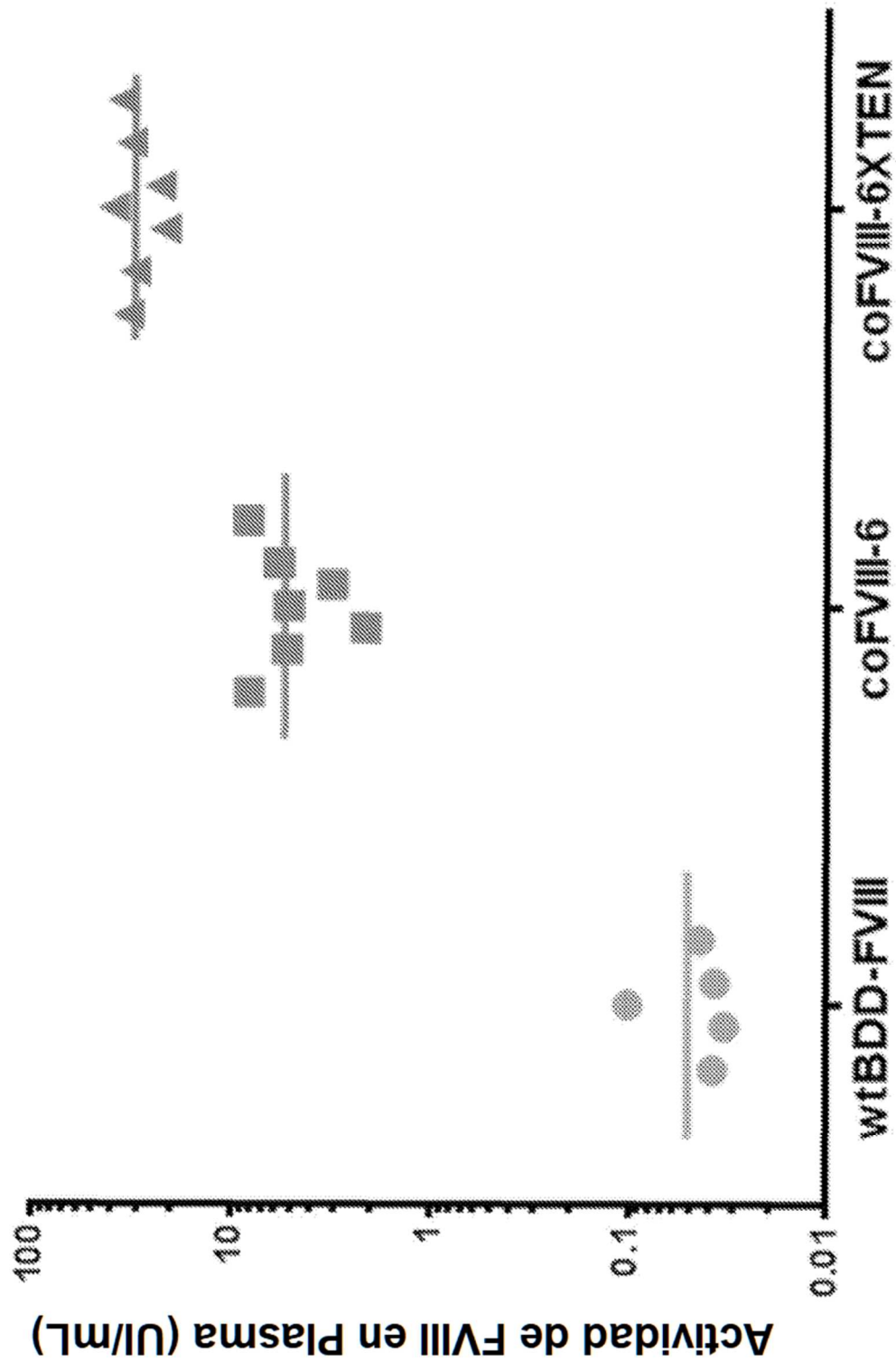
CTCGAGGTCAATTACGCGAGTTAATAATTACCAGCGCGGGCCAAATAATAATCCGCGAGGGGCAGGTGACGTTTGCCAGCGCGC  
GCTGGTAATTATTAACCTCGCGAATATTGATTGAGGCCGCGATTGCCGCAATCGCGAGGGGCAGGTGACCTTTGCCAGCGCGCT  
TCGCCCCGCCCGGACGGTATCGATAAGCTTAGGAGCTTGGGCTGCAGGTCGAGGGCACTGGGAGGATGTTGAGTAAGATGGAAAA  
CTACTGATGACCTTGCAGAGACAGAGTATTAGGACATGTTTGAACAGGGGCCGGGCGATCAGCAGGTAGCTCTAGAGGATCCCCGT  
CTGTCTGCACATTCGTAGAGCGAGTGTCCGATACTCTAATCTCCCTAGGCAAGGTTTATTTGTAGGTTACTTATTCTCTTTTG  
TTGACTAAGTCAATAATCAGAAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATAAAA  
GCCCCCTTACCAGGAGAAGCCGTACACAGATCCACAAGCTCCTGCCACCATGG

