



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.1

NORMA ST.25

NORMA PARA LA PRESENTACIÓN DE LISTAS DE SECUENCIAS DE NUCLEÓTIDOS Y AMINOÁCIDOS EN SOLICITUDES DE PATENTE

Revisión adoptada por el Grupo de Trabajo del SCIT sobre Normas y Documentación en su undécima sesión el 30 de octubre de 2009

Se recomienda a las Oficinas que apliquen, *mutatis mutandis*, las disposiciones de la “Norma para la presentación de listas de secuencias de nucleótidos y aminoácidos en solicitudes de patente en virtud del Tratado de Cooperación en materia de Patentes (PCT)” en la forma establecida en el Anexo C de las Instrucciones Administrativas en virtud del PCT, a todas las solicitudes de patente distintas de las solicitudes internacionales en virtud del PCT, observando que ciertas disposiciones específicas a los procedimientos y exigencias del PCT podrán no ser aplicables a solicitudes de patente que no sean solicitudes internacionales en virtud del PCT^(*). El texto de esa Norma del PCT se reproduce en las páginas siguientes.

^(*) Si al 1 de julio de 2009 la legislación y la práctica nacionales aplicables por una Oficina no son compatibles con las disposiciones del párrafo 3.i) de la “Norma para la presentación de listas de secuencias de nucleótidos y aminoácidos en solicitudes internacionales de patente en virtud del Tratado de Cooperación en materia de Patentes (PCT)”, que establecen que, cuando una lista de secuencias figure en la solicitud internacional tal como ésta fue presentada, “deberá presentarse como una parte distinta de la descripción, figurar al final de la solicitud, llevar por título, preferiblemente, ‘lista de secuencias’, comenzar en una nueva página y tener paginación independiente”, dicha Oficina podrá no adherirse a esas disposiciones hasta tanto no cese dicha incompatibilidad.



ANEXO C

NORMA PARA LA PRESENTACIÓN DE LISTAS DE SECUENCIAS DE NUCLEÓTIDOS Y AMINOÁCIDOS EN SOLICITUDES INTERNACIONALES DE PATENTE EN VIRTUD DEL PCT

INTRODUCCIÓN

1. Esta Norma ha sido elaborada para normalizar la presentación de listas de secuencias de nucleótidos y aminoácidos en solicitudes internacionales de patente. El objetivo de la Norma es permitir que el solicitante establezca una única lista de secuencias que sea aceptable para todas las Oficinas receptoras, las Administraciones encargadas de la búsqueda internacional y las Administraciones encargadas del examen preliminar internacional a los efectos de la fase internacional, y para todas las Oficinas designadas y elegidas a los fines de la fase nacional. Está destinada a mejorar la precisión y calidad de la presentación de las secuencias de nucleótidos y aminoácidos que figuran en solicitudes internacionales, permitir una presentación y divulgación de las secuencias que sea más fácil para los solicitantes, el público y los examinadores, facilitar la búsqueda de datos en las secuencias y permitir el intercambio electrónico de datos de las secuencias y la introducción de esos datos en bases de datos informatizadas.

DEFINICIONES

2. A los efectos de la presente Norma:

i) se entenderá por “lista de secuencias” la lista de secuencias de nucleótidos o aminoácidos que ofrece una divulgación detallada de las secuencias de nucleótidos o aminoácidos así como otras informaciones disponibles;

i-bis) se entenderá por “lista de secuencias que forma parte de la solicitud internacional” la lista de secuencias que figure en la solicitud internacional tal como fue presentada (según se menciona en el párrafo 3) o la lista de secuencias incluida en la solicitud internacional por medio de una corrección conforme a la Regla 26, una rectificación conforme a la Regla 91 o una modificación conforme al artículo 34.2)b) de la descripción en relación con las secuencias que figuran en la solicitud internacional tal como fue presentada (según se menciona en el párrafo 3*bis*);

i-ter) se entenderá por “lista de secuencias que no forma parte de la solicitud internacional” la lista de secuencias que no forma parte de la solicitud internacional pero que se suministra a los efectos de la búsqueda internacional o del examen preliminar internacional (según se menciona en los párrafos 4 y 4*bis*);

ii) las secuencias que se incluyan serán las secuencias lineales de cuatro o más aminoácidos o las secuencias lineales de diez o más nucleótidos. Quedan específicamente excluidas de la presente definición las secuencias ramificadas, las secuencias con menos de cuatro aminoácidos o nucleótidos específicamente definidos, así como las secuencias que comprendan nucleótidos o aminoácidos distintos de los que figuran en los Cuadros 1, 2, 3 y 4 del Apéndice 2;

iii) por “nucleótidos” se entenderán únicamente los nucleótidos que puedan representarse utilizando los símbolos que figuran en el Cuadro 1 del Apéndice 2. Se podrán describir las modificaciones, por ejemplo, las bases metiladas, en la forma establecida en el Cuadro 2 del Apéndice 2, pero no figurarán explícitamente en la secuencia de nucleótidos;



iv) por “aminoácidos” se entenderán los L-aminoácidos que se encuentran comúnmente en las proteínas naturales y que se enumeran en el Cuadro 3 del Apéndice 2. Las secuencias de aminoácidos que contienen al menos un D-aminoácido no están comprendidas en la presente definición. Toda secuencia de aminoácidos que contenga aminoácidos modificados posteriormente a la traducción, puede describirse como la secuencia de aminoácidos que se traduce inicialmente utilizando los símbolos que figuran en el Cuadro 3 del Apéndice 2, con las posiciones modificadas, por ejemplo, las hidroxilaciones o las glicosilaciones, que se describen en el Cuadro 4 del Apéndice 2, pero esas modificaciones no aparecerán explícitamente en la secuencia de aminoácidos. Quedan comprendidos en esta definición los péptidos o proteínas que puedan expresarse en forma de secuencia utilizando los símbolos del Cuadro 3 del Apéndice 2, y que al mismo tiempo contenga una descripción, por ejemplo, de los enlaces anormales, los entrecruzamientos (por ejemplo, el puente disulfuro) y las cofias terminales, los enlaces no peptídicos, etc.;

v) por “identificador de secuencia” se entenderá un número entero único que corresponde al SEQ ID NO asignado a cada secuencia de la lista;

vi) por “identificador numérico” se entenderá un número de tres dígitos que representa a un elemento de dato específico;

vii) por “vocabulario lingüísticamente neutro” se entenderá un vocabulario controlado que se utiliza en la lista de secuencias para representar los términos científicos en la forma prescrita por los proveedores de bases de datos de secuencias (incluidos el nombre científico, los calificadores y sus valores en relación con el vocabulario controlado, los símbolos de los Cuadros 1, 2, 3 y 4 del Apéndice 2, y las claves de caracterización que figuran en los Cuadros 5 y 6 del Apéndice 2);

viii) por “administración competente” se entenderá la Administración encargada de la búsqueda internacional que efectúa la búsqueda internacional y emite la opinión escrita de la Administración encargada de la búsqueda internacional respecto de la solicitud internacional, o la Administración encargada del examen preliminar internacional que efectúa el examen preliminar internacional respecto de la solicitud internacional.

LISTAS DE SECUENCIAS

Lista de secuencias que forma parte de la solicitud internacional

3. La lista de secuencias que figura en la solicitud internacional tal como fue presentada:

i) deberá presentarse como una parte distinta de la descripción, figurar al final de la solicitud, llevar por título, preferiblemente, “Lista de secuencias”, comenzar en una nueva página y tener paginación independiente³⁴; preferiblemente, la lista de secuencias no deberá reproducirse en ninguna otra parte de la solicitud; a reserva de lo dispuesto en el párrafo 36, resulta innecesario describir las secuencias en ninguna otra parte de la descripción;

ii) deberá exponer las secuencias representadas en la lista de secuencias y otras informaciones disponibles en la lista de secuencias de conformidad con lo dispuesto en los párrafos 5 a 35;

iii) si figura en una solicitud internacional presentada en formato electrónico, deberá figurar en un formato de documento electrónico y presentarse por un medio de transmisión que se halle en conformidad con lo dispuesto en el párrafo 37.

3bis. Toda corrección conforme a la Regla 26, toda rectificación conforme a la Regla 91 o toda modificación conforme al artículo 34.2)b) de la descripción presentada en relación con las secuencias

³⁴ *Nota del editor:* No es necesaria la paginación independiente cuando la lista de secuencias figure en una solicitud internacional presentada en formato electrónico y en el formato de documento electrónico mencionado en el párrafo 40.



que figuran en una solicitud internacional presentada en formato electrónico deberá presentarse en forma de lista de secuencias en formato electrónico que comprenda la lista entera con la corrección, rectificación o modificación pertinente. Esa lista de secuencias:

i) deberá llevar por título, preferiblemente, “Lista de secuencias - Corrección”, “Lista de secuencias - Rectificación” o “Lista de secuencias - Modificación”, según sea el caso, y tener paginación independiente;

ii) deberá exponer las secuencias representadas en la lista de secuencias y otras informaciones disponibles en la lista de secuencias de conformidad con lo dispuesto en los párrafos 5 a 35; cuando proceda, deberá mantenerse la numeración original de las secuencias en la solicitud internacional tal como fue presentada (según se menciona en el párrafo 5); en caso contrario, las secuencias deberán numerarse de conformidad con lo dispuesto en el párrafo 5;

iii) deberá figurar en un formato de documento electrónico y presentarse por un medio de transmisión que se halle en conformidad con lo dispuesto en el párrafo 38.

Lista de secuencias que no forma parte de la solicitud internacional

4. La lista de secuencias proporcionada en virtud de la Regla 13^{ter} a los fines de la búsqueda internacional y del examen preliminar internacional:

i) deberá llevar por título, preferiblemente, “Lista de secuencias - Regla 13^{ter}”;

ii) deberá exponer las secuencias representadas en la lista de secuencias y otras informaciones disponibles en la lista de secuencias de conformidad con lo dispuesto en los párrafos 5 a 35; cuando proceda, deberá mantenerse la numeración original de las secuencias en la solicitud internacional tal como fue presentada (según se menciona en el párrafo 5); en caso contrario, las secuencias deberán numerarse de conformidad con lo dispuesto en el párrafo 5;

iii) si se presenta en papel de conformidad con la Regla 13^{ter}.1)b), deberá tener paginación independiente;

iv) si se presenta en formato electrónico, deberá figurar en un formato de documento electrónico y presentarse por un medio de transmisión que se halle en conformidad con lo dispuesto en el párrafo 39;

v) si se presenta en formato electrónico junto con la solicitud internacional, deberá ser idéntica a la lista de secuencias que figura en la solicitud e ir acompañada de una declaración en el sentido de que “la información registrada en formato electrónico presentada en virtud de la Regla 13^{ter} es idéntica a la lista de secuencias que figura en la solicitud internacional”;

vi) si se entrega posteriormente a la presentación de la solicitud internacional, no podrá exceder la divulgación contenida en la solicitud internacional tal como fuera presentada e irá acompañada de una declaración a tal efecto; esa lista de secuencias deberá contener únicamente las secuencias divulgadas en la solicitud internacional tal como fue presentada.

4bis. Toda corrección conforme a la Regla 26, toda rectificación conforme a la Regla 91 o toda modificación conforme al artículo 34.2)b) de la descripción presentada en relación con las secuencias que figuran en la solicitud internacional tal como fue presentada deberá ir acompañada, a los fines de la búsqueda internacional o del examen preliminar internacional, de una lista de secuencias en formato electrónico en un formato de documento electrónico que se halle en conformidad con lo dispuesto en el párrafo 39, que comprenda la lista entera, incluida toda corrección, rectificación o modificación, siempre que lo exija la autoridad competente, salvo que esa autoridad ya disponga de ese tipo de lista en formato electrónico de una forma y manera que sean aceptables para ella. Esa lista de secuencias en formato electrónico:



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.5

i) deberá llevar por título, preferiblemente, “Lista de secuencias – Corrección - Regla 13^{ter}”, “Lista de secuencias – Rectificación - Regla 13^{ter}” o “Lista de secuencias – Modificación - Regla 13^{ter}”, según sea el caso;

ii) deberá exponer las secuencias representadas en la lista de secuencias y otras informaciones disponibles en la lista de secuencias de conformidad con lo dispuesto en los párrafos 5 a 35; cuando proceda, deberá mantenerse la numeración original de las secuencias en la solicitud internacional tal como fue presentada (según se menciona en el párrafo 5); en caso contrario, las secuencias deberán numerarse de conformidad con lo dispuesto en el párrafo 5;

iii) deberá presentarse por un medio de transmisión que se halle en conformidad con lo dispuesto en el párrafo 39;

iv) deberá ser idéntica a la lista de secuencias presentada en virtud del párrafo 3^{bis} como una corrección conforme a la Regla 26, una rectificación conforme a la Regla 91 o una modificación conforme al artículo 34.2)b) de la descripción e ir acompañada de una declaración en el sentido de que “la información registrada en formato electrónico presentada en virtud de la Regla 13^{ter} es idéntica a la lista de secuencias suministrada como una corrección conforme a la Regla 26 (o una rectificación conforme a la Regla 91 o una modificación conforme al artículo 34.2)b), según sea el caso) de la descripción en relación con las secuencias que figuran en la solicitud internacional tal como fue presentada.”

Si la autoridad competente no dispone de esa lista de secuencias en formato electrónico y, cuando proceda, de esa declaración, dicha autoridad únicamente tendrá que tener en cuenta ese tipo de corrección, rectificación o modificación a los fines de la búsqueda internacional o del examen preliminar internacional en la medida en que se pueda efectuar una búsqueda significativa o un examen preliminar significativo sin disponer de esa lista de secuencias en formato electrónico.

PRESENTACIÓN DE SECUENCIAS

5. Se asignará a cada secuencia un identificador de secuencia distinto. Los identificadores de secuencia comenzarán con el 1 e irán aumentando en forma consecutiva por números enteros. Si a un identificador de secuencia no correspondiese una secuencia, se utilizará el código 000 en el identificador numérico <400>, comenzando en la línea siguiente al SEQ ID NO. La respuesta al identificador numérico <160> deberá incluir el número total de SEQ ID NO, con independencia de si van seguidos de una secuencia o del código 000.

6. En la descripción, las reivindicaciones o los dibujos de la solicitud, se deberá hacer referencia a las secuencias representadas en la lista de secuencias mediante el identificador de secuencia precedido de “SEQ ID NO:”.

7. Las secuencias de nucleótidos y de aminoácidos deben estar representadas por una de las tres posibilidades siguientes, como mínimo:

- i) una secuencia de nucleótidos pura;
- ii) una secuencia de aminoácidos pura;
- iii) una secuencia de nucleótidos y la correspondiente secuencia de aminoácidos.

En las secuencias divulgadas según el formato especificado en el punto iii), la secuencia de aminoácidos deberá divulgarse separadamente en la lista de secuencias como una secuencia pura de aminoácidos y con un identificador de secuencia distinto, compuesto por un número entero.



Secuencias de nucleótidos

Símbolos que deben utilizarse

8. Toda secuencia de nucleótidos deberá representarse únicamente por una cadena simple, en el sentido 5' a 3' y de izquierda a derecha. Los valores 3' y 5' no deberán representarse en la secuencia.
9. Se deberán representar las bases de una secuencia de nucleótidos utilizando el código de una letra correspondiente a las secuencias de este tipo. Se utilizarán únicamente letras minúsculas conforme a la lista que se proporciona en el Cuadro 1 del Apéndice 2.
10. Las bases modificadas se representarán mediante las bases no modificadas correspondientes contenidas en la propia secuencia (o mediante la letra "n") si la base modificada es una de las que figuran en el Cuadro 2 del Apéndice 2, y se describirá posteriormente la modificación en la sección de características de la lista de secuencias utilizando los códigos que figuran en el Cuadro 2 del Apéndice 2. Se podrán utilizar esos códigos en la descripción o en la sección de características de la lista de secuencias, pero no así en la propia secuencia (véase también el párrafo 32). El símbolo "n" equivale a un solo nucleótido desconocido o modificado.

Formato que debe utilizarse

11. Toda secuencia de nucleótidos estará compuesta por un máximo de 60 bases por línea, con un espacio entre cada grupo de 10 bases.
12. Las bases de una secuencia de nucleótidos (incluidos los intrones) figurarán en la lista en grupos de 10 bases, salvo en las regiones codificadoras de la secuencia. Las bases restantes, en número inferior a 10, situadas en los extremos de las regiones no codificadoras de una secuencia, deben estar agrupadas y separadas de los grupos adyacentes mediante un espacio.
13. Las bases de las regiones codificadoras de una secuencia de nucleótidos deben figurar como tripletes (codones).
14. La enumeración de los nucleótidos deberá comenzar en la primera base de la secuencia con el número 1. Deberá ser continua en toda la secuencia en el sentido 5' a 3'. Deberá figurar en el margen derecho, en la línea que contiene los códigos de una letra correspondientes a las bases, y deberá indicar el número de la última base de dicha línea. Este método de numeración de las secuencias de nucleótidos es aplicable a las secuencias de nucleótidos de configuración circular, con la salvedad de que la designación de la primera base de la secuencia de nucleótidos puede quedar a elección del solicitante.
15. Toda secuencia compuesta por uno o más segmentos no contiguos de una secuencia más grande o por segmentos de diferentes secuencias, deberá numerarse como una secuencia distinta y con un identificador de secuencia diferente. Una secuencia que presente uno o más espacios deberá numerarse como una serie de secuencias distintas y con identificadores diferentes, siendo el número de las secuencias distintas igual al número de series continuas de datos de secuencia.

Secuencias de aminoácidos

Símbolos que deben utilizarse

16. Los aminoácidos de una secuencia proteínica o peptídica deberán relacionarse en el sentido del grupo amino al grupo carboxilo, y de izquierda a derecha. Los grupos amino y carboxilo no deberán representarse en la secuencia.
17. Los aminoácidos se representarán empleando el código de tres letras, siendo la primera una letra mayúscula, con arreglo a la lista que figura en el Cuadro 3 del Apéndice 2. Toda secuencia de



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.7

aminoácidos que contenga un espacio en blanco o un símbolo interno de terminación (por ejemplo, “Ter” o “*” o “.”) no podrá representarse como una secuencia única de aminoácidos, sino que deberá presentarse como diferentes secuencias de aminoácidos (véase el párrafo 22).

18. Los aminoácidos modificados y poco comunes se representarán como los aminoácidos no modificados correspondientes (o como “Xaa”) en la propia secuencia, si el aminoácido modificado es uno de los que figura en la lista del Cuadro 4 del Apéndice 2, y se deberá describir con detalle la modificación en la sección de características de la lista de secuencias, utilizando los códigos del Cuadro 4 del Apéndice 2. Se podrán emplear estos códigos en la descripción o en la sección de características de la lista de secuencias pero no así en la propia secuencia (véase también el párrafo 32). El símbolo “Xaa” equivale a un solo aminoácido desconocido o modificado.

Formato que debe utilizarse

19. Toda secuencia proteínica o peptídica deberá tener un máximo de 16 aminoácidos por línea, dejando un espacio entre cada aminoácido.

20. Los aminoácidos correspondientes a los codones de las regiones codificadoras de una secuencia de nucleótidos deberán figurar inmediatamente debajo de los codones correspondientes. Cuando un codón esté interrumpido por un intrón, el símbolo del aminoácido figurará debajo de la porción del codón que contenga dos nucleótidos.

21. La numeración de los aminoácidos comenzará en el primer aminoácido de la secuencia, con el número 1. Alternativamente, los aminoácidos que precedan a la proteína madura, por ejemplo las presecuencias, las prosecuencias y las preprosecuencias, así como las secuencias señal, cuando existan, podrán tener números negativos contados en forma regresiva desde el aminoácido adyacente al número 1. No se empleará el cero (0) cuando la numeración de los aminoácidos emplee números negativos para distinguir la proteína madura. El número figurará debajo de la secuencia cada cinco aminoácidos. El método de numeración de las secuencias de aminoácidos descrito anteriormente es aplicable a las secuencias de aminoácidos de configuración circular, con la salvedad de que la designación del primer aminoácido de la secuencia puede quedar a elección del solicitante.

22. Toda secuencia de aminoácidos compuesta por uno o más segmentos no contiguos de una secuencia más grande o de segmentos de secuencias diferentes, deberá numerarse como una secuencia distinta y con un identificador de secuencia diferente. Una secuencia que presente uno o varios espacios se numerará como una serie de secuencias distintas con identificadores de secuencia diferentes, siendo el número de secuencias distintas igual al número de series continuas de datos de secuencia.

OTRAS INFORMACIONES DISPONIBLES EN LA LISTA DE SECUENCIAS

23. El orden de los elementos de información en las listas de secuencia se ajustará al orden en que esos elementos están relacionados en la lista de identificadores numéricos de elementos de datos, definidos en el Apéndice 1.

24. Sólo se emplearán los identificadores numéricos de los elementos de datos definidos en el Apéndice 1 para la presentación de los elementos de información en la lista de secuencias. No se emplearán las descripciones de los identificadores numéricos correspondientes. La información suministrada deberá figurar inmediatamente después del identificador numérico, y sólo deben aparecer en la lista de secuencias los identificadores numéricos respecto de los cuales se suministra información. Las dos excepciones a esta regla la constituyen los identificadores numéricos <220> y <300>, que sirven de encabezamiento a los elementos titulados “características” e “información sobre la publicación”, respectivamente, y que están asociados con la información contenida en los identificadores numéricos <221> a <223> y <301> a <313>, respectivamente. Cuando se suministra



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.8

información en la lista de secuencias bajo esos identificadores numéricos en relación con los elementos titulados “característica” e “información sobre la publicación”, deben incluirse los identificadores numéricos <220> y <300>, respectivamente, pero dejándolos en blanco. En general, debe insertarse una línea en blanco entre los identificadores numéricos cuando cambie el dígito en la primera o segunda posición del identificador numérico. La excepción a esta regla general la constituye el identificador numérico <310>, al que no precederá línea en blanco ninguna. Además, todo identificador numérico repetido deberá ir precedido por una línea en blanco.

Elementos de datos obligatorios

25. La lista de secuencias deberá incluir, además de, e inmediatamente antes de la secuencia de nucleótidos o de aminoácidos, los siguientes elementos de información definidos en el Apéndice 1 (elementos de datos obligatorios):

<110>	Nombre del solicitante
<120>	Título de la invención
<160>	Número de SEQ ID NO
<210>	SEQ ID NO: x
<211>	Longitud
<212>	Tipo
<213>	Organismo
<400>	Secuencia

Cuando el nombre del solicitante (identificador numérico <110>) esté escrito en caracteres que no pertenezcan al alfabeto latino, también deberá aparecer en caracteres del alfabeto latino, sea como simple transliteración del nombre o mediante su traducción al inglés.

Excepto para los identificadores numéricos <110>, <120> y <160>, se deberán repetir los elementos de datos para cada secuencia incluida en la lista de secuencias. Cuando no haya una secuencia asociada a un identificador de secuencia (véase el párrafo 5 y el SEQ ID NO: 4 en el ejemplo que se ofrece en el Apéndice 3 de esta Norma) sólo son obligatorios los elementos de datos correspondientes a los identificadores numéricos <210> y <400>.

26. Además de los elementos de datos indicados en el párrafo 25, cuando se suministre una lista de secuencias en cualquier momento antes de la asignación del número de solicitud, se incluirá el siguiente elemento de dato en la lista de secuencias:

<130>	Referencia del expediente
-------	---------------------------

27. Además de los elementos de datos indicados en el párrafo 25, cuando se suministre una lista de secuencias en cualquier momento después de la asignación del número de solicitud, se incluirán los siguientes elementos de datos en la lista de secuencias:

<140>	Solicitud de patente en trámite
<141>	Fecha de presentación de la solicitud en trámite



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.9

28. Además de los elementos de datos indicados en el párrafo 25, cuando se presente una lista de secuencias en relación con una solicitud en la que se reivindica la prioridad de una solicitud anterior, se incluirán los siguientes elementos de datos en la lista de secuencias:

<150>	Solicitud de patente anterior
<151>	Fecha de presentación de la solicitud anterior

29. Si se emplea en la secuencia “n” o “Xaa” o una base modificada o un L-aminoácido modificado o poco común, los siguientes elementos de datos serán obligatorios:

<220>	Característica
<221>	Nombre/clave
<222>	Localización
<223>	Otra información

30. Si el organismo (identificador numérico <213>) es una “Secuencia artificial” o es “Desconocido,” son obligatorios los siguientes elementos de datos:

<220>	Característica
<223>	Otra información

Elementos de datos facultativos

31. Todos los elementos de datos definidos en el Apéndice 1, que no están mencionados en los párrafos 25 a 30, son facultativos (elementos de datos facultativos).

Presentación de las características

32. Las características de las secuencias (es decir, el identificador numérico <220>), se describirán mediante las “claves de caracterización” que figuran en los Cuadros 5 y 6 del Apéndice 2.³⁵

Texto libre

33. Por “texto libre” se entiende la descripción de las características de una secuencia conforme al identificador numérico <223> (otra información) en la que se emplea un vocabulario distinto del vocabulario lingüísticamente neutro mencionado en el párrafo 2.vii).

34. La utilización del texto libre se limitará a unos pocos términos cortos que sean indispensables para entender la secuencia. Cada elemento de dato no excederá de cuatro líneas con un máximo de 65 caracteres por línea, cuando esté escrito en inglés. Se incluirá toda información adicional en la parte principal de la descripción en el idioma de la misma.

35. Preferentemente, todo texto libre deberá estar redactado en inglés.

REPETICIÓN DE TEXTO LIBRE EN LA PARTE PRINCIPAL DE LA DESCRIPCIÓN

36. Si la lista de secuencias que forma parte de la solicitud internacional contiene texto libre, dicho texto deberá repetirse en la parte principal de la descripción en el idioma de ésta. Se recomienda

³⁵ *Nota del editor:* Estos cuadros contienen fragmentos del cuadro de características del DDBJ/EMBL/GenBank (secuencias de nucleótidos) y del cuadro de características de SWISS PROT (secuencias de aminoácidos).



colocar el texto libre que figure en el idioma de la parte principal de la descripción en una sección específica de la descripción titulada “Texto libre de la lista de secuencias”.

LISTA DE SECUENCIAS EN FORMATO ELECTRÓNICO

37. Toda lista de secuencias mencionada en el párrafo 3 que figura en una solicitud internacional presentada en formato electrónico deberá figurar en un formato de documento electrónico y presentarse por un medio de transmisión que haya sido especificado por la Oficina receptora a los fines de la presentación de solicitudes internacionales en formato electrónico, siempre y cuando ese tipo de lista de secuencias figure preferiblemente en el formato de documento electrónico especificado en el párrafo 40 y sea presentada, si es posible, por un medio de transmisión que haya sido especificado por la Oficina receptora y la autoridad competente^{36, 37}.

38. Toda lista de secuencias en formato electrónico mencionada en el párrafo 3*bis* deberá figurar en un formato de documento electrónico que haya sido especificado por la Oficina receptora (en el caso de una corrección) o por la autoridad competente (en el caso de una rectificación o modificación) a los fines de la presentación de solicitudes internacionales en formato electrónico, siempre y cuando esa lista figure preferiblemente en el formato de documento electrónico especificado en el párrafo 40. Ese tipo de lista deberá presentarse por medio de un medio de transmisión que haya sido especificado por la Oficina receptora o la autoridad competente, según proceda, a los fines del presente párrafo; si es posible, deberá presentarse preferiblemente por un medio de transmisión que haya sido especificado por la Oficina receptora y la autoridad competente.³⁸

39. Toda lista de secuencias en formato electrónico mencionada en los párrafos 4 y 4*bis* suministrada a los fines de la búsqueda internacional o del examen preliminar internacional deberá figurar en el formato de documento electrónico especificado en el párrafo 40 y presentarse por un medio de transmisión que haya sido especificado por la autoridad competente a los fines del presente párrafo.

40. A los fines de la búsqueda internacional y del examen preliminar internacional, toda lista de secuencias en formato electrónico deberá figurar en un archivo electrónico codificado por medio de la página de código 437 de IBM³⁹, la página de código 932 de IBM⁴⁰ o una página de código compatible para representar la lista de secuencias en la forma expuesta en los párrafos 5 a 36 sin ningún otro código incluido. En una página de código compatible necesaria, por ejemplo, en el caso de los

³⁶ *Nota del editor:* Cuando la autoridad competente no disponga de una lista de secuencias en formato electrónico que cumpla con lo dispuesto en la presente Norma de una forma y manera que sean aceptables para ella (es decir, en particular, cuando la autoridad competente no disponga de la lista en el formato de documento electrónico especificado en el párrafo 40), podrá requerir al solicitante que le entregue esa lista de secuencias en formato electrónico (véase la Regla 13*ter*).

³⁷ *Nota del editor:* Independientemente del formato de documento electrónico de la lista de secuencias, deberá mantenerse la relación espacial (por ejemplo, columnas e hileras) de los elementos de datos incluidos en la lista de secuencias y el formato de las secuencias de nucleótidos o aminoácidos, en la forma que se especifica en el presente Anexo.

³⁸ *Nota del editor:* Cuando la autoridad competente no disponga de una lista de secuencias de reemplazo en formato electrónico en la que figuren correcciones, rectificaciones o modificaciones de una forma y manera que sean aceptables para ella (es decir, en particular, cuando no disponga de ella en el formato de documento electrónico especificado en el párrafo 40), únicamente tendrá que tener en cuenta ese tipo de corrección, rectificación o modificación a los fines de la búsqueda internacional o del examen preliminar internacional en la medida en que se pueda efectuar una búsqueda significativa o un examen preliminar significativo sin disponer de la lista de secuencias de reemplazo (véase el párrafo 4*bis*). Véase asimismo la Nota del editor N° 38, que se aplica igualmente a toda lista de secuencias de reemplazo en formato electrónico mencionada en el párrafo 3*bis*.

³⁹ *Nota del editor:* IBM es una marca registrada de *International Business Machine Corporation* (Estados Unidos de América).

⁴⁰ *Nota del editor:* Las páginas de código especificadas constituyen de hecho las de uso corriente en las computadoras personales.



caracteres japoneses, chinos, cirílicos, árabes, griegos o hebreos, se asigna el alfabeto y los caracteres latinos a las mismas posiciones hexadecimales que en las páginas de código especificadas.

41. Toda lista de secuencias en el formato de documento electrónico especificado en el párrafo 40 deberá estar creada, preferiblemente, mediante un programa informático especializado, como *PatentIn*.

PROCEDIMIENTO ANTE LAS OFICINAS DESIGNADAS Y ELEGIDAS

42. A los fines del procedimiento ante una Oficina designada o elegida ante la que se haya iniciado la tramitación de una solicitud internacional que contenga la divulgación de una o más secuencia de nucleótidos o de aminoácidos (véase la Regla 13^{ter}.3):

i) toda referencia a la Oficina receptora o a la autoridad competente deberá interpretarse como una referencia a la Oficina designada o elegida en cuestión;

ii) toda referencia a una lista de secuencias que figura en la solicitud internacional por medio de una corrección conforme a la Regla 26, una rectificación conforme a la Regla 91 o una modificación conforme al Artículo 34.2)b) de la descripción en relación con secuencias que figuran en la solicitud tal como fue presentada deberá interpretarse en el sentido de que incluye asimismo toda lista de secuencias que figure en la solicitud, en virtud de la legislación nacional aplicada por la Oficina designada o elegida en cuestión, por medio de una corrección (de un defecto de forma), una rectificación (de un error evidente) o una modificación de la descripción en relación con las secuencias que figuran en la solicitud tal como fue presentada;

iii) toda referencia a una lista de secuencias suministrada a los fines de la búsqueda internacional o del examen preliminar internacional deberá interpretarse en el sentido de que incluye asimismo toda lista suministrada a la Oficina designada o elegida en cuestión a los fines de la búsqueda o el examen nacional por esa Oficina;

iv) la Oficina designada o elegida en cuestión podrá requerir al solicitante que le entregue, dentro de un plazo que deberá ser razonable con arreglo a las circunstancias, a los fines de la búsqueda o el examen nacional, una lista de secuencias en formato electrónico que se halle en cumplimiento de la presente Norma, salvo que esa Oficina ya disponga de ese tipo de lista en formato electrónico de una forma y manera que sean aceptables para ella.



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.12

Apéndices

Apéndice 1: Identificadores numéricos

Apéndice 2: Símbolos de nucleótidos y aminoácidos y Cuadro de características

Cuadro 1: Lista de nucleótidos

Cuadro 2: Lista de nucleótidos modificados

Cuadro 3: Lista de aminoácidos

Cuadro 4: Lista de aminoácidos modificados y poco comunes

Cuadro 5: Lista de claves de caracterización de secuencias de nucleótidos

Cuadro 6: Lista de claves de caracterización de secuencias proteínicas

Apéndice 3: Ejemplo de lista de secuencias

[Siguen los Apéndices 1 a 3 del Anexo C]



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.13

Anexo C, Apéndice 1 Identificadores Numéricos

En las listas de secuencias presentadas en las solicitudes sólo podrán emplearse los identificadores numéricos que se definen a continuación. El texto de los encabezamientos de los siguientes elementos de datos no debe figurar en las listas de secuencias.

Los identificadores numéricos de los elementos de datos obligatorios, es decir, los elementos de datos que deben incluirse en todas las listas de secuencias (véase el párrