

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 394 722**

51 Int. Cl.:

**C07K 16/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.12.2006 E 06851829 (9)**

97 Fecha y número de publicación de la solicitud europea: **15.10.2008 EP 1979378**

54 Título: **Nuevo conjunto de regiones HCDR3 y utilizaciones de las mismas**

30 Prioridad:

**20.12.2005 US 751633 P**

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

**05.02.2013**

73 Titular/es:

**MORPHOSYS AG (100.0%)  
LENA-CHRIST-STRASSE 48  
82152 MARTINSRIED/MÜNCHEN, DE**

72 Inventor/es:

**ENZELBERGER, MARKUS;  
THIEL, STEFANIE;  
PRASSLER JOSEF;  
URLINGER STEFANIE y  
ROTHE CHRISTINE**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

**ES 2 394 722 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

Nuevo conjunto de regiones HCDR3 y utilizaciones de las mismas.

**Antecedentes de la Invención**

5 La tercera región determinante de la complementariedad de la cadena pesada de inmunoglobulina (H-CDR3) constituye el centro del sitio clásico de fijación de antígeno y juega a menudo un papel dominante en la determinación de la especificidad y afinidad del anticuerpo. En todas las especies conocidas con un sistema inmune adaptativo, H-CDR3 es más diversa que cualquiera de las otras cinco regiones CDR (H-CDR1, H-CDR2, L-CDR1, L-CDR2 o L-CDR3) que, juntas, forman el borde exterior del sitio de fijación de antígeno (*Tonegawa, 1983; Chothia et al., 1989*).

10 Se ha encontrado que tanto la longitud media de H-CDR3 como el intervalo de longitudes de H-CDR3, que pueden tener gran influencia en la gama de estructuras de fijación de antígeno disponibles para las especies y por consiguiente la función del repertorio de anticuerpos (*Johnson & Wu, 1998; Collis et al., 2003*), aumenta de los ratones a los humanos y a ganado vacuno (*Wu et al., 1993*). La diversidad de H-CDR3 puede estar regulada también dentro de una especie. Por ejemplo, en el hombre y el ratón, la gama de longitudes de H-CDR3 y la  
15 diversidad de H-CDR3 aumenta durante la ontogenia (*Feeney, 1992, Zemlin et al., 2002*).

Debido a su papel prominente en la fijación de antígeno, la región H-CDR3 ha sido utilizada como vehículo para la introducción de hipervariabilidad en bibliotecas de anticuerpos sintéticos, que pueden utilizarse para la transformación por ingeniería genética de anticuerpos terapéuticos (*Knappik et al., 2000*). La aleatorización de esta  
20 región con cebadores degenerados produce una diversidad enorme de secuencias, pero muchas, sino la mayoría, de estas estructuras están distorsionadas y no son funcionales (*Zemlin et al., 2003*). Por consiguiente, se requieren bibliotecas de anticuerpos sintéticos muy grandes para obtener afinidades satisfactorias contra una diana dada.

Zemlin observó que las regiones H-CDR3 en humanos exhiben una mayor gama de longitudes comparadas con las regiones H-CDR3 murinas, después del análisis por Zemlin de regiones H-CDR3 publicadas, singulares y  
25 funcionales murinas y humanas. Zemlin observó también que la frecuencia de ciertos aminoácidos cambia a medida que aumenta la longitud de la H-CDR3. Por ejemplo, Zemlin consignó que la frecuencia de serina aumentaba con la longitud para secuencias de 8-14 residuos de aminoácidos, pero exhibía un patrón mixto en secuencias más largas que 14 residuos de aminoácidos.

Recientemente, Hoet et al., *Nature Biotech* 23:(3), marzo 2005, describieron la construcción de una biblioteca de anticuerpos que contenía secuencias variables pesadas (VH) que eran parcialmente sintéticas, pero contenían  
30 regiones HCDR3 capturadas de donantes humanos. En particular, la porción de la cadena VH que no incluye la región HCDR3 (FR1-H-CDR1-FR2-H-CDR2-FR3) se obtuvo sintéticamente y se incorporó en un gen VH; y la porción de la cadena VH que incluye la región HCDR3 (H-CDR3-FR4) se derivaba de pacientes autoinmunes humanos existentes naturalmente.

Hoet optó por incorporar secuencias HCDR3 capturadas de donantes humanos sobre la base de que un grado  
35 similar de diversidad funcional de biblioteca y por consiguiente de calidad, no podría conseguirse por incorporación de una región HCDR3 creada sintéticamente en una cadena VH. Esto no era sorprendente, dado que otros han elegido el enfoque de generar las así llamadas bibliotecas HCDR3 enfocadas (véase, v.g., la solicitud de patente publicada US 2006/0257937A1), en la cual la diversidad está controlada (por limitación del análisis a segmentos V, D y J conocidos) en un esfuerzo para reducir la concentración de ligantes no funcionales que, con arreglo al saber  
40 convencional, se espera que sea alta en las bibliotecas que están orientadas a incluir la mayor diversidad posible, haciendo con ello el cribado de las bibliotecas más complicado de lo que tendría que ser (véase la solicitud de patente '937A1). De hecho, Hoet indicó no sólo que en los anticuerpos humanos existentes naturalmente, HCDR3 varía desde 4 a más de 35 residuos y tiene una diversidad de secuencia no aleatoria, sino también que es imposible sintetizar DNA que codifique a la vez la secuencia y la diversidad de longitud encontradas en los repertorios de H-  
45 CDR3 naturales. A este fin, Hoet pretende describir el uso de una biblioteca de HCDR3 dado a conocer en la solicitud de patente '937A1 indicando que los segmentos D de la línea germinal han sido seleccionados para fomentar el plegado y la fijación apropiados de los miembros de la biblioteca HCDR3, y que es deseable la inclusión de segmentos D en una biblioteca. Un resultado práctico de una biblioteca enfocada expuesta en la solicitud de patente '937A1, consistente en que la misma está construida con vistas a un conjunto de datos limitado, a saber los  
50 segmentos conocidos V, D y J, es que aquélla no consigue obtener una diversidad óptima de diversificación para cada residuo de aminoácidos.

En contraste con la doctrina de Hoet y la solicitud de patente '937A1, la presente invención proporciona, *inter alia*, un conjunto de regiones H-CDR3 de anticuerpos sintéticos humanos que tienen una diversidad esencialmente como  
55 la encontrada en los repertorios naturales de H-CDR3, mimetizando la distribución natural de aminoácidos por sesgo de la distribución aleatoria completa de aminoácidos en la secuencia de DNA codificante de H-CDR3. En este sentido, la invención proporciona una biblioteca de anticuerpos que está construida de una manera que se creía hasta ahora imposible de construir.

**Sumario de la Invención**

La invención proporciona un conjunto de regiones H-CDR3 diversas de anticuerpos humanos, en las cuales la diversidad del conjunto se genera por diversificación de regiones H-CDR3 que tienen una longitud comprendida dentro de un primer intervalo de aminoácidos de acuerdo con un primer factor de diversidad, por diversificación de regiones H-CDR3 que tienen una longitud comprendida dentro de un segundo intervalo de aminoácidos de acuerdo con un segundo factor de diversidad y por diversificación de regiones H-CDR3 que tienen una longitud dentro de un tercer intervalo de aminoácidos de acuerdo con un tercer factor de diversidad, en donde los factores de diversidad son diferentes, y en donde dicho primer intervalo de aminoácidos es de 4 a 7 aminoácidos, el segundo intervalo de aminoácidos es de 8 a 14 aminoácidos, y el tercer intervalo de aminoácidos es de 15 a 23 aminoácidos, y en donde dicho primer intervalo de aminoácidos tiene una composición como la representada en la Tabla 4, dicho segundo intervalo de aminoácidos tiene una composición como la representada en la Tabla 9 y dicho tercer intervalo de aminoácidos tiene una composición como la representada en la Tabla 14.

La presente invención proporciona también un conjunto de diversos anticuerpos humanos o fragmentos funcionales de los mismos, que comprende un conjunto de cadenas pesadas variables de anticuerpos humanos que comprende un conjunto de diversas regiones H-CDR3 de anticuerpos humanos de acuerdo con la presente invención.

Se describe también en esta memoria una biblioteca sintética de anticuerpos humanos, que comprende un conjunto de diversos anticuerpos humanos o fragmentos funciones de los mismos que comprenden un conjunto de cadenas pesadas variables de anticuerpos humanos de acuerdo con la presente invención.

La invención se refiere también a un conjunto de ácidos nucleicos que codifican un conjunto de regiones H-CDR3 de anticuerpos humanos diversos de acuerdo con la presente invención.

Se describe también un método de obtención de un conjunto de diversas regiones H-CDR3 de anticuerpos humanos de longitudes variables que comprenden la expresión del conjunto de ácidos nucleicos.

En el contexto de la presente invención, debe entenderse que una o más posiciones de aminoácidos en una región H-CDR3 diversa por lo demás pueden mantenerse constantes. El operario experto apreciará que para la eficiencia de la síntesis de mezclas de trinucleótidos a fin de introducir diversidad dentro de las regiones H-CDR3 y para eficiencia de clonación de las diversas regiones H-CDR3 diversificadas, puede ser deseable insertar una o más posiciones constantes dentro de la región H-CDR3.

Se describe también en esta memoria un conjunto de diversas regiones H-CDR3 humanas, en donde si los miembros del conjunto tienen una longitud de 4 a 8 aminoácidos, entonces el o los factores de diversidad que se aplican a dichos miembros que tienen dicha longitud de 4 a 8 aminoácidos definen 0, 1, 2 ó 3 posiciones constantes de aminoácidos dentro de las regiones H-CDR3 de dichos miembros que tienen dicha longitud. Se describe también un conjunto de diversas regiones H-CDR3 humanas o humanizadas, en donde si los miembros del conjunto tienen una longitud de 9 a 10 aminoácidos, entonces el o los factores de diversidad que se aplican a dichos miembros que tienen dicha longitud de 9 a 10 aminoácidos definen 0, 1, 2, ó 3 posiciones constantes de aminoácidos dentro de las regiones H-CDR3 de dichos miembros que tienen dicha longitud.

Se describe también un conjunto de diversas regiones H-CDR3 humanas o humanizadas, en donde si los miembros del conjunto tienen una longitud de 11 a 13 aminoácidos, entonces el o los factores de diversidad que se aplican a dichos miembros que tienen dicha longitud de 11 a 13 aminoácidos definen 0, 1, 2 ó 3 posiciones constantes de aminoácidos dentro de las regiones H-CDR3 de dichos miembros que tienen dicha longitud. Se describe también un conjunto de diversas regiones H-CDR3 humanas o humanizadas, en donde si los miembros del conjunto tienen una longitud de 14 aminoácidos, entonces el o los factores de diversidad que se aplican a dichos miembros que tienen dicha longitud de 14 aminoácidos definen 0, 1, 2, 3, 4 ó 5 posiciones constantes de aminoácidos dentro de las regiones H-CDR3 de dichos miembros que tienen dicha longitud.

Se describe también un conjunto de diversas regiones H-CDR3 humanas, en donde si los miembros del conjunto tienen una longitud de 15 a 16 aminoácidos, entonces el o los factores de diversidad que se aplican a dichos miembros que tienen dicha longitud de 15 a 16 definen 0, 1, 2, 3, 4, 5, 6, 7 u 8 posiciones constantes de aminoácidos dentro de las regiones H-CDR3 de dichos miembros que tienen dicha longitud.

Se describe también un conjunto de diversas regiones H-CDR3 humanas o humanizadas, en donde si los miembros del conjunto tienen una longitud de 16 a 17 aminoácidos, entonces el o los factores de diversidad que se aplican a dichos miembros que tienen dicha longitud de 16 a 17 aminoácidos definen 0, 1, 2, 3, 4, 5, 6, 7 u 8 posiciones constantes de aminoácidos dentro de las regiones H-CDR3 de dichos miembros que tienen dicha longitud.

Se describe también un conjunto de diversas regiones H-CDR3 humanas o humanizadas, en donde si los miembros del conjunto tienen una longitud de 18 aminoácidos, entonces el o los factores de diversidad que se aplican a dichos miembros que tienen dicha longitud de 18 aminoácidos definen 0, 1, 2, 3, 4, 5, 6, 7 u 8 posiciones constantes de aminoácidos dentro de las regiones H-CDR3 de dichos miembros que tienen dicha longitud.

Se describe también un conjunto de diversas regiones H-CDR3 humanas o humanizadas, en donde si los miembros del conjunto tienen una longitud de 19 aminoácidos o mayor, entonces la diversidad de dichos miembros que tienen dicha longitud de 19 aminoácidos o mayor definen 0, 1, 2, 3, 4, 5, 6 ó 7 posiciones constantes de aminoácidos dentro de las posiciones H-CDR3 de dichos miembros que tienen dicha longitud.

5 **Sumario de las figuras:**

Fig. 1: Alineación a la izquierda; cada paso en la “escalera” descrita en la Figura 1b representa la secuencia de consenso de 380 secuencias por término medio con una longitud H-CDR3 definida. En la alineación izquierda, las longitudes diferentes se disponen por entrada de residuos de aminoácidos adicionales a lo largo de la diagonal como en el caso de una longitud creciente de izquierda a derecha.

10 Fig. 2: Alineación derecha; cada paso en la “escalera” descrita en la Figura 2b representa la secuencia de consenso de 380 secuencias por término medio con una longitud H-CDR3 definida. En la alineación derecha, las longitudes diferentes se disponen por entrada de residuos de aminoácidos adicionales a lo largo de la diagonal como en el caso de una longitud creciente de derecha a izquierda.

15 Fig. 3a + 3b + 3c: Alineación centrada (región flanqueante, región de bucle, mixta-centrada); cada capa en la figura representada en la figura de la derecha representa la secuencia de consenso de 380 secuencias por término medio con una longitud H-CDR3 definida. En la alineación de la región flanqueante, las diferentes longitudes se disponen por entrada de residuos de aminoácidos adicionales alternantes a lo largo de las diagonales como en el caso de una longitud creciente desde la pare central externa a la interna (desde la región flanqueante hacia la región de bucle). Los cuadrados grises oscuros son ejemplos de nuevos residuos añadidos.

20

Fig. 4: Ejemplos de esquemas de numeración de regiones variables de anticuerpos.

**Descripción detallada de la invención:**

En un aspecto, la presente invención proporciona, *inter alia*, un conjunto de regiones H-CDR3 sintéticas de anticuerpos humanos que tienen una diversidad esencialmente como la encontrada en los repertorios H-CDR3 naturales, mimetizando la distribución natural de aminoácidos; esto se consigue por sesgo de la distribución aleatoria de aminoácidos en la secuencia de DNA codificante de H-CDR3 de acuerdo con la longitud de la región H-CDR particular.

25

En lugar de utilizar sólo una serie limitada de datos, a saber segmentos V, D y J conocidos, pueden analizarse secuencias de anticuerpos humanos asequibles de bases de datos disponibles comercialmente (*v.g.*, Ig-BLAST) y utilizarse como base para determinación de la distribución de aminoácidos en las regiones H-CDR3.

30

En este sentido, la invención proporciona una biblioteca de anticuerpos que se construye de una manera que se consideraba hasta ahora imposible de construir.

En el contexto de la presente invención, el término “factor de diversidad” se define como, para cualquier intervalo dado de residuos de aminoácidos en cada región H-CDR3, la tasa de frecuencia para la cual los aminoácidos aparecen en una o más posiciones dadas en dicha región H-CDR3. En el contexto de la presente invención, la “tasa de frecuencia” de un aminoácido en una posición dada se define como la probabilidad de que dicho aminoácido aparezca en tal posición, donde la frecuencia puede ser de 0% a 100%, de acuerdo con la doctrina expuesta en esta memoria. En los anticuerpos humanos existentes naturalmente, las regiones H-CDR3 varían desde aproximadamente 4 a 35 aminoácidos de longitud. De acuerdo con ello, la descripción contempla que la longitud de los residuos de aminoácidos en cada región H-CDR3 puede estar comprendida entre 4 y 35, y preferiblemente es 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 ó 22 residuos de aminoácidos de longitud, y el intervalo de aminoácidos al que se aplica un factor de diversidad particular puede variar, como se describe más detalladamente en esta memoria. Los anticuerpos y/o regiones H-CDR3 de anticuerpos de la invención, que pueden ser humanos, pueden utilizarse en muchos contextos, que se describen más detalladamente en esta memoria.

35

40

La diversidad en un conjunto de regiones de anticuerpos humanos H-CDR3 diversas puede generarse por aplicación de: un primer factor de diversidad para los miembros de la región H-CDR3 que tengan una longitud de aproximadamente 4 a aproximadamente 9 aminoácidos, un segundo factor de diversidad para los miembros de la región H-CDR3 que tengan una longitud de aproximadamente 10 a aproximadamente 16 aminoácidos, y un tercer factor de diversidad para los miembros de la región H-CDR3 que tengan una longitud de aproximadamente 17 a aproximadamente 22 aminoácidos.

45

50

El operario experto apreciará que el número de factores de diversidad puede totalizar el número de posiciones de aminoácidos en un miembro de la región H-CDR3 que tenga una longitud definida. Preferiblemente, se utilizan hasta 19 factores de diversidad, y los conjuntos de regiones H-CDR3 diversificadas de la invención pueden combinarse para formar diversos conjuntos de regiones H-CDR3 diversificadas, sea solas o como parte de un dominio de cadena pesada variable o una cadena de inmunoglobulina de longitud total o longitud sustancialmente total.

55

**EJEMPLOS**

Los ejemplos no limitantes siguientes definen más particularmente diversas realizaciones de la invención, especificando los parámetros para ciertos factores de diversidad que pueden aplicarse a regiones H-CDR3 que tengan una longitud dentro de un intervalo particular de aminoácidos (aa). El operario experto apreciará que se permite cierta desviación en la diversidad (frecuencia de aminoácidos) con respecto a los diseños teóricos mencionados más adelante, sin desviarse del alcance de la invención, dado que podría existir, por ejemplo, un deseo de implementar ciertas eficiencias de clonación en la construcción de una biblioteca basada, al menos en parte, en tales diseños teóricos.

5 permite cierta desviación en la diversidad (frecuencia de aminoácidos) con respecto a los diseños teóricos mencionados más adelante, sin desviarse del alcance de la invención, dado que podría existir, por ejemplo, un deseo de implementar ciertas eficiencias de clonación en la construcción de una biblioteca basada, al menos en parte, en tales diseños teóricos.

Recuperación y edición de secuencias para preparación de los diseños H-CDR3:

10 Las series de datos para análisis se obtuvieron de Ig-BLAST en el Servidor NCBI. Se descargó el archivo siguiente, que contiene registros de datos de Ig-BLAST: <ftp://ftp.ncbi.nih.gov/blast/db/FASTA/igSeqProt.gz>. Se descargaron en total 40.808 registros de datos, que eran humanos, de ratón, o no estaban asignados en el identificador de nombres de las entradas de datos en el formato FASTA. Por filtración de los datos con la palabra clave "humano" en el identificador de nombres en MS-Word, se extrajeron 22.500 entradas de anticuerpos humanos. De las 22.500  
15 entradas de anticuerpos humanos, 13.235 estaban asignadas a cadenas pesadas y 5490 a cadenas ligeras. El resto de los anticuerpos se asignaron ulteriormente a hc, lc-kappa e lc-lambda, dado que no se encontraba palabra clave alguna en el identificador de nombres. Para las cadenas pesadas, se recopiló un conjunto H-CDR3 dependiente de la longitud utilizando diferentes hojas de cálculo para diferentes longitudes en MS-Excel. En el análisis de H-CDR3, pudieron recogerse para análisis 8.886 de 13.235 secuencias para la cadena pesada.

20 Diseño de H-CDR3 dependiente de la longitud:

Como ejemplo, se generaron los tres diseños separados siguientes de acuerdo con tres factores de diversidad:

- 4aa-7aa
- 8aa-14aa
- 15aa-23aa

25 El contenido de Y en la posición Kabat 102 muestra una disminución a lo largo de la longitud 8aa-23aa de 50% a 5%, mientras que el contenido de V en esta posición aumenta desde 12% a 65%. Además, el contenido de D en la posición 101 aumenta desde 75% a 100%. Para implementar los hallazgos de los diferentes aa a lo largo de la longitud, pero sin limitar el alcance de la presente invención, se decidió que son preferibles 2 factores de diversidad diferentes bajo un diseño basado en bucle: uno de 8-14aa y otro de 15-23aa, mientras que la aplicación del factor de  
30 diversidad a las regiones H-CDR3 que tienen 4-7aa de longitud puede estar de acuerdo con un diseño a la izquierda.

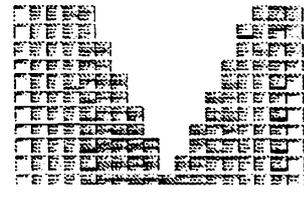
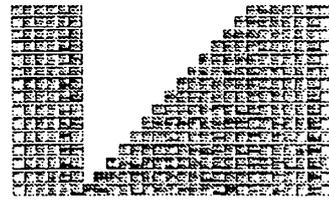
Tipo de alineación:

Para resumir un diseño de longitudes de varios aa, los tipos siguientes de alineaciones son alternativas representativas para el diseño de miembros de biblioteca que contienen 8-14aa y 15-23aa (tabla 1): alineación a la izquierda, alineación a la derecha y alineación simétrica.

**Alineación: izquierda**

**Alineación: derecha**

**Alineación: simétrica**



35 **Tabla 1: Tipos de alineación; las tres primeras posiciones y la última posición son posiciones constantes de aminoácidos de las regiones de entramado que flanquean la región H-CDR3.**

Cada paso de la "escalera" representa la secuencia de consenso de una cierta longitud. Las 3 primeras posiciones representan partes de las regiones de entramado vecinas: la región CAR en las posiciones Kabat 92, 93 y 94 y la W del comienzo de FW4 (WGQG) en la posición Kabat 103.

40 Alineación izquierda (Tabla 1 + Figura 1):

Las tres primeras posiciones (posiciones Kabat 96, 97 y 98) se alinearon como un bloque. Estas posiciones se encuentran en el borde entre el segmento V y el segmento D y podrían ser codificadas por V (= línea germinal). Así, estas posiciones podrían diferir en la composición del resto de la HCDR3 y se mantuvieron separadas. Las  
45 posiciones 96 y 97 difieren de otras posiciones de H-CDR3, dado que el contenido de Y se ha reducido, mientras que en la posición 96, el contenido de D se ha incrementado. Las cinco últimas posiciones (posiciones Kabat l, m, n,

101 y 102) con inclusión de las secuencias FDY dominantes, se alinearon como un bloque. Las posiciones 101 y 102 están codificadas a menudo en el segmento J y, por tanto, exhiben distribuciones de aminoácidos similares.

5 En primer lugar, después de llenar las tres primeras y las cinco últimas posiciones con aa, la alineación a la izquierda se inició comenzando con la posición Kabat 99 y llenando sucesivamente la laguna entre las posiciones Kabat 99 y "I".

Alineación a la derecha (Tabla 1 + Figura 2):

10 También en este caso, las posiciones Kabat 96, 97 y 98, así como las posiciones Kabat l, m, n, 101 y 102 se alinearon como un bloque, siguiendo las mismas consideraciones descritas para la alineación a la izquierda. En primer lugar, después de rellenar las tres primeras y las cinco últimas posiciones con aa, la alineación a la derecha se inició comenzando con la posición Kabat "I" y rellenando sucesivamente la laguna entre las posiciones Kabat "I" y 99.

Alineación centrada (Tabla 1 + Figura 3a, 3b, 3c):

15 En lo que concierne a los aspectos estructurales, una alineación simétrica es un tipo preferido de alineación y permite además una identificación rápida de los aa que interaccionan en el bucle. En este caso se añaden nuevos aminoácidos simétricamente a una cierta posición. La alineación simétrica se realizó centrada en la posición AHO 123 como se describe en "Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling an analysis tool, A. Honegger, A. Plückthun, J.Mol.Biol (2001) 309, 657-670.

20 Son posibles dos alineaciones simétricas: Una alineación, disponiendo una laguna en la posición 123, desplazando las secuencias en la región de entramado y otra alineación desplazando las secuencias hacia la posición 123, con la posición 123 como eje de simetría, dividiendo las secuencias en dos partes con el mismo número de aa. La primera alineación describe la distribución de aa de las regiones flanqueantes (Fig. 3a), mientras que la segunda alineación describe la punta del bucle de la región H-CDR3 (Fig. 3b), aunque las mismas secuencias pueden tomarse y disponerse de otro modo. Para el bucle, se introducen nuevos aminoácidos simétricamente en la base del bucle. En la alineación centrada mixta (Fig. 3c), las secuencias se orientaron simétricamente a la posición AHO 123, pero combinando las alineaciones para el bucle y para las regiones flanqueantes. Un bloque de 5aa, que rodeaba la posición 123, se mantenía constante, mientras que las otras secuencias se dispusieron en las regiones flanqueantes. Ejemplos de sistemas de numeración que pueden utilizarse para los dominios variables de anticuerpos se muestran en la Figura 4.

30 La Tabla 2 muestra el resultado de una alineación centrada (región flanqueante) para regiones H-CDR3 que tienen una longitud de 4-7aa. "Total" describe una distribución igual para todos los 19aa sin cisteína.

Pos. Kabat	95	96	97	m	n	101	102
Pos. AHO	109	110	111	112	113	137	138
Nombre	RECORTE 1	RECORTE 2	RECORTE 3		RECORTE 4	RECORTE 5	RECORTE 6
Diseño	<b>25% GD</b> <b>5% AEVLRS</b> 1,1% total	<b>10% RGSYA</b> 2,6% total	<b>25% G</b> <b>10% YS</b> <b>7.5% AD</b> <b>5% T</b> 2,3% total		<b>30% F</b> <b>15% L</b> <b>10% MG</b> 1,8% total	<b>65%D</b> <b>10%G</b> 1,3% total	<b>50% Y</b> <b>15% V</b> <b>10% I</b> 1,3% total
Distribución natural de aa	<b>33% G</b> <b>17% D</b> 6% A 6% E 6% V 5% L 5% R 5% S 2% H 2% I 2% N 2% P 2% T 2% W 2% Y 1% K 1% M 0% C 0% F 0% Q	<b>11% R</b> <b>10% G</b> <b>10% S</b> 8% Y 7% A 5% D 5% H 5% L 5% T 4% F 4% I 4% N 4% P 4% V 4% W 3% E 2% K 2% Q 1% M 0% C	<b>25% G</b> <b>11% Y</b> 9% S 7% D 6% A 6% T 5% N 5% R 5% W 4% L 3% P 3% V 2% E 2% F 2% Q 1% H 1% I 1% K 1% M 0% C	<b>27% G</b> <b>16% A</b> 7% S 7% Y 6% V 5% D 5% E 4% T 3% F 3% N 3% P 3% R 3% W 2% H 2% M 2% Q 1% I 1% K 1% L 0% C	<b>31% F</b> <b>16% L</b> <b>11% M</b> 9% G 4% I 4% P 4% S 3% A 3% Y 2% D 2% E 2% R 2% T 2% V 2% W 1% C 1% H 1% K 1% N 0% Q	<b>63% D</b> 7% G 5% S 4% A 3% R 2% E 2% L 2% N 2% V 1% H 1% K 1% P 1% T 0% C 0% F 0% I 0% M 0% Q 0% W	<b>47% Y</b> <b>14% V</b> 9% I 4% H 4% P 3% D 3% F 3% L 3% S 2% G 2% R 1% A 1% N 1% T 0% C 0% E 0% K 0% M 0% Q 0% W

Tabla 2: Alineación centrada para regiones H-CDR3 con longitudes de 4-7aa

La Tabla 3 muestra el resultado de una alineación centrada (región de bucle) para regiones de H-CDR3 que tienen una longitud de 4-7aa. "Total" describe una distribución igual para todos los 19aa sin cisteína.

**Resultado de todas las secuencias**

**Regiones de bucle, H-CDR-3, 4aa-7aa**

Pos. Kabat	95	96	97	m	n	101	102
Pos. AHO	109	110	111	112	113	137	138
Nombre	RECORTE 1	RECORTE 2	RECORTE 3		RECORTE 4	RECORTE 5	RECORTE 6
Diseño	<b>40%G</b> <b>15%D</b> <b>5%AEHSV</b> 1,1% total	<b>10% RSGYA</b> <b>5% HLTV</b> 1,6% total	<b>20%G</b> <b>10%A</b> <b>7.5%YSFDL</b>		<b>25% D</b> <b>20% F</b> <b>10% L</b> <b>5%GYM</b> 1,6% total	<b>40% D</b> <b>20% Y</b> <b>10% V</b> 1,6 total	<b>50% Y</b> <b>15% V</b> <b>10% I</b> 1,3% total
Distribución natural de aa	<b>40% G</b> <b>14% D</b> 5% E 5% A 5% L 5% S 5% V 4% R 2% H 2% I 2% N 2% P 2% T 2% W 2% Y 1% K 0% C 0% F 0% M 0% Q	<b>12% R</b> <b>11% D</b> <b>10% G</b> <b>10% S</b> 7% A 7% Y 5% H 5% L 5% T 5% V 4% I 4% N 3% E 3% F 2% K 2% W 1% M 1% Q 0% C	<b>22% G</b> <b>10% Y</b> 7% A 7% S 6% D 6% T 6% W 5% L 5% N 5% R 4% F 4% P 3% V 2% E 2% M 2% Q 1% H 1% I 1% K 0% C	<b>21% G</b> <b>11% F</b> <b>10% A</b> 7% D 7% L 6% M 6% S 5% Y 4% E 4% P 4% V 3% T 3% W 2% H 2% I 2% N 2% R 1% K 1% Q 0% C	<b>26% D</b> <b>18% F</b> <b>10% L</b> 6% G 6% Y 5% M 4% A 4% I 4% R 4% S 4% V 2% P 1% C 1% E 1% H 1% K 1% N 1% Q 1% T 1% W	<b>41% D</b> <b>21% Y</b> 8% V 4% G 4% P 3% F 3% H 3% I 3% S 2% L 2% N 2% T 1% A 1% R 0% C 0% E 0% K 0% M 0% Q 0% W	<b>51% Y</b> <b>15% V</b> <b>12% I</b> 4% D 3% P 3% S 2% H 2% L 1% A 1% F 1% G 1% N 1% R 1% T 0% C 0% E 0% K 0% M 0% Q 0% W

**Tabla 3: Alineación centrada de regiones H-CDR3 con longitudes de 4-7aa.**

La Tabla 4 muestra el resultado de una alineación a la izquierda para regiones H-CDR3 que tienen una longitud de 4-7aa. "Total" describe una distribución igual para todos los 19aa sin cisteína.

Pos. Kabat	95	96	97	m	n	101	102
Pos. AHO	109	110	111	112	113	137	138
Nombre	RECORTE 1	RECORTE 2	RECORTE 3	RECORTE 4	RECORTE 5	RECORTE 6	RECORTE 7
Diseño	<b>25% GD</b> <b>5% AEVLR</b> 1,1% total	<b>10% RGSY</b> <b>5% ADHIT</b> 1,8% total	<b>25% G</b> <b>10% YSDR</b> 1,8% total	<b>25% G</b> <b>10% YSA</b> 2,4% total	<b>30% F</b> <b>15% L</b> <b>10% MG</b> 1,8% total	<b>65% D</b> <b>10% G</b> 1,3% total	<b>50% Y</b> <b>15% V</b> <b>10% I</b> 1,3% total
Distribución natural de aa	<b>33% G</b> <b>17% D</b> 6% A 6% E 6% V 5% L 5% R 5% S 2% H 2% I 2% N 2% P 2% T 2% W 2% Y 1% K 1% M 0% C 0% F 0% Q	<b>12% R</b> <b>12% S</b> <b>10% G</b> 8% Y 6% A 6% D 6% H 5% I 4% F 4% L 4% N 4% V 3% E 3% P 3% W 3% W 2% K 2% Q 1% M 0% C	<b>23% G</b> <b>11% Y</b> 9% D 8% S 7% R 6% A 6% W 5% N 4% L 4% T 3% E 3% F 3% P 3% V 2% H 2% Q 1% I 1% K 0% C 0% M	<b>26% G</b> <b>13% A</b> 8% Y 7% S 6% T 4% D 4% N 4% P 4% V 4% W 3% E 3% F 3% R 2% H 2% L 2% M 2% Q 1% I 1% K 0% C	<b>29% F</b> <b>16% L</b> <b>10% M</b> 9% G 4% I 4% P 4% Y 3% A 3% S 3% V 3% W 2% D 2% E 2% R 2% T 1% C 1% H 1% K 1% N 0% Q	<b>63% D</b> 7% G 5% S 4% A 3% R 2% E 2% L 2% N 2% V 2% Y 1% H 1% K 1% P 1% T 0% C 0% F 0% I 0% M 0% Q 0% W	<b>47% Y</b> <b>14% V</b> 9% I 4% H 4% P 3% D 3% F 3% L 3% S 2% G 2% R 1% A 1% N 1% T 0% C 0% E 0% K 0% M 0% Q 0% W

Tabla 4: Alineación a la izquierda de regiones H-CDR3 con longitudes de 4-7aa.

5 La Tabla 5 muestra el resultado de una alineación a la derecha para regiones H-CDR3 que tienen una longitud de 4-7aa. "Total" describe una distribución igual para todos los 19aa sin cisteína.

Pos. Kabat	95	96	97	m	n	101	102
Pos. AHO	109	110	111	112	113	137	138
Nombre	RECORTE 1	RECORTE 2	RECORTE 3		RECORTE 4	RECORTE 5	RECORTE 6
Diseño	<b>30%G</b> <b>20%D</b> <b>5%AEVLR</b> 1,1% total	<b>10% RSGYA</b> <b>5% DHIT</b> 1,6% total	<b>25%G</b> <b>10%YS</b> <b>7.5% DRA</b> 1,8% total		<b>30% F</b> <b>15% L</b> <b>10% MG</b> 1,8% total	<b>65%D</b> <b>10%G</b> 1,3% total	<b>50% Y</b> <b>15% V</b> <b>10% I</b> 1,3% total
Distribución natural de aa	<b>33% G</b> <b>17% D</b> 6% A 6% E 6% V 5% L 5% R 5% S 2% H 2% I 2% N 2% P 2% T 2% W 2% Y 1% K 1% M 0% C 0% F 0% Q	<b>12% R</b> <b>12% S</b> <b>10% G</b> 8% Y 7% A 6% D 6% H 5% I 5% T 4% F 4% L 4% N 4% V 3% E 3% P 3% W 2% K 2% Q 1% M 0% C	<b>23% G</b> <b>11% Y</b> 9% D 8% S 7% R 6% A 6% W 5% N 4% L 4% T 3% E 3% F 3% P 3% V 2% H 2% Q 1% I 1% K 0% C 0% M	<b>26% G</b> <b>13% A</b> 8% Y 7% S 6% T 4% D 4% N 4% P 4% V 3% E 3% F 3% R 2% H 2% L 2% M 2% Q 1% I 1% K 0% C	<b>29% F</b> <b>16% L</b> <b>10% M</b> 9% G 4% I 4% P 4% Y 3% A 3% S 3% V 3% W 2% D 2% E 2% R 2% T 1% C 1% H 1% K 1% N 0% Q	<b>63% D</b> 7% G 5% S 4% A 3% R 2% E 2% L 2% N 2% V 1% H 1% K 1% P 1% T 0% C 0% F 0% I 0% M 0% Q 0% W	<b>47% Y</b> <b>14% V</b> 9% I 4% H 4% P 3% D 3% F 3% L 3% S 2% G 2% R 1% A 1% N 1% T 0% C 0% E 0% K 0% M 0% Q 0% W

Tabla 5: Alineación a la derecha de regiones H-CDR3 con longitudes de 4-7aa.

5 La Tabla 6 muestra el resultado de una alineación centrada (región flanqueante) para regiones H-CDR3 que tienen una longitud de 8-14aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 7 muestra el resultado de una alineación centrada (región de bucle) para regiones H-CDR3 que tienen una longitud de 8-14aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 8 muestra el resultado de una alineación centrada (centrada-mixta) para regiones H-CDR3 que tienen una longitud de 8-14aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 9 muestra el resultado de una alineación a la izquierda para regiones H-CDR3 que tienen una longitud de 8-14aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 10 muestra el resultado de una alineación a la derecha para regiones H-CDR3 que tienen una longitud de 8-14aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 11 muestra el resultado de una alineación centrada (región flanqueante) para regiones H-CDR3 que tienen una longitud de 15-23aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 12 muestra el resultado de una alineación centrada (región de bucle) para regiones H-CDR3 que tienen una longitud de 15-23aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 13 muestra el resultado de una alineación centrada (centrada-mixta) para regiones H-CDR3 que tienen una longitud de 15-23aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 14 muestra el resultado de una alineación a la izquierda para regiones H-CDR3 que tienen una longitud de 15-23aa. "Total" describe una distribución igual de todos los 19aa sin cisteína. La Tabla 15 muestra el resultado de una alineación a la derecha para regiones H-CDR3 que tienen una longitud de 15-23aa. "Total" describe una distribución igual de todos los 19aa sin cisteína.

Pos. Kabat	95	96	97	98	99	100	a	b	c	d	m	n	101	102				
Pos. AHO	109	110	111	112	113	114	115	116	117	118	119	120	137	138				
Nombre	RECORTÉ 1	RECORTÉ 2	RECORTÉ 3					RECORTÉ 4					REC. 5	REC. 6	RECORTÉ 7			
Diseno	20% DG 8% VEAS 1.5% total	10% G 10% RSLP 2.6% total	15% GYS 5% DRP VALIT 0.8% total						20% Y 15% G 10% S 2.9% total						20% YA 15% G 10% W 5% P 2.6% total	50% F 10% ML 1.6% total	80% D 20% A	45% Y 15% V 10% P 3% FILNS
Distribución natural de aa	22% D 19% G 9% E 7% V 6% A 6% R 6% S 4% L 3% H 3% T 2% I 2% N 2% P 2% Q 1% C 1% F 1% K 1% M 1% W 1% Y	14% G 12% R 9% S 8% L 8% P 6% T 5% A 5% V 5% Y 4% D 4% I 3% E 3% F 3% N 3% Q 2% H 2% K 2% W 1% C 1% M	17% G 13% S 9% Y 7% A 7% D 6% T 6% V 5% D 5% L 5% R 5% P 4% T 4% W 3% E 3% F 3% N 3% F 3% N 2% F 2% H 2% Q 1% C 1% K 1% M 1% Q	18% G 15% S 10% Y 7% A 7% T 6% V 5% D 4% L 4% R 3% E 3% I 3% N 3% W 2% F 2% F 2% P 1% C 1% H 1% K 1% M 1% Q	17% G 15% S 10% Y 8% A 7% T 6% D 5% L 5% R 5% V 4% P 4% W 3% N 2% E 2% F 2% F 2% I 1% H 1% K 1% M 0% C 0% Q	16% G 15% S 12% Y 8% T 6% A 6% V 5% D 5% W 4% L 4% R 3% F 3% N 3% P 2% E 2% E 2% H 2% I 1% C 1% K 1% M 1% Q	15% G 14% Y 12% S 7% T 6% A 6% V 5% L 5% R 4% D 4% N 4% P 4% W 3% N 2% C 2% E 2% F 2% H 2% I 2% I 2% Q 1% C 1% K 1% M 1% Q	17% Y 13% G 11% S 6% A 6% L 6% T 5% D 5% R 4% N 4% P 4% V 4% W 3% F 2% E 2% H 2% V 2% I 2% Q 1% C 1% K 1% M 1% H	20% Y 15% G 11% D 8% S 7% N 5% T 5% W 4% A 4% R 3% L 3% P 2% E 2% F 2% H 2% V 1% C 1% I 1% K 1% Q 0% M	20% Y 15% G 11% D 8% S 7% N 5% T 5% W 4% A 4% R 3% L 3% P 2% E 2% F 2% H 2% V 1% C 1% I 1% K 1% Q 0% M	21% Y 19% A 15% G 7% W 6% P 6% S 3% D 3% H 3% L 3% R 3% T 2% E 2% F 2% H 2% V 1% A 1% H 1% T 1% W 0% C 0% E 0% G 0% K 0% M 0% R 0% Q 0% R 0% T 0% W	55% F 11% M 10% L 3% I 3% Y 2% G 2% P 2% S 2% V 1% A 1% S 1% H 0% C 0% F 0% I 0% C 0% M 0% D 0% E 0% K 0% M 0% R 0% T 0% Q 0% R 0% T 0% W	86% D 2% A 2% E 2% G 1% H 1% L 1% N 1% P 1% Q 1% S 0% C 0% F 0% I 0% K 0% M 0% R 0% T 0% V 0% W 0% Y	45% Y 16% I 13% V 8% P 4% L 3% F 3% H 3% S 1% D 1% N 0% A 0% C 0% E 0% G 0% K 0% M 0% Q 0% R 0% T 0% W				

Tabla 6: Alineación centrada de regiones H-CDR3 con longitudes de 8-14 aa.

Pos. Kebab	95	96	97	98	99	100	a	b	c	d	m	n	101	102
Pos. AHO	109	110	111	112	113	114	115	116	117	118	119	120	137	138
Nombre	RECORTÉ 1	RECORTÉ 2	RECORTÉ 3	RECORTÉ 4				REC. 5				REC. 5	RECORTÉ 7	
Diseño	20%GD 9%AEVS 0.70%	20%GD 10%REV 5%SALP 0.5% total	15%GS 10%DRV 5%ALPEITY 0.26% total	15%GSY 7.5%DRALTY 0.5% total	19%G 14%S 8%Y 7%A 7%T 6%D 6%L 5%L 5%V 6%R 4%W 4%I 4%P 3%E 3%I 3%Q 2%F 2%N 2%Q 2%Y 1%C 1%K 1%W 1%Y	16%G 13%S 9%Y 7%R 6%A 6%D 6%L 6%T 6%V 5%L 5%R 4%N 4%W 3%E 3%I 3%N 3%F 3%N 3%W 2%H 2%Q 2%Q 1%H 1%K 1%M 1%M	18%G 15%S 11%Y 8%A 6%D 6%T 6%V 5%L 5%R 5%R 4%W 4%V 3%F 3%N 3%P 2%E 2%F 2%H 2%K 2%N 2%Q 2%Y 1%C 1%K 1%M 1%Q	18%G 14%Y 12%S 7%A 6%D 6%T 5%L 5%R 4%N 4%V 4%W 3%P 2%E 2%F 2%I 2%Q 1%C 1%K 1%M 1%Q	20%Y 15%G 10%D 7.5%ASN 1.3% total	15%DYF 10%AG 5%WLP 1.1% total	22%D 22%F 13%Y 7%A 7%G 4%L 4%M 3%P 3%W 2%H 2%S 2%N 2%I 2%S 2%V 1%E 1%N 1%R 1%T 0%C 0%K 0%Q 0%K 0%Q	35%D 25%F 15%Y 5%IML 0.52% total	40%D 25%Y 10%IV 5%PLF 0.52% total	45%Y 15%VI 10%P 3%FHLNS
Distribución natural de aa	25%D 16%G 9%A 9%E 6%V 5%R 4%H 4%L 4%L 4%S 3%P 3%T 2%I 2%M 2%N 1%C 1%F 1%K 1%Q 1%W 1%Y	19%G 18%D 9%R 7%E 7%V 6%S 5%A 5%L 5%L 4%P 4%T 3%H 3%I 2%F 2%N 2%Q 2%Y 1%C 1%K 1%M 1%W	15%G 10%D 10%R 8%S 7%V 6%A 6%L 6%P 5%E 4%I 4%T 3%Y 3%Q 2%F 2%H 2%K 2%N 2%Q 2%Y 1%C 1%K 1%W 1%M	17%G 10%S 9%D 8%R 7%Y 6%A 6%L 6%T 6%V 5%L 5%R 4%I 4%E 3%I 2%F 2%H 2%K 2%N 2%Q 2%W 2%W 1%C 1%M	19%G 14%S 8%Y 7%A 7%T 6%D 6%L 5%L 5%V 4%R 4%W 3%E 3%I 3%N 3%P 3%N 2%F 2%Q 2%Q 1%H 1%K 1%M 0%C	16%G 13%S 9%Y 7%R 6%A 6%D 6%L 6%T 6%V 5%L 5%R 4%N 4%W 3%E 3%I 3%N 3%F 3%N 3%W 2%H 2%Q 2%Q 1%H 1%K 1%M 1%M	18%G 15%S 11%Y 8%A 6%D 6%T 6%V 5%L 5%R 5%R 4%W 4%V 3%F 3%N 3%P 2%E 2%F 2%H 2%K 2%N 2%Q 2%Y 1%C 1%K 1%M 1%Q	18%G 14%Y 12%S 7%A 6%D 6%T 5%L 5%R 4%N 4%V 4%W 3%P 2%E 2%F 2%I 2%Q 1%C 1%K 1%M 1%Q	18%Y 13%G 9%A 8%S 7%D 6%F 5%L 5%T 5%W 4%N 4%P 3%N 3%T 2%E 2%H 2%I 2%H 2%M 2%I 1%C 1%K 1%M 1%Q	17%Y 16%F 11%G 10%D 8%A 5%W 4%L 4%P 4%S 3%N 3%T 2%E 2%H 2%I 2%M 2%R 2%V 1%C 1%Q 0%K 0%Q	22%D 22%F 13%Y 7%A 7%G 4%L 4%M 3%P 3%W 2%H 2%S 2%N 2%I 2%S 2%V 1%E 1%N 1%R 1%T 0%C 0%K 0%Q 0%K 0%Q	33%D 23%F 14%Y 6%I 5%M 4%L 3%V 2%P 2%S 1%A 1%E 1%G 1%N 1%Q 0%A 0%C 0%K 0%M 0%Q 0%R 0%T 0%W	43%D 24%Y 7%I 7%V 5%P 3%L 2%F 2%H 2%S 1%E 1%G 1%N 1%Q 0%A 0%C 0%K 0%M 0%Q 0%R 0%T 0%W	39%Y 17%V 16%I 9%P 5%L 3%H 3%S 2%F 2%D 1%N 0%A 0%C 0%G 0%K 0%M 0%Q 0%R 0%T 0%W

Tabla 7: Alineación centrada de regiones H-CDR3 con longitudes de 8-14 aa.

Pos. Kabet.	95	96	97	98	99	100	a	b	c	d	m	n	101	102
Pos. AHO	109	110	111	112	113	114	115	116	117	118	119	120	137	138
Nombre	RECORTÉ 1	RECORTÉ 2	RECORTÉ 3	RECORTÉ 4	RECORTÉ 3		RECORTÉ 4		RECORTE 5		REC. 6	REC. 7	REC. 8	REC. 9
Diseño	20% DG 8% VEAS 11.5% total	10% G 10% RSLP 2.6% total	15% GYS 10% AR 5% DP VLIT 0.8 % total	15% GSY 7.5% DAPVT 0.9% total					20% Y 15% G 10% D 7.5% ASN 1.3% total		20% YA 15% G 10% W 5% P 2.6% total	50% F 10% ML 1.6% total	80% D 20% A	45% Y 15% VI 10% P 3% FHLNS
Distribución natural de aa	22% D 19% G 9% E 7% V 6% A 6% R 6% S 4% L 3% H 3% T 2% I 2% N 2% P 2% Q 1% C 1% F 1% K 1% M 1% W 1% Y	14% G 13% R 9% S 8% L 8% P 6% T 5% A 5% V 5% Y 4% D 4% I 3% E 3% K 3% Q 2% F 2% H 2% N 2% W 1% C 1% M	16% G 11% S 8% R 8% Y 7% V 6% A 6% D 5% L 5% P 5% T 4% E 4% I 3% E 3% F 3% W 2% C 2% H 2% N 2% Q 1% K 1% M 1% M	14% G 13% S 9% D 9% Y 8% A 6% P 6% V 5% L 5% T 4% E 4% I 4% R 3% K 2% C 2% F 2% N 2% W 1% H 1% M 1% Q	16% G 13% S 9% Y 7% R 6% A 6% D 6% L 6% T 6% V 5% L 4% R 4% W 3% E 3% I 3% N 3% P 3% N 2% F 2% H 2% Q 1% H 1% K 1% M 1% M	19% G 14% S 8% Y 7% A 7% T 6% D 5% L 5% V 4% R 4% W 3% E 3% I 3% N 3% P 3% N 2% F 2% H 2% Q 1% H 1% K 1% M 0% C	18% G 15% S 11% Y 8% A 6% D 6% T 6% V 5% L 5% R 4% N 4% W 3% F 3% N 3% P 2% E 2% H 2% I 1% C 1% K 1% M 1% Q	16% G 14% Y 12% S 7% A 6% D 6% T 5% L 5% R 4% V 4% W 4% W 3% P 2% E 2% F 2% I 2% Q 1% C 1% H 1% K 1% M 1% M	18% Y 13% G 9% A 8% S 7% D 6% F 5% L 5% T 5% W 4% N 4% P 3% R 3% V 2% E 2% H 2% I 1% C 1% K 1% M 1% Q	25% Y 14% G 12% D 7% N 6% S 5% W 4% R 4% T 3% L 3% P 2% A 2% C 2% E 2% F 2% H 2% I 2% Q 1% K 1% M 0% M	23% Y 18% A 15% G 8% W 6% P 4% S 3% H 3% T 2% D 2% E 2% F 2% L 2% N 2% R 2% V 1% C 1% I 0% K 0% M 0% Q	60% F 11% M 9% L 3% I 3% Y 2% G 2% P 2% S 2% V 1% A 1% T 1% W 0% C 0% I 0% E 0% G 0% K 0% M 0% R 0% T 0% V 0% W 0% R	86% D 2% A 2% E 2% G 1% H 1% L 1% N 1% P 1% Q 1% S 0% C 0% F 0% I 0% E 0% G 0% K 0% M 0% R 0% T 0% V 0% W 0% Y	45% Y 16% I 13% V 8% P 4% L 3% F 3% H 3% S 1% D 1% N 0% A 0% C 0% E 0% G 0% K 0% M 0% R 0% T 0% V 0% W

Tabla 8: Alineación centrada de regiones H-CDR3 con longitudes de 8-14 aa.

Pos. Kabat	95	96	97	98	99	100	a	b	c	d	m	n	101	102				
Pos. AHO	109	110	111	112	113	114	115	116	117	118	119	120	137	138				
Nombre	REC. 1	REC. 2	RECORTIE 3					RECORTIE 4					REC. 5	REC. 6	REC. 7			
Diseño	20% DG 8% VEAS 1.5% total	10% G 10% RSLP 2.6% total	15% GYS 5% DRP VALIT 0.8 % total						20% Y 15% G 10% S 2.9% total						20% YA 15% G 10% W 5% P 2.6% total	50% F 10% ML 1.6% total	80% D 20% A 3% FHLS	45% Y 15% VI 10% P 3% FHLS
Distribución natural de aa	22% D 19% G 9% E 7% V 6% A 6% R 6% S 4% L 3% H 3% T 2% I 2% N 2% P 2% Q 1% C 1% F 1% K 1% M 1% W 1% Y	14% G 12% R 9% S 8% L 8% P 6% T 5% A 5% V 5% Y 4% D 4% I 4% I 3% E 3% F 3% N 3% Q 3% Q 2% H 2% K 2% W 1% C 1% M	17% G 12% S 8% Y 7% R 6% A 6% V 5% D 5% L 5% P 5% T 4% I 4% W 3% E 3% N 3% F 2% F 2% H 2% Q 1% C 1% K 1% M 1% M	17% G 13% S 9% Y 7% A 7% D 6% T 6% V 5% L 5% R 4% P 3% E 3% I 3% N 3% W 2% F 2% P 2% H 2% Q 2% Q 1% C 1% K 1% M 1% M	18% G 15% S 10% Y 7% A 7% T 6% V 5% D 4% L 4% R 3% E 3% I 3% N 3% W 2% F 2% P 1% C 1% H 1% K 1% M 1% M	17% G 14% S 11% Y 7% A 7% T 6% D 6% L 5% V 4% P 4% R 3% E 3% I 3% N 2% E 2% F 2% I 1% H 1% K 1% M 1% Q 0% C	14% G 13% S 13% Y 8% T 6% A 6% L 6% V 5% D 5% R 4% P 4% W 4% W 3% F 3% N 2% E 2% I 2% Q 1% C 1% H 1% K 1% M 1% Q 0% C	17% Y 13% Q 12% S 6% A 6% T 5% L 4% D 4% N 4% P 4% R 4% V 4% W 4% W 3% F 3% F 3% N 2% C 2% E 2% H 2% I 2% K 2% K 2% Q 1% M	22% Y 13% G 10% S 6% A 6% L 5% P 5% T 4% D 4% R 4% V 4% W 3% F 3% I 3% N 2% E 2% V 2% H 1% C 1% I 1% K 1% M 1% Q	20% Y 17% G 11% D 8% S 7% N 5% W 4% A 4% R 4% T 3% L 2% E 2% F 2% H 2% P 2% V 1% A 1% H 0% C 0% F 0% I 0% K 0% M 1% C 1% I 1% K 1% Q 0% K 0% M	21% Y 19% A 15% G 7% W 6% P 6% S 3% D 3% H 3% L 3% R 3% T 2% E 2% F 2% H 2% P 2% V 1% A 1% H 0% C 0% F 0% I 0% K 0% M 1% C 1% I 1% K 1% Q 0% K 0% M	59% F 11% M 10% L 3% I 3% Y 2% G 2% P 2% S 2% V 1% A 1% H 0% C 0% F 0% I 0% K 0% M 0% E 0% G 0% K 0% M 0% R 0% T 0% V 0% W 0% Y	86% D 2% A 2% E 2% G 1% H 1% L 1% N 1% P 1% Q 1% S 1% N 0% C 0% F 0% I 0% K 0% M 0% R 0% T 0% V 0% W 0% Y	45% Y 16% I 13% V 8% P 4% L 3% F 3% H 3% S 1% D 1% N 0% A 0% C 0% E 0% G 0% K 0% M 0% Q 0% R 0% T 0% V 0% W 0% Y				

Tabla 9: Alineación izquierda de regiones H-CDR3 con longitudes de 8-14 aa.

Pos. Kabat	95	96	97	98	99	100	a	b	c	d	m	n	101	102
Pos. AHO	109	110	111	112	113	114	115	116	117	118	REC. 5	REC. 6	137	138
Nombre	RECORTES 1-6													
Diseño	RECORTES 1-6													
	20% DG 8% VEAS 1.5% total	10% G 10% RSLP 2.6% total	15% GYS 5% DRP VALIT 0.8 % total						15% GS 7.5% DATRV 1.7% total		20% YA 15% G 10% W 5% P 2.6% total	50% F 10% ML 1.6% total	60% D 20% A	45% Y 15% VI 10% P 3% FHLNS
Distribución natural de aa	22% D 19% G 9% E 7% V 6% A 6% R 6% S 4% L 3% H 3% T 2% I 2% N 2% P 1% C 1% F 1% K 1% M 1% W 1% Y	14% G 12% R 9% S 8% L 8% P 6% T 5% A 5% V 5% Y 4% D 4% I 3% E 3% F 3% N 3% Q 2% H 2% K 2% W 1% C 1% M	16% G 11% S 9% R 7% V 7% Y 6% A 6% D 6% P 5% L 5% V 5% Y 4% D 4% I 4% E 3% F 2% C 2% H 2% M 2% N 2% W 1% K 1% Q	14% G 11% S 9% Y 7% D 6% A 6% L 6% P 6% R 6% T 6% V 5% L 5% R 4% W 4% E 4% I 3% F 2% C 2% H 2% M 2% N 2% W 1% K 1% M	16% G 13% S 11% Y 7% A 6% R 6% T 6% V 5% D 5% L 5% L 4% I 4% P 4% E 3% F 2% C 2% H 2% N 2% Q 2% W 1% K 1% M	15% G 15% S 8% Y 7% A 7% D 7% T 6% V 5% D 5% L 4% R 4% W 3% E 3% I 3% N 3% P 2% F 2% H 2% Q 1% H 1% K 1% M 0% C	19% G 14% S 10% Y 7% A 6% T 6% Y 5% D 5% L 4% R 4% W 3% F 3% I 3% N 2% E 2% F 2% E 2% H 1% C 1% H 1% K 1% M 1% Q	17% G 14% S 11% Y 7% A 7% T 6% V 5% D 5% R 4% L 4% P 4% W 3% N 3% I 2% E 2% F 2% I 1% C 1% H 1% K 1% M 1% Q	15% G 15% Y 12% S 7% A 6% L 6% T 5% D 5% R 5% V 4% N 4% P 4% W 3% F 3% I 3% N 2% E 2% F 2% E 2% H 2% I 2% Q 1% C 1% K 1% M	20% Y 17% G 11% D 8% S 7% N 5% W 4% A 4% R 4% T 3% L 3% E 2% F 2% H 2% P 2% Y 1% C 1% I 1% K 1% Q 0% M	21% Y 19% A 15% G 7% W 6% P 6% S 3% D 3% H 3% L 3% R 3% T 2% E 2% F 2% N 2% V 1% C 1% E 1% W 0% C 0% D 0% M 0% E 0% R 0% T 0% N 0% Q 0% M	59% F 11% M 10% L 3% I 3% Y 2% G 2% P 2% S 2% V 1% A 1% H 1% T 1% W 0% C 0% D 0% K 0% M 0% E 0% R 0% T 0% N 0% Q 0% R	66% D 2% A 2% E 2% G 1% H 1% L 1% N 1% P 1% Q 1% S 0% C 0% F 0% I 0% K 0% M 0% R 0% T 0% V 0% W 0% Y	45% Y 15% I 13% V 8% P 4% L 3% F 3% H 3% S 1% D 1% N 0% A 0% C 0% E 0% G 0% K 0% M 0% Q 0% R 0% T 0% W 0% Y

Tabla 10: Alineación derecha de regiones H-CDR3 con longitudes de 8-14 aa.



Pos. Kabat	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100					
Pos. AHO	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148					
Nombre	REC.1		REC.2		REC.3			REC.4		REC.5		REC.6			REC.7		REC.8		REC.9			REC.10			REC.11																										
Diseño	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	
	0.8% total	1.3% total	1.3% total	1.1% total	1.3% total	1.6% total	0.8% total	1.3% total	1.6% total																																										
	34R0	34R1	34R2	34R3	34R4	34R5	34R6	34R7	34R8	34R9	34R0	34R1	34R2	34R3	34R4	34R5	34R6	34R7	34R8	34R9	34R0	34R1	34R2	34R3	34R4	34R5	34R6	34R7	34R8	34R9	34R0	34R1	34R2	34R3	34R4	34R5	34R6	34R7	34R8	34R9	34R0	34R1	34R2	34R3	34R4	34R5	34R6	34R7	34R8	34R9	
	34R0	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G
	34R1	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	
	34R2	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G		
	34R3	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G			
	34R4	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G				
	34R5	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G					
	34R6	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G						
	34R7	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G							
	34R8	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G								
	34R9	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G	15R0G	15R1G	15R2G	15R3G	15R4G	15R5G	15R6G	15R7G	15R8G	15R9G									

Distribución natural de aa

Tabla 12: Alineación centrada de regiones H-CDR3 con longitudes de 15-23 aa.

Pos. Kabal	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92																																													
	REC.1	REC.2	RECORTE 3			RECORTE 4			RECORTE 5			RECORTE 6			RECORTE 7			REC.8	REC.9	REC.10	REC.11																																																														
Nombre	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92
Diseño	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92
	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92
	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92

Distribución natural de aa

Tabla 13: Alineación centrada de regiones H-CDR3 con longitudes de 15-23 aa.

Pos. Kebab	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52																																	
Pos. AHO	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467																																	
Nombre	REC. 1			REC. 2			REC. 3			REC. 4			REC. 5			REC. 6			REC. 7			REC. 8			REC. 9			REC. 10																							
Diseño	20% DG 15% AS 15% VS 1.5% total			15% AS 15% VS 1.5% total																																															
	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500
Distribución natural de aa	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500

Tabla 14: Alineación izquierda de regiones H-CDR3 con longitudes de 15-23 aa.

Pos. Kabat	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112			
Pos. AHO	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117			
Nombre	REC. 1	REC. 2	REC. 3	REC. 4	RECORTE 5				RECORTE 6				RECORTE 7	RECORTE 8	REC. 9	REC. 10	REC. 11	Bambol- REC. 12			
Diseño	10% DS 10% VS 10% LS 1.5% total	10% G US 10% GP 10% BS 2.6% total	10% GS 10% BF 10% SF 2.1% total	10% GP 10% BS 10% SF 2.1% total	10% GS 10% BF 10% SF 2.1% total																
Distribución natural de aa	10% D 10% G 10% A 10% E 10% V 10% S 10% L 10% R 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C	10% G 10% R 10% N 10% W 10% S 10% A 10% Y 10% D 10% H 10% I 10% P 10% C

Tabla 15: Alineación derecha de regiones H-CDR3 con longitudes de 15-23 aa.

En otro ejemplo no limitante, la presente descripción proporciona un conjunto de diversas regiones H-CDR3 de longitud variable, en las cuales la diversidad de las regiones H-CDR3 puede generarse por diversificación de regiones H-CDR3 que tienen una longitud dentro de un primer intervalo de aminoácidos de aproximadamente 4 a aproximadamente 8 aminoácidos de acuerdo con un primer factor de diversidad, donde el primer factor de diversidad requiere que: alanina tenga una tasa de frecuencia de aproximadamente 80% en la posición 1, arginina tenga una tasa de frecuencia de aproximadamente 60% en la posición 2, glicina tenga una tasa de frecuencia de aproximadamente 40% en la posición 3, ácido aspártico tenga una tasa de frecuencia de aproximadamente 50% en la posición 7 y tirosina tenga una tasa de frecuencia de aproximadamente 40% en la posición 8 (Tabla 16).

5

Tabla 16: La tabla muestra un ejemplo para un factor de diversidad para diversificación de las regiones H-CDR3 que tienen una longitud dentro de un primer intervalo de aminoácidos de aproximadamente 4 a aproximadamente 8 aminoácidos.

10

Posición de aa HCDR3	1	2	3	4	5	6	7	8
	A 80	R 60	G 40	19 sin C	19 sin C	19 sin C	D 50	Y 40
	T 10	18 sin Cys	18 sin Cys				G 10	V 20
	17 sin Cys						A10	17 sin Cys
							16 sin C	

En otro ejemplo no limitante, la presente descripción proporciona un conjunto de diversas regiones H-CDR3 de anticuerpos humanos o humanizados de longitud variable, en las cuales la diversidad de las regiones H-CDR3 puede generarse por diversificación de regiones H-CDR3 que tienen una longitud comprendida dentro de un primer intervalo de aminoácidos de aproximadamente 9 a aproximadamente 15 aminoácidos de acuerdo con un primer factor de diversidad, donde el primer factor de diversidad requiere que: alanina tenga una tasa de frecuencia de aproximadamente 90% en la posición 1, arginina tenga una tasa de frecuencia de aproximadamente 70% en la posición 2, tirosina tenga una tasa de frecuencia de aproximadamente 20% en la posición 10, tirosina tenga una tasa de frecuencia de aproximadamente 20% en la posición 11, tirosina tenga una tasa de frecuencia de aproximadamente 30% en la posición 12, fenilalanina tenga una tasa de frecuencia de aproximadamente 60% en la posición 13, ácido aspártico tenga una tasa de frecuencia de aproximadamente 80% en la posición 14 y tirosina tenga una tasa de frecuencia de aproximadamente 40% en la posición 15. Preferiblemente, la diversidad de una región H-CDR3 humana o humanizada que tiene una longitud de aminoácidos de aproximadamente 9 a aproximadamente 15 se genera por un contenido elevado de glicina y serina dentro de la región H-CDR3 (Tabla 17).

15

20

25

Tabla 17: La tabla muestra un ejemplo para un factor de diversidad para diversificación de regiones H-CDR3 que tienen una longitud dentro de un primer intervalo de aminoácidos de aproximadamente 9 a aproximadamente 15 aminoácidos.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A80	R70	O20	R15	O20	O20	O20	O20	O20	Y20	Y20	Y20	F80	F80	Y40
T5	F20	O21	G15	S16	S10	S10	S10	S10	G10	G35	G36	L10	12 sin C	V30
V5	1/7 sin C	17 sin C	17 sin C	17 sin C	V10	V10	V10	V10	S10	S10	S10	17/ sin C		P10
					16 sin C	18 sin C	18 sin C	16 sin C	19 sin C	19 sin C	19 sin C			18 sin C

30

En otra realización adicional no limitante, la presente descripción proporciona un conjunto de diversas regiones H-CDR3 de anticuerpos humanos o humanizados de longitud variable, en donde la diversidad de las regiones H-CDR3 puede generarse por diversificación de regiones H-CDR3 que tienen una longitud de aproximadamente 16 a aproximadamente 22 aminoácidos de acuerdo con un primer factor de diversidad, donde el primer factor de diversidad requiere que: Alanina tenga una tasa de frecuencia de aproximadamente 90% en la posición 1, Arginina tenga una tasa de frecuencia de aproximadamente 60% en la posición 2, Ácido aspártico tenga una tasa de frecuencia de aproximadamente 30% en la posición 3, Glicina tenga una tasa de frecuencia de aproximadamente 20% en la posición 4, Arginina tenga una tasa de frecuencia de aproximadamente 10% en la posición 5, Arginina tenga una tasa de frecuencia de aproximadamente 10% en la posición 6, Tirosina tenga una tasa de frecuencia de aproximadamente 20% en la posición 7, Tirosina tenga una tasa de frecuencia de aproximadamente 40% en la posición 15, Tirosina tenga una tasa de frecuencia de aproximadamente 50% en la posición 16, Tirosina tenga una tasa de frecuencia de aproximadamente 50% en la posición 17, Tirosina tenga una tasa de frecuencia de aproximadamente 60% en la posición 18, Tirosina tenga una tasa de frecuencia de aproximadamente 40% en la

35

40

posición 19, Metionina tenga una tasa de frecuencia de aproximadamente 50% en la posición 20, Ácido aspártico tenga una tasa de frecuencia de aproximadamente 95% en la posición 21 y Valina tenga una tasa de frecuencia de aproximadamente 60% en la posición 22 (tabla 18).

Posición de aa	1	2	3	4	5	6	7	8	9	10	11	12
HCDR3	A 90	R 80	D 30	G 20	R 10	R 10	Y 20	S 20	R 10	G 20	S 20	S 20
	T 5	K 20	G 20	P 10	P 10	Y 10	G 10	D 10	D 10	G 10	G 10	Y 10
	V 5	T 20	17 sin C	R 10	G 10	G 10	F 10	I 10	G 10	17 sin C	V 20	G 10
				F 10	F 10	I 10	M 10	Y 10	V 10		18 sin C	V 10
				15 sin C		15 sin C						

	13	14	15	16	17	18	19	20	21	22
	E 10	Y 20	Y 40	Y 50	Y 50	Y 60	Y 40	M 50	D 95	V 60
	S 10	T 10	G 10	G 10	18 sin C	18 sin C	G 40	F 40	Q 5	16 sin C
	G 10	17 sin C	17 sin C	17 sin C			17 sin C	17 sin C		
	16 sin C									

5 **Tabla 18: La tabla muestra un ejemplo para un factor de diversidad para diversificación de regiones H-CDR3 que tienen una longitud comprendida dentro de un primer intervalo de aminoácidos de aproximadamente 16 a aproximadamente 22 aminoácidos.**

Adicionalmente, la diversidad de un conjunto de regiones H-CDR3 humanas o humanizadas que tienen una longitud de aminoácidos de aproximadamente 16 a aproximadamente 22 aminoácidos puede generarse por un contenido elevado de serina en la región H-CDR3, un contenido elevado de aminoácidos básicos en la parte final (C-terminal) de la región H-CDR3, ácido aspártico en la parte frontal (N-terminal) de la región H-CDR3 y un contenido elevado de aminoácidos aromáticos en la región H-CDR3 total.

En el contexto de la presente invención, se entiende que en cualquier posición de aminoácido dada en una región H-CDR3, la existencia de aminoácidos no especificados con una tasa de frecuencia definida en una posición de aminoácido definida puede presentarse en ratios sustancialmente iguales unos con respecto a otros, o algunos aminoácidos pueden tener una mayor tasa de frecuencia comparados con otros aminoácidos en dicha posición y, en algunos casos, ciertos aminoácidos pueden omitirse en una o más posiciones de aminoácidos definidas (v.g. cisteína).

En el contexto de la presente invención, el diseño de diversas regiones H-CDR3 puede hacerse por mimetización de la distribución natural conocida de aminoácidos por sesgo de la distribución aleatoria completa de aminoácidos en la secuencia de DNA codificante de H-CDR3. La generación de diversas regiones H-CDR3 puede hacerse por utilización de la tecnología de mutagénesis de trinucleótidos (TRIM) (Virnekäs et al., 1994), que permite sintetizar a voluntad cualquier mezcla de aminoácidos deseada en cada posición individual, introduciendo con ello cualquier sesgo de aminoácidos en cualquier posición. Generalmente, la tecnología TRIM está basada en la utilización de trinucleótidos pre-ensamblados como entrada para la síntesis de DNA estándar, evitando con ello desplazamientos de marco, codones de parada, o aminoácidos no deseados, y produce una diversidad de secuencias enorme. Por tanto, se requieren tamaños de biblioteca muy grandes para generar un espacio suficientemente amplio de secuencias funcionales. Correspondientemente, el conocimiento de la composición de aminoácidos dependiente de la longitud en combinación con la tecnología TRIM permitiría un diseño de un conjunto mucho menor de regiones H-CDR3 con un contenido funcional mucho mayor.

Adicionalmente, la invención proporciona un conjunto de diversos anticuerpos humanos o fragmentos funcionales de los mismos, cada uno de los cuales contiene una cadena pesada variable de anticuerpo humano de acuerdo con la invención. En el contexto de la presente invención, debe entenderse que para la generación de un conjunto de anticuerpos humanos diversos o fragmentos de los mismos, el profesional experto podrá modificar el diseño teórico de un conjunto de diversas regiones H-CDR3. El operario experto apreciará que puede ser aceptable modificar el diseño con respecto a los aminoácidos que existen a una tasa de frecuencia baja (v.g. aa de una frecuencia inferior a 5% o preferiblemente inferior a 3%), por adición de aquellos aminoácidos que aparecen por debajo de dicha tasa de frecuencia en concentraciones equimolares unos frente a otros, por ejemplo; de este modo pueden generarse colecciones de diversos anticuerpos humanos o humanizados o fragmentos funcionales de los mismos por inclusión de mezclas de tales aminoácidos a dichas concentraciones.

Se describe también en esta memoria una biblioteca sintética de anticuerpos humanos que comprende un conjunto de diversos anticuerpos humanos o fragmentos funcionales de los mismos. Las secuencias de consenso sintéticas pueden cubrir el repertorio estructural de anticuerpos codificados en el genoma humano. Una biblioteca sintética de anticuerpos humanos puede estar basada en anticuerpos o fragmentos de anticuerpos, preferiblemente fragmentos Fv, Fv enlazado por disulfuro, Fv monocatenario (scFv), o Fab, los cuales pueden utilizarse como fuentes de especificidades contra nuevos antígenos diana.

En el contexto de la presente invención, una biblioteca sintética de anticuerpos humanos o humanizados puede estar basada en secuencias de aminoácidos que han sido diseñados en simulación por computadora y codificados

por ácidos nucleicos que se crean sintéticamente. El diseño de simulación por computadora de una secuencia de anticuerpos se realiza, por ejemplo, por análisis de una base de datos de secuencias humanas e ideación de una secuencia de polipéptidos utilizando los datos obtenidos a partir de ello. Por análisis de las secuencias de aminoácidos conocidas de la línea germinal de dominios variables de inmunoglobulina humana, se agrupan los genes de anticuerpos en sub-familias de acuerdo con la homología de secuencia y con estructuras canónicas (Chothia, et al., 1992; Tomlinson et al., 1995). Debido a la alta homología en términos de secuencia de aminoácidos y estructuras canónicas dentro de una sub-familia, pueden diseñarse 7 secuencias de consenso VH y 7 VL (genes maestros) y combinarse para producir 49 combinaciones de entramado maestras (Knappik et al., 2000). Los genes maestros podrían sintetizarse por tanto químicamente generando una estructura génica completamente modular por introducción de sitios de restricción únicos flanqueantes de todos los módulos funcionales.

Métodos para el diseño y la obtención en secuencias coloreadas por simulación en computadora se describen, por ejemplo, en Knappik et al., J. Mol. Biol. (2000) 296:57; Krebs et al., J. Immunol. Methods. (2001) 254:67; y en la Patente U.S. No. 6.300.064 otorgada a Knappik et al.

Se describen también en esta memoria métodos de preparación de un conjunto de moléculas de ácido nucleico que codifican diversas regiones H-CDR3 humanas. Esto puede hacerse, por ejemplo, por síntesis de una pluralidad de moléculas de DNA, donde cada una codifica una región H-CDR3, y donde las regiones H-CDR3 son de gamas variables.

En un aspecto, la presente invención proporciona un conjunto de ácidos nucleicos que codifican un conjunto de diversas regiones H-CDR3 de anticuerpos humanos de acuerdo con la presente invención.

Se describen también en esta memoria métodos de obtención de un conjunto de diversas regiones H-CDR3 de anticuerpos humanos o humanizados. Esto puede hacerse por expresión del conjunto de ácidos nucleicos que codifican un conjunto de diversas regiones H-CDR3 de anticuerpos humanos o humanizados.

Se describen también métodos de identificación de uno o más genes que codifican una o más regiones H-CDR3 de la invención que se fijan a una diana dada. Esto se realiza, v.g., por expresión de las regiones H-CDR3, seguido por cribado de las mismas para aislar una o más regiones H-CDR3 que se fijan a una molécula diana dada. Una región H-CDR3 puede estar comprendida dentro de una cadena pesada variable de anticuerpo. El cribado puede realizarse utilizando uno de los métodos bien conocidos por los expertos en la técnica, tales como presentación de fago o CysDisplay (descrito, v.g. en la patente US No. 6.753.136).

En el contexto de la presente descripción, una biblioteca de anticuerpos humanos puede estar basada en la combinación de fragmentos génicos de donantes humanos con DNA sintético diseñado. Como ejemplo, fragmentos procedentes de donantes humanos que codifican la región variable de la cadena ligera y una gama de regiones de cadena pesada se combinan con DNA sintético que codifica secuencias de anticuerpos humanos con diversidad introducida en sitios específicos.

Se describen también anticuerpos mutados u optimizados (o fragmentos de los mismos) derivados de un anticuerpo o fragmento del mismo, como se describe en esta memoria. Un anticuerpo está compuesto de dos cadenas peptídicas, cada una de las cuales contiene uno (cadena ligera) o tres (cadena pesada) dominios constantes y una región variable (VL, VH), la última de las cuales está constituida en cada caso por cuatro regiones FR y tres CDRs intercaladas. El sitio de fijación de antígeno está formado por una o más CDRs; en cambio, las regiones FR proporcionan el entramado estructural para las CDRs y pueden jugar también un papel importante en la fijación de antígeno. Por alteración de uno o más residuos de aminoácido en una región CDR o FR, el operario experto puede generar rutinariamente secuencias de anticuerpos mutadas o diversificadas, que pueden cribarse contra el antígeno, respecto a propiedades nuevas o mejoradas, por ejemplo.

Asimismo, como se utiliza en esta memoria, una "inmunoglobulina" (Ig) se define por tanto como una proteína perteneciente a la clase IgG, IgM, IgE, IgA, o IgD (o cualquier subclase de las mismas), e incluye todos los anticuerpos convencionalmente conocidos y fragmentos funcionales de los mismos. Un "fragmento funcional" de un anticuerpo/inmunoglobulina se define en esta memoria como un fragmento de un anticuerpo/inmunoglobulina (v.g., una región variable de una IgG) que conserva la región de fijación de antígeno. Una "región de fijación de antígeno" de un anticuerpo se encuentra típicamente en una o más regiones hipervariables de un anticuerpo, es decir, las regiones CDR-1, -2, y/o -3; sin embargo, las regiones variables "de entramado" pueden jugar también un papel importante en la fijación de antígeno, por ejemplo por proporcionar un armazón para las CDRs. Preferiblemente, la "región de fijación de antígeno" comprende al menos los residuos de aminoácidos 4 a 103 de la cadena ligera variable (VL) y 5 a 109 de la cadena variable pesada (VH), más preferiblemente los residuos de aminoácidos 3 a 107 de VL y 4 a 111 de VH, prefiriéndose particularmente las cadenas VL y VH completas (posiciones de aminoácidos 1 a 109 de VL y 1 a 113 de VH; numeración de acuerdo con WO 97/08320). Una clase preferida de inmunoglobulinas para uso en la presente invención es IgG. "Fragmentos funcionales" de la invención incluyen el dominio de un fragmento F(ab')<sub>2</sub>, un fragmento Fab y scFv. El F(ab')<sub>2</sub> o Fab puede modificarse por ingeniería genética para minimizar o eliminar por completo las interacciones intermoleculares disulfuro que existen entre los dominios CH1 y CL.

El término "región hipervariable", cuando se utiliza en esta memoria, hace referencia a los residuos de aminoácidos de un anticuerpo que son responsables de la fijación de antígeno. La región hipervariable comprende residuos de aminoácidos de una "región determinante de la complementariedad" o "CDR" (es decir los residuos 24-34 (L1), 50-56 (L2) y 89-97 (L3) en el dominio variable de la cadena ligera y 31-35 (H1), 50-65 (H2) y 95-102 (H3) en el dominio variable de la cadena pesada; Kabat et al., Sequences of Proteins of Immunological Interest, 5ª edición, Public Health Service, National Institutes of Health, Bethesda, Md. (1991)) y/o aquellos residuos de un "bucle hipervariable" (es decir los residuos 26-32 (L1); 50-52 (L2) y 91-96 (L3) en el dominio variable de la cadena ligera y 26-32 (H1), 53-55 (H2) y 96-101 (H3) en el dominio variable de la cadena pesada; Chothia y Lesk J. Mol. Biol. 196; 901-917 (1987)). Los residuos "de entramado" o "FR" son aquellos residuos de dominio variable distintos de los residuos de la región hipervariable, como se definen en esta memoria.

Un anticuerpo "humano" o fragmento de anticuerpo humano funcional se define en esta memoria como uno que no es quimérico (v.g., no "humanizado") y no procede (totalmente o en parte) de una especie no humana. Un anticuerpo o fragmento de anticuerpo funcional humano puede derivarse de un humano o puede ser un anticuerpo sintético humano. Un "anticuerpo sintético humano" se define en esta memoria como un anticuerpo que tiene una secuencia derivada, totalmente o en parte, por simulación en computadora a partir de secuencias sintéticas que están basadas en el análisis de secuencias conocidas de anticuerpos humanos. El diseño por simulación en computadora de una secuencia de anticuerpos humanos o fragmentos de la misma puede realizarse por ejemplo, por análisis de una base de datos de anticuerpos o secuencias de fragmentos de anticuerpos humanos e ideación de una secuencia de polipéptidos que utilice los datos obtenidos de este modo. Otro ejemplo de un anticuerpo o fragmento de anticuerpo funcional humano es uno que está codificado por un ácido nucleico aislado de una biblioteca de secuencias de anticuerpos de origen humano (es decir, estando basada dicha biblioteca en anticuerpos tomados de una fuente humana natural).

Un "anticuerpo humanizado" o fragmento de anticuerpo humanizado funcional se define en esta memoria como uno que (i) se deriva al menos en parte de una fuente no humana (v.g., un ratón transgénico que lleva un sistema inmune heterólogo), anticuerpo que está basado en una secuencia de línea germinal humana; o (ii) es quimérico, en donde al menos una porción del dominio variable se deriva de un origen no humano y el dominio constante se deriva de un origen humano, o (iii) está injertado en CDR, en donde al menos una porción de las CDRs del dominio variable son de origen no humano, mientras que uno o más entramados del dominio variable son de origen humano y el dominio constante (en su caso) es de origen humano.

30

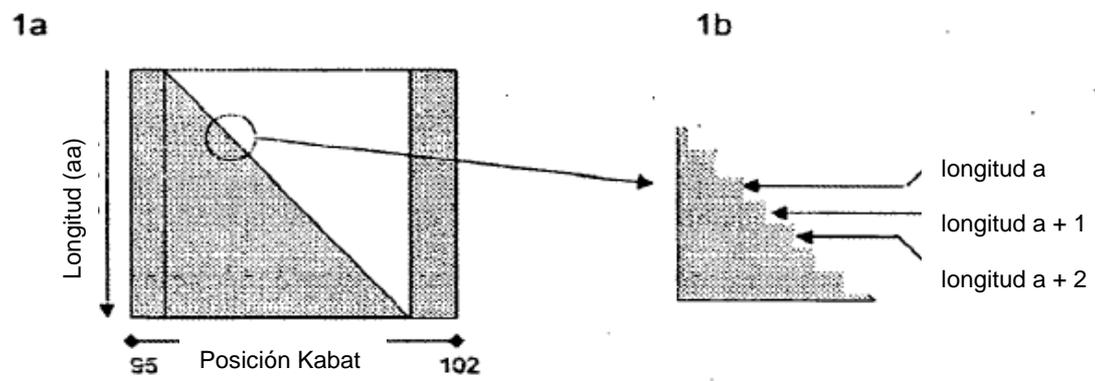
**REIVINDICACIONES**

1. Una conjunto de regiones H-CDR3 diversas de anticuerpos humanos de intervalos variables de aminoácidos, en donde la diversidad del conjunto se genera por: diversificación de regiones H-CDR3 que tienen una longitud comprendida dentro de un primer intervalo de aminoácidos de acuerdo con un primer factor de diversidad, 5 diversificación de regiones H-CDR3 que tienen una longitud dentro de un segundo intervalo de aminoácidos de acuerdo con un segundo factor de diversidad, y diversificación de regiones H-CDR3 que tienen una longitud dentro de un tercer intervalo de aminoácidos de acuerdo con un tercer factor de diversidad, en donde dichos factores de diversidad son diferentes,
- 10 en donde dicha primer intervalo de aminoácidos es de 4 a 7 aminoácidos, en donde dicho segundo intervalo de aminoácidos es de 8 a 14 aminoácidos, y en donde dicho tercer intervalo de aminoácidos es de 15 a 23 aminoácidos, y
- en donde dicho primer intervalo de aminoácidos tiene una composición como se representa en la Tabla 4, dicho segundo intervalo de aminoácidos tiene una composición como se representa en la Tabla 9, y dicho tercer intervalo de aminoácidos tiene una composición como se representa en la Tabla 14.
- 15 2. Una conjunto de diversos de anticuerpos humanos o fragmentos funcionales de los mismos, que comprende un conjunto de cadenas pesadas variables de anticuerpo humano que comprenden un conjunto de diversas regiones HCDR3 de anticuerpos humanos de acuerdo con la reivindicación 1.
3. Una conjunto de ácidos nucleicos que codifica un conjunto de diversas regiones H-CDR3 de anticuerpos humanos de acuerdo con la reivindicación 1.

20

**Fig. 1a + 1b**

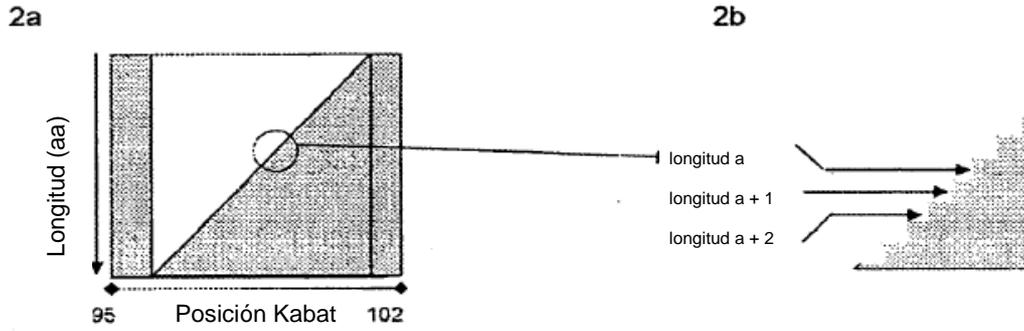
Alineación a la izquierda



Gris: secuencias  
Blanco: Holgura

Fig. 2a + 2b

Alineación a la derecha

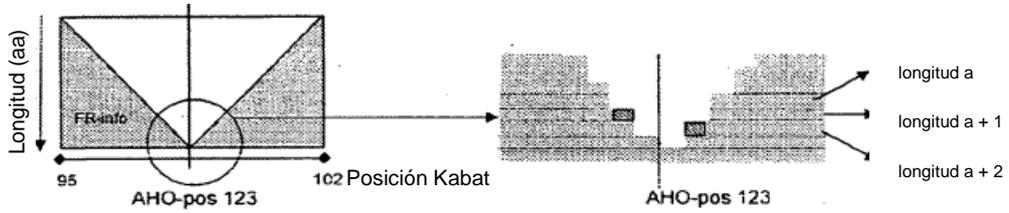


Gris: secuencias  
Blanco: Holgura

Fig. 3a + 3b + 3c

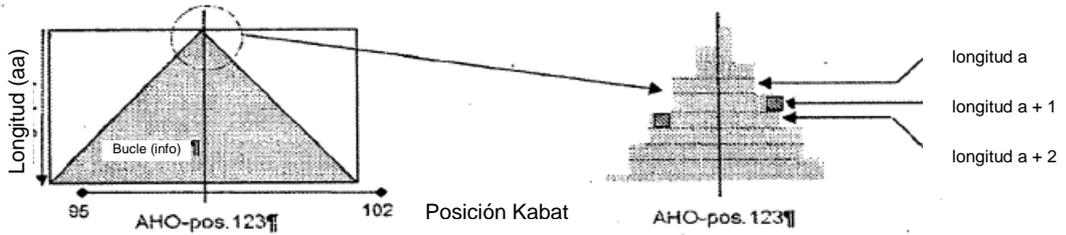
3a

Alineación - Región flanqueante (FR)



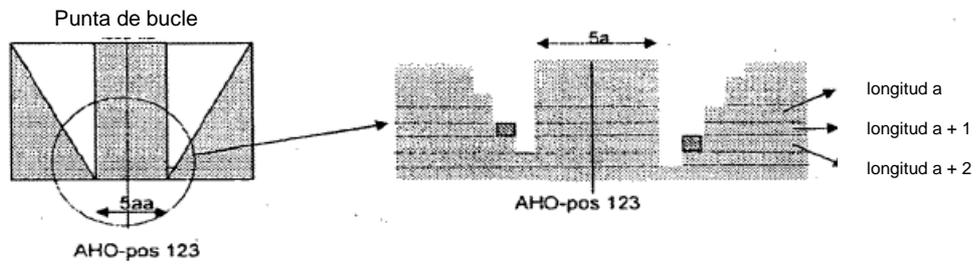
3b

Alineación - Región de bucle (LR)



3c

Alineación centrada mixta



Gris: secuencias  
Blanco: Holgura

Fig. 4