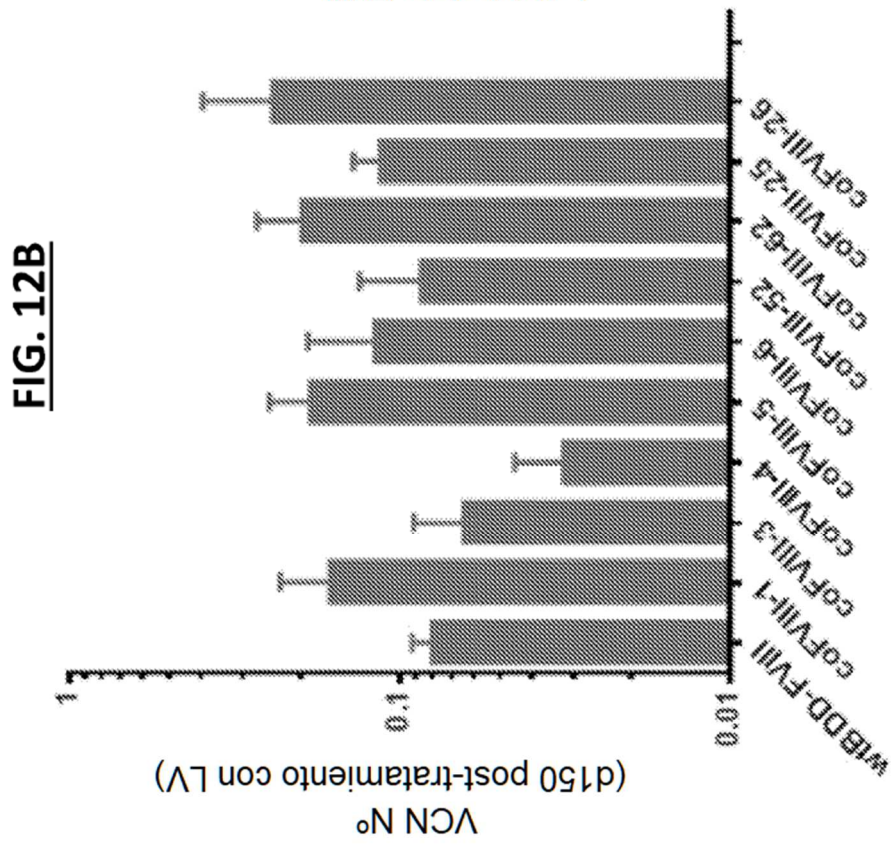
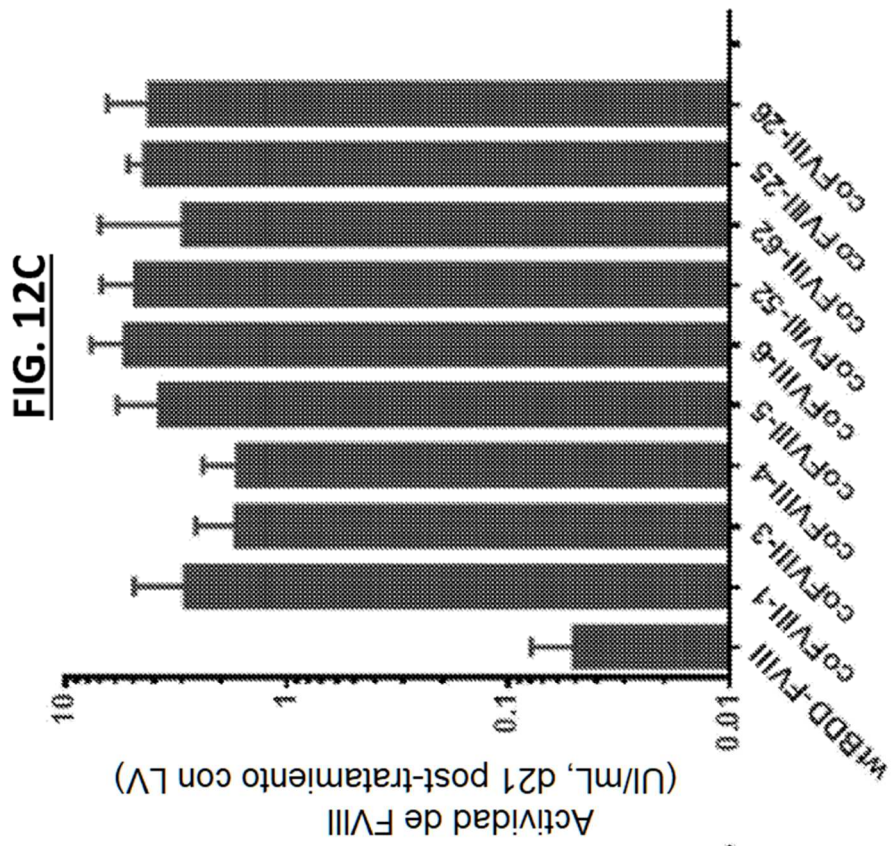
**FIG. 11Z: coFVIII-1 – Secuencia ADN (SEQ ID NO: 68)**

ATGCAGATTGAGCTGTCTACTTGTCTTTTCTGTGCCTGCTGAGGTTTTGCTTTTCCGCTACACGAAGGTATTATCTGGGGGCTGTGGA  
 ACTGTCTTGGGATTACATGCAGAGTGACCTGGGAGAGCTGCCAGTGGACGCAAGGTTTCCCCCTAGAGTCCCTAAGTCATTCCCCCTTC  
 AACACTAGCGTGGTCTACAAGAAAACACTGTTCTGTGGAGTTTACTGATCACCTGTTCAACATCGCAAAGCCTAGGCCACCCTGGATGG  
 GACTGCTGGGGCCAACAATCCAGGCCGAGGTGTACGACACCGTGGTCATTACACTTAAGAACATGGCCTCACACCCCGTGAGCCTGC  
 ATGCTGTGGGCGTCAGCTACTGGAAGGCTTCCGAAGGAGCAGAGTATGACGATCAGACTTCCAGAGAGAAAAAGAGGACGATAAG  
 GTGTTTCTGGCGGATCTCATACCTACGTGTGGCAGGTCTGAAAGAGAATGGCCCTATGGCCTCCGACCCTCTGTGCCTGACCTACT  
 CTTATCTGAGTCACGTGGACCTGGTCAAGGATCTGAACAGCGCCCTGATCGGAGCCCTGCTGGTGTGCAGGGAAGGAAGCCTGGCTA  
 AGGAGAAAACCCAGACACTGCATAAGTTCATTCTGCTGTTCCGCTGTTTGACGAAGGGAAATCATGGCACAGCGAGACAAAGAATA  
 GTCTGATGCAGGACAGGGATGCCGCTTCAGCCAGAGCTTGGCCCAAATGCACACTGTGAACGGCTACGTCAATCGCTCACTGCCTG  
 GGCTGATCGGCTGCCACCGAAAGAGCGTGTATTGGCATGTATCGGGATGGGCACCACACCTGAAGTGCATCCATTTCTCGGAGG  
 GACATACCTTTCTGGTCCGCAACCCAGACAGGCTTCCCTGGAGATCTCTCAATTACCTTCTGACAGCACAGACTCTGCTGATGGAC  
 CTGGGGCAGTTCCTGCTGTTTGGCCATCAGCTCCACCAGCATGATGGCATGGAGGCTTACGTGAAAGTGGACTCTTGTCCCGAGG  
 AACCTCAGCTGCGGATGAAGAACAATGAGGAAGCAGAAGACTATGACGATGACCTGACCGACTCCGAGATGGATGGTCCGATTTC  
 GATGACGATAACAGCCCTCTTTATCCAGATTAGATCTGTGGCCAAGAAACACCTAAGACATGGGTCCATTACATCGCAGCCGAGG  
 AAGAGGACTGGGATTATGCACCACTGGTCTGGCACCAGACGATCGCTCCTACAAATCTCAGTATCTGAACAATGGGCCACAGAGGA  
 TTGGCAGAAAGTACAAGAAAGTGCGGTTCATGGCATATACCGATGAGACCTTCAAGACTCGCAAGCCATCCAGCACGAGAGCGGCA  
 TCCTGGGACCACTGCTGTACGGAGAAGTGGGAGACACCTGCTGATCATTTTCAAGAACCAGGCCAGCCGGCCTTACAATATCTATCC  
 ACATGGGATTACAGATGTGCGCCCTCTGTACAGCAGGAGACTGCCAAAGGGCGTCAAACACCTGAAGGACTTCCCAATCTGCCCGG  
 AGAAATCTTCAAGTACAAGTGGACTGTACCCGTCGAGGATGGCCCCACTAAGAGCGACCTCGGTGCCTGACCCGCTACTATTCTAGT  
 TTCGTGAATATGGAAGAGATCTGGCAAGCGGACTGATCGGACCACTGCTGATTGTTACAAAGAGAGCGTGGATCAGAGAGGCAAC  
 CAGATCATGTCCGACAAGCGGAATGTGATTCTGTTCAGTGTCTTTGACGAAAACAGGTGATGGTACCTGACCGAGAACATCCAGAGAT  
 TCCTGCCTAATCCAGCTGGGGTGCAGCTGGAAGATCCTGAGTTTCAGGCATCTAACATCATGCATAGTATTAATGGCTACGTGTTTCA  
 CAGTTTGCAGCTGAGCGTGTGCTGCACGAGGTGCTTACTGGTATATCTGAGCATTGGGGCACAGACAGATTTCTGAGCGTGTTCT  
 TTTCCGGCTACACTTTTAAGCATAAAATGGTCTATGAGGACACACTGACTCTGTTCCCTTCAGCGCGAAACCGTGTTTATGAGCA  
 GGAGAATCCCGGACTGTGGATTCTGGGGTGCCACAACAGCGATTTCAGAAATCGCGGAATGACTGCCCTGTGAAAGTGTCAAGCTG  
 TGACAAGAACACCGGGGACTACTATGAAGATTCATACGAGGACATCAGCGCATATCTGCTGTCCAAAACAATGCCATTGAACCCCG  
 GTCTTTTGTAGTCAAGTCTCCAGTGTGAAGAGGCACAGAGGGAGATCACCCGCACTACCCTGCAGAGTGATCAGGAAGAGATCGA  
 CTACGACGATACAATTTCTGTGGAATGAAGAAAGAGGACTTCGATATCTATGACGAAGATGAGAACCAGAGTCTCTGATCATTCCAG  
 AAGAAAACCAAGGCATTACTTTATTGCCGAGTGGAGCGGCTGTGGGATTATGGCATGTCCTCTAGTCTCACGTGCTGCGAAATAGG  
 GCCCAGTCAGGAAGCGTCCCACAGTTCAAGAAAGTGGTCTTCCAGGAGTTTACAGACGGGTCTTTACTCAGCCACTGTACAGGGGC  
 GAAGTGAACGAGCACCTGGGACTGCTGGGGCCCTATATCAGAGCAGAAGTGGAGGATAACATTATGGTCACTTCAGAAATCAGGCC  
 TCTCGGCCTTACAGTTTTTATTCAAGCTGATCTTACGAAGAGACACGACAGGAGCTGAACACGAAAAAACTTCGTGAAGC  
 CTAATGAGACCAAAACATATTTTGAAGGTGCAGACCATATGGCCCCAACAAAAGACGAGTTTCGATTGCAAGGCATGGGCCTATTT  
 TTCTGACGTGGATCTGGAGAAGGACGTGCACAGTGGCCTGATTGGCCCACTGCTGGTGTGCCATACTAACCCCTGAATCCAGCCAC  
 GGCCGGCAGGTCACTGTCCAGGAGTTCGCTGTTCTTTACCATCTTTGATGAGACAAAGAGCTGGTACTTCACCGAAAAATGGAGC  
 GAAATTGCAGGGCTCCATGTAACATTAGATGGAAGACCCACATTCAAGGAGAACTACCGCTTTCATGCTATCAATGGATACATCAT  
 GGATACTGTGCCGGGCTGGTGCATGGCACAGGACAGAGAATCCGGTGGTATCTGCTGAGCATGGGCAGCAACGAGAATATCCACTC  
 AATTCAATTCAGCGGGCACGTGTTTACTGTGAGGAAGAAAGAGTACAAGATGGCCCTGTACAACCTGTATCCCGCGGTGTTTCGAA  
 ACCGTCGAGATGCTGCCTAGCAAGGCCGGAATCTGGAGAGTGGAAATGCCTGATTGGAGAGCACCTGCATGCTGGGATGTCTACCTG  
 TTTCTGGTGTACAGTAATAAGTGTGACACCCCTGGGAATGGCATCCGGGCATATCAGGGATTTCCAGATTACCGCATCTGGACAGT  
 ACGGACAGTGGGCACCTAAGCTGGCTAGACTGCATATTCCGGATCTATCAACGCTTGGTCCACAAAAGAGCCTTTCTCTTGATTAA  
 GGTGGACCTGCTGGCCCCAATGATCATTCATGCCATCAAAACCTAGGAGCTCGGCAGAAAGTTCCTCTCTGTACATCTCACAGTTTA  
 TCATCATGTACAGCCTGGATGGGAAGAAATGGCAGACATACCGCGGCAATAGCACAGGAACCTGATGGTGTCTTTGGCAACGTGG  
 ACAGCAGCGGAATCAAGCACAACTTTCAATCCCCATCATTTGCTAGATACATCCGGCTGCACCAACCCATTATTCTATTGAAAGT  
 AACTGAGGATGGAAGTGTGGGATGCGATCTGAACAGTTGTTCAATGCCCTGGGGATGGAGTCAAGGCAATCTCTGACGCCAG  
 ATTACCGCAGCTCCTACTTCACTAATATGTTTGTACCTGGAGCCCTTCCAAAGCAAGACTGCACCTGCAAGGCCGACGCAACGCATG  
 GCGACCACAGGTGAACAATCCCAAGGAGTGGTTGCAGGTGATTTTCAGAAAATATGAAGGTGACCGGGGTCAAACTCAGGGCG  
 TGAAAAGTCTGCTGACCTCAATGTACGTCAAGGAGTTCCTGATCTCTAGTTACAGGACGGACATCAGTGGACACTGTTCTTTCAGAA  
 CGGGAAGGTGAAAGTCTTCCAGGGCAATCAGGATTCCTTTACACCTGTGGTCAACAGTCTAGACCTCCACTGCTGACCAGATACCTG  
 AGAATCCACCCTCAGTCTGGGTGCACAGATTGCCCTGAGAATGGAAGTGTGGGATGCGAGGCCAGGATCTGTACTGA

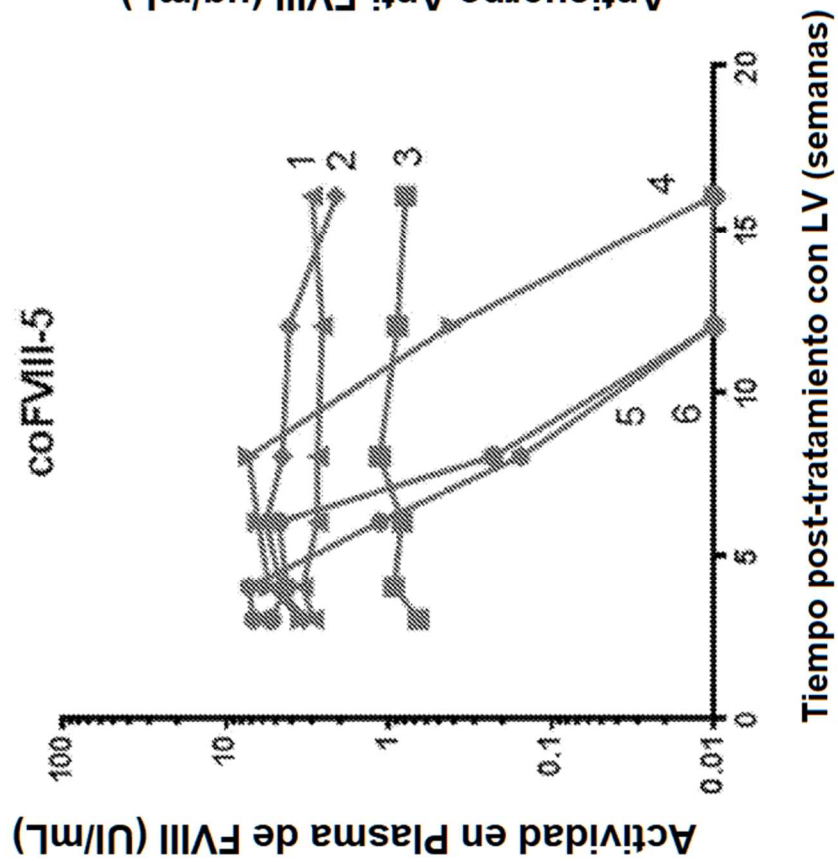


FIG. 12A: Inyección IV de Ratones HemA con LV-FVIII

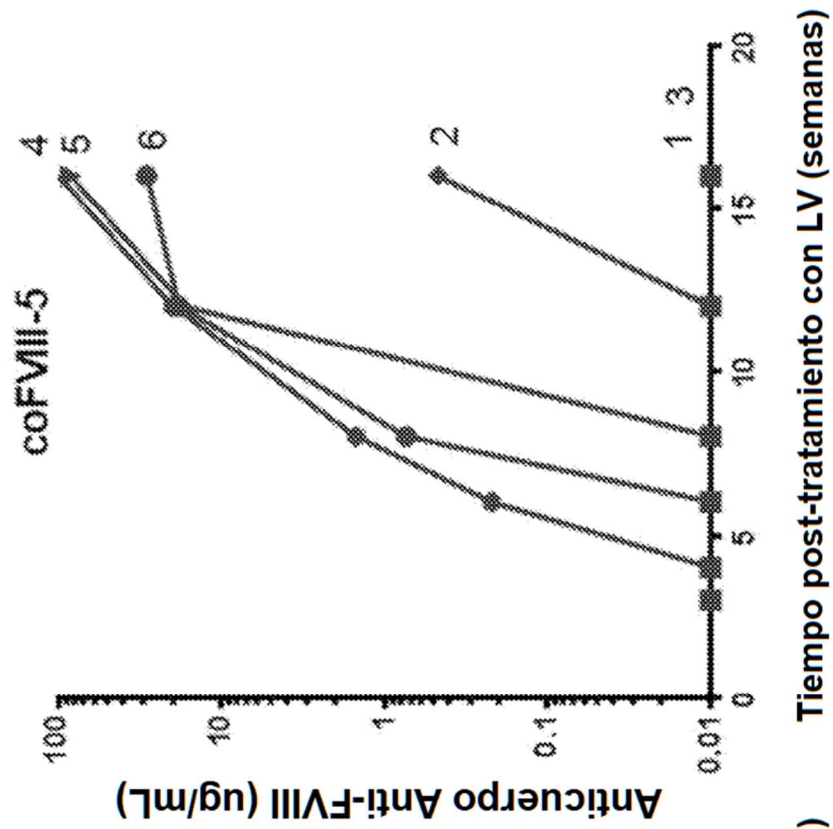


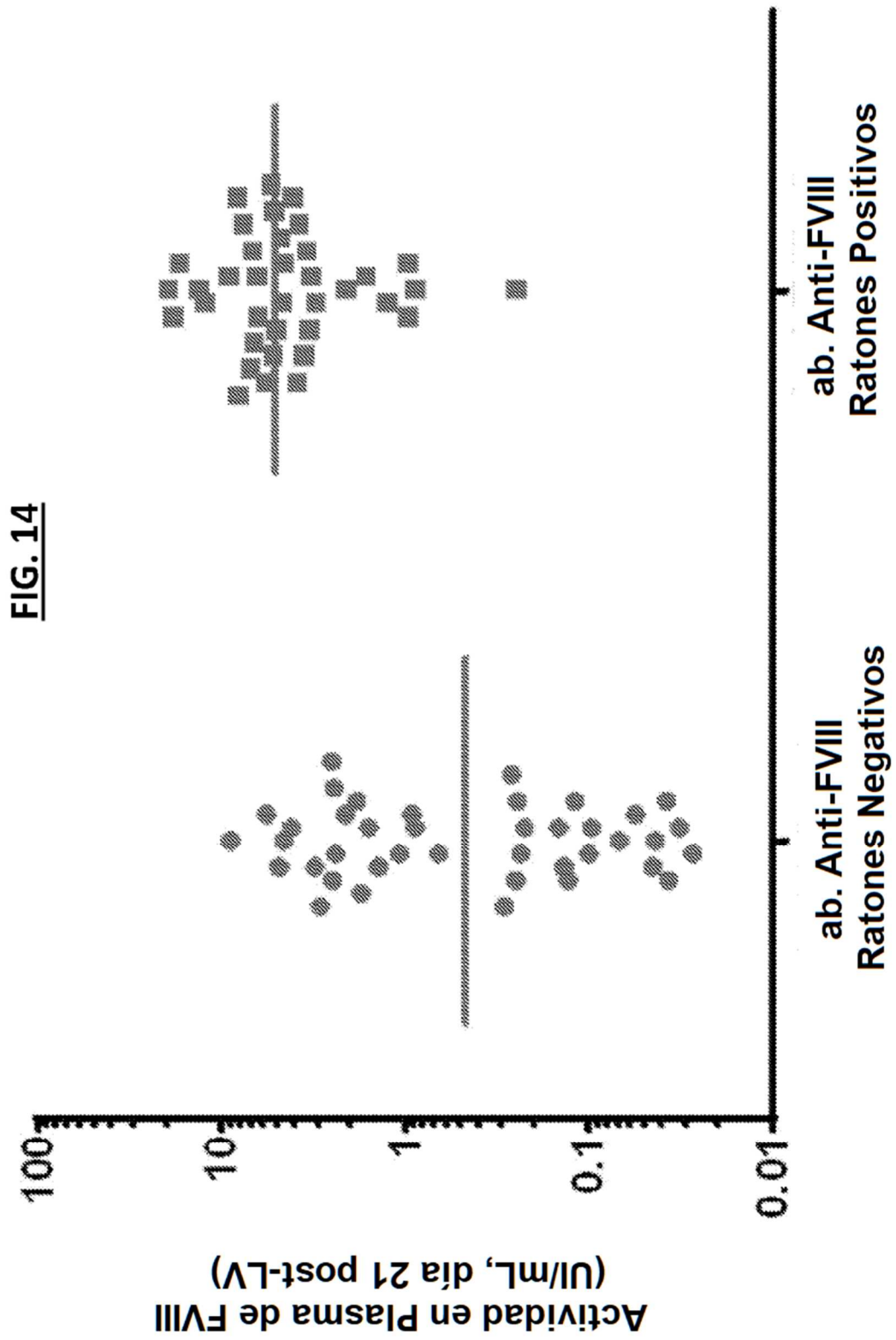


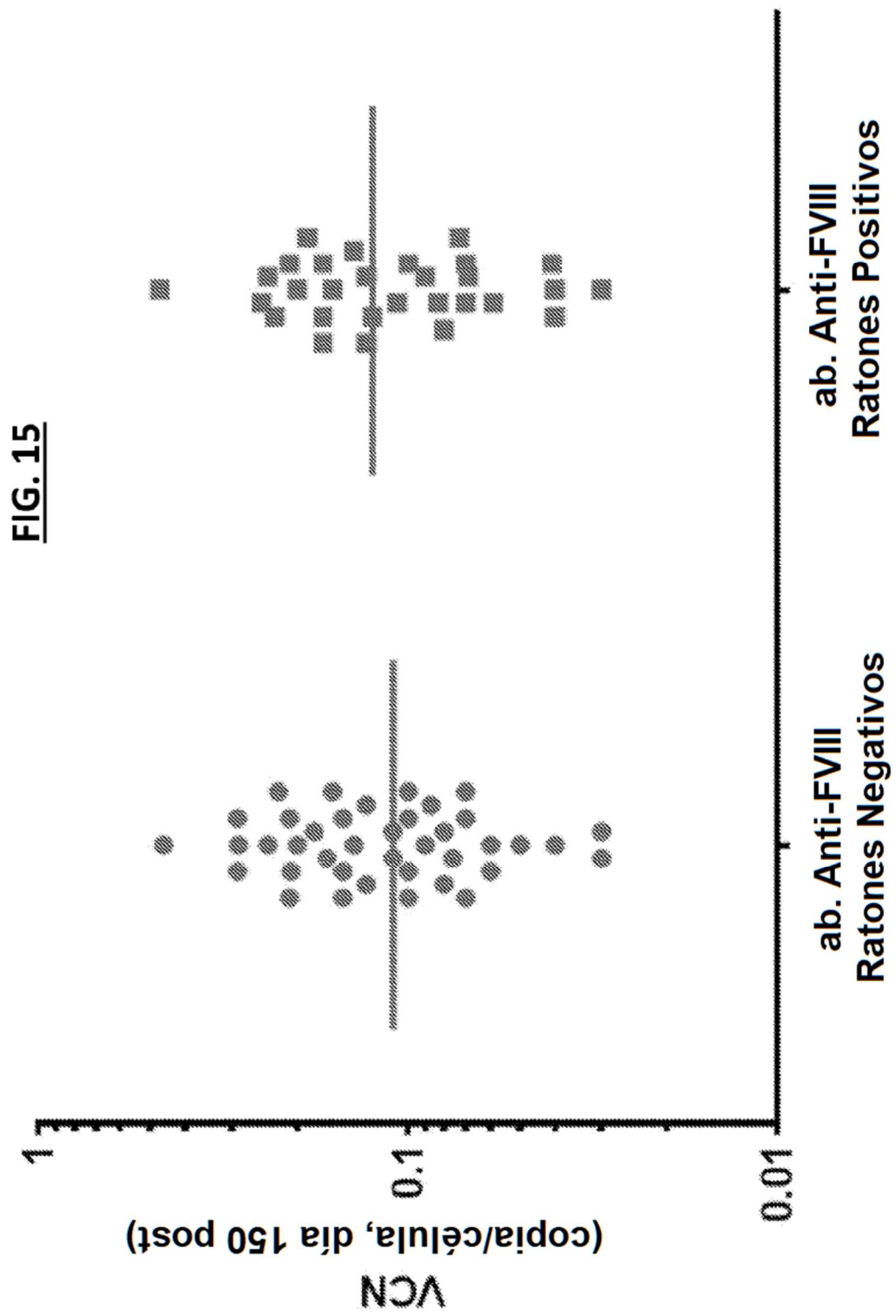
**FIG. 13A**



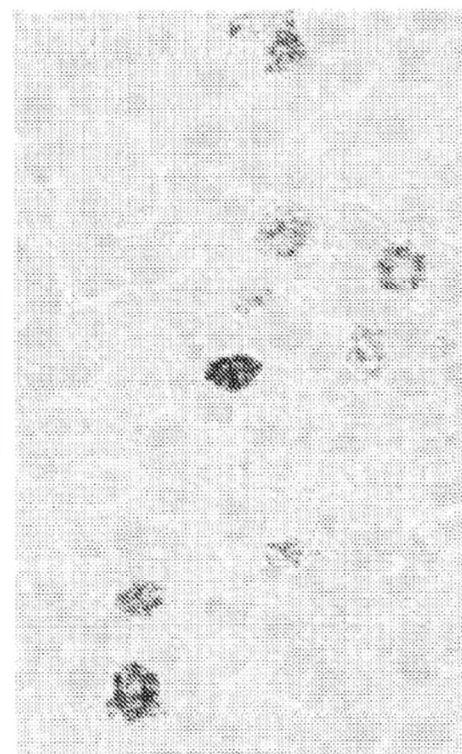
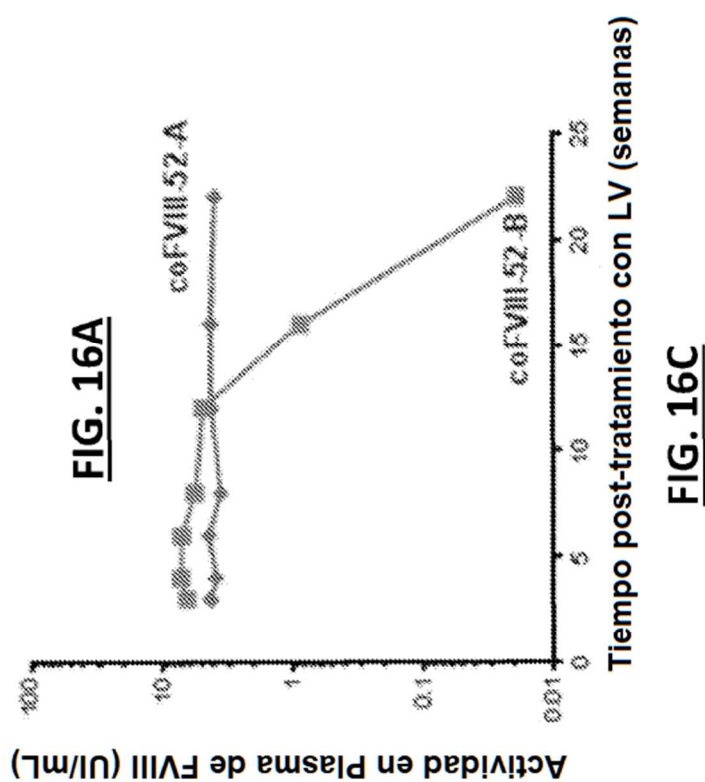
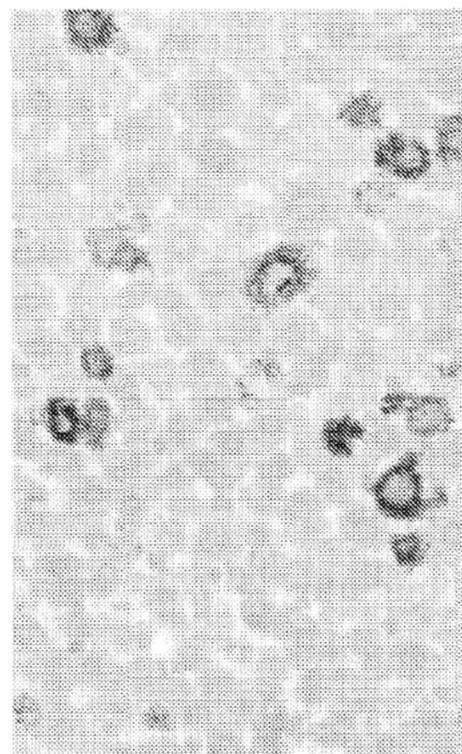
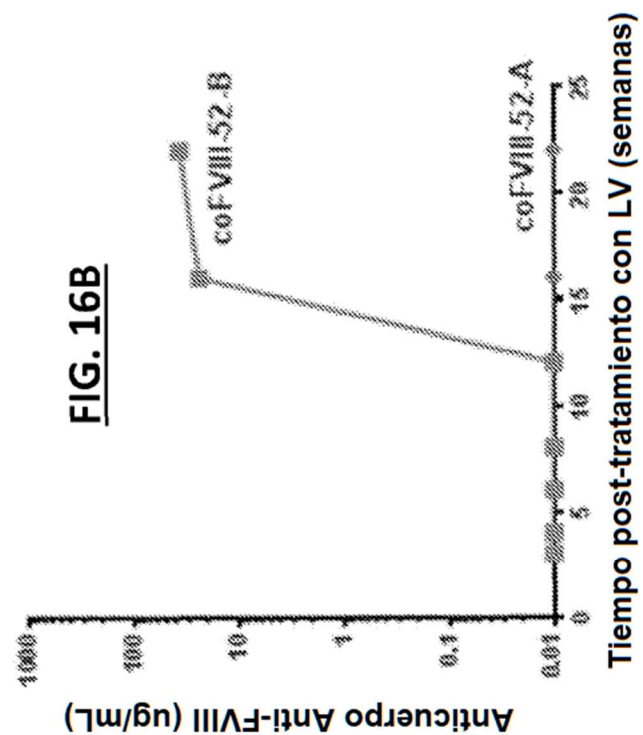
**FIG. 13B**

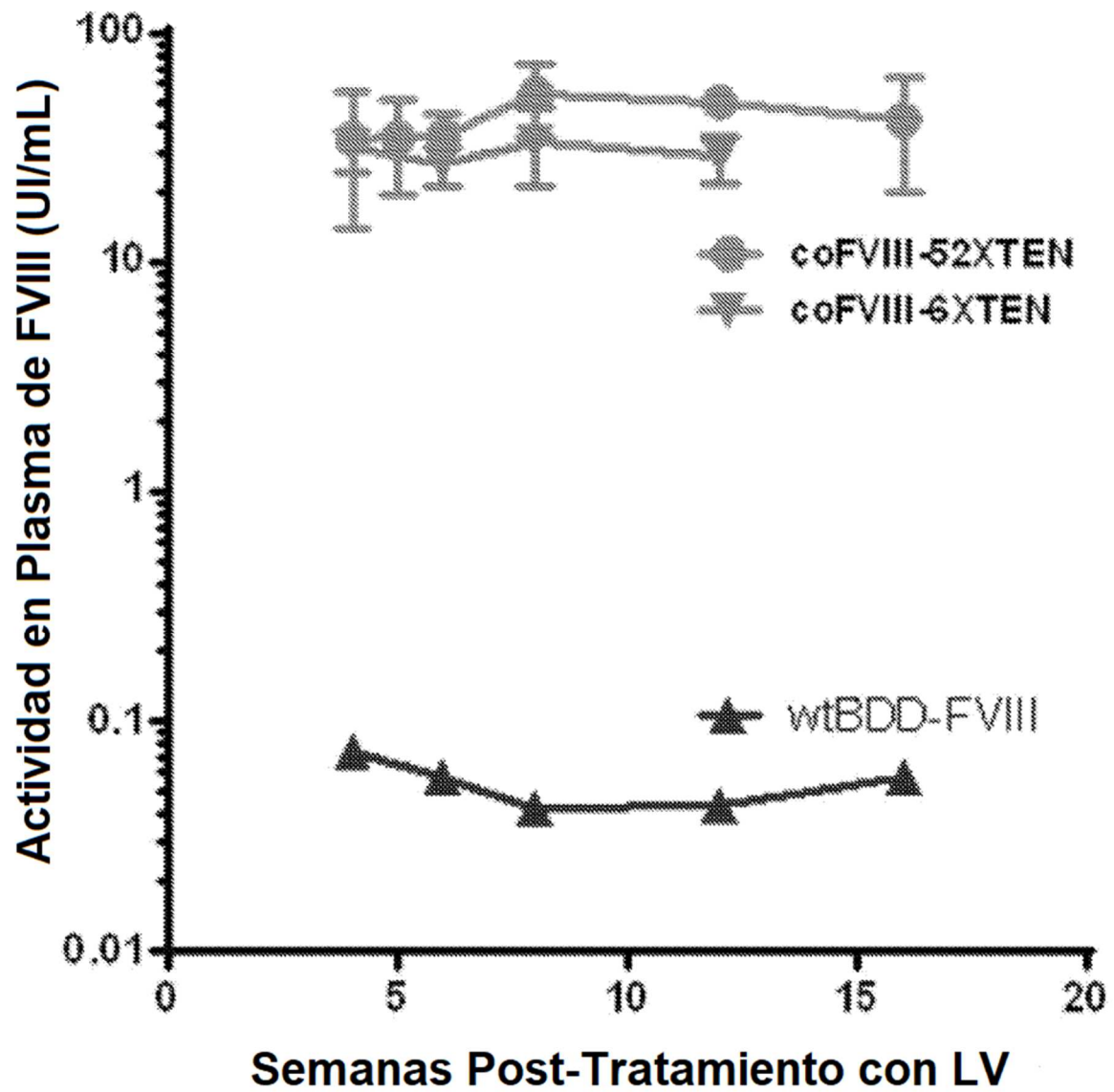


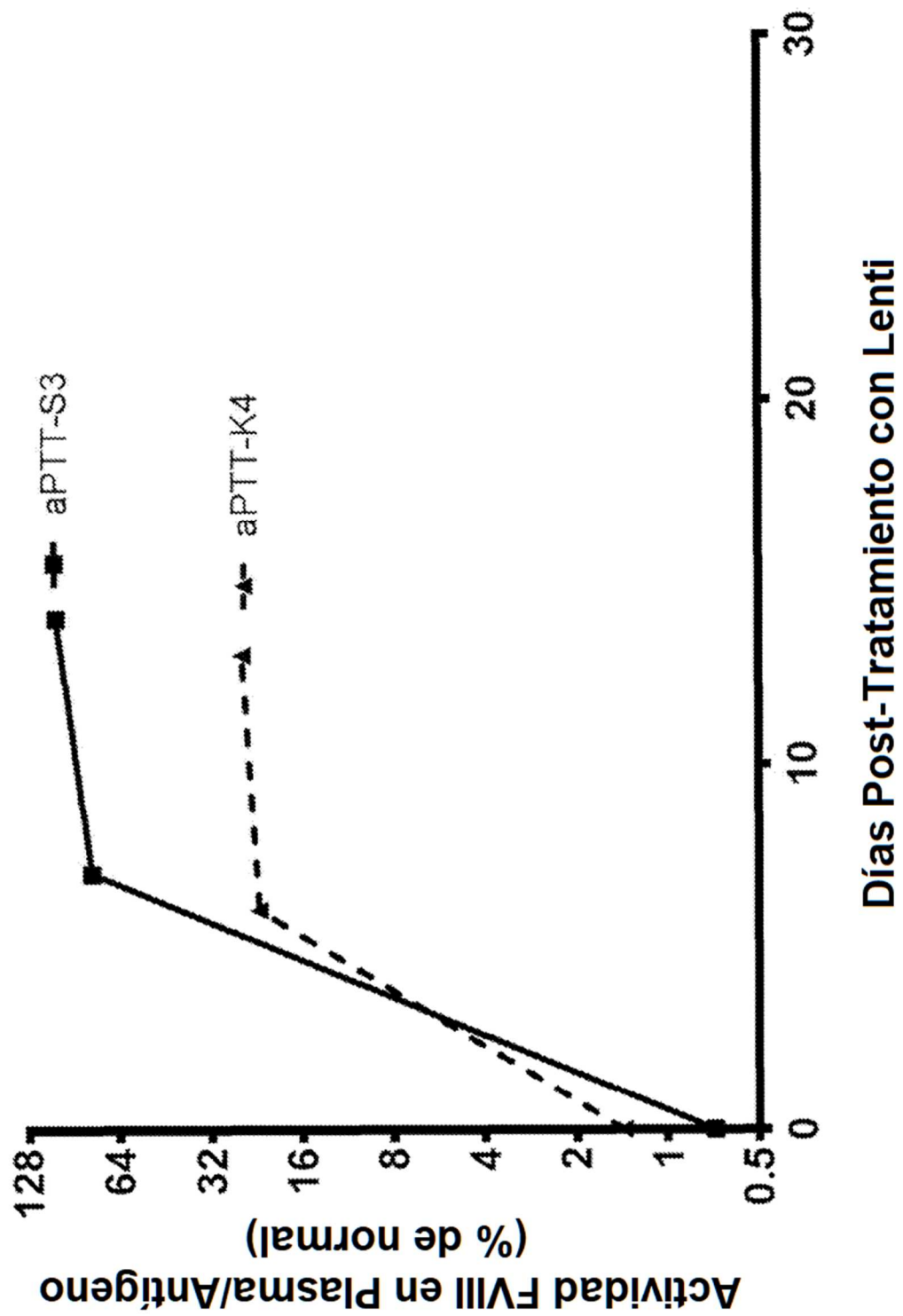






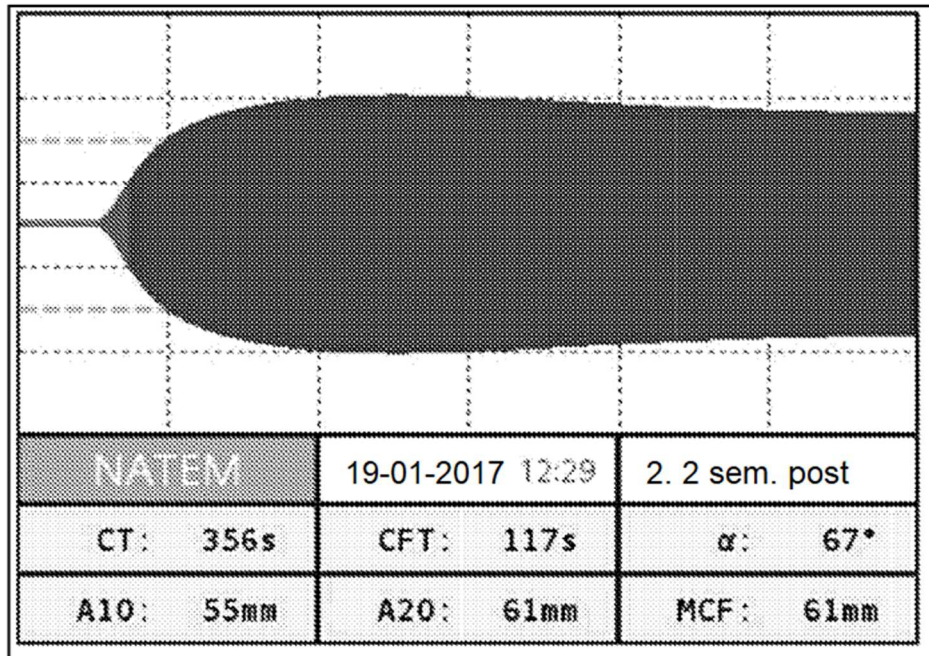


**FIG. 17**

**FIG. 18**





**FIG. 19C****FIG. 19D**

Parámetro	Descripción parámetro	Intervalo normal
CT (s)	tiempo coagulación	120-480
CFT(s)	tiempo formación coágulo	60-240
$\alpha(^{\circ})$	ángulo alfa	27-60
A5 (mm)	amplitud 5 min tras CT	na
A20 (mm)	amplitud 20 min tras CT	na
MCF (mm)	firmeza máxima del coágulo	40-60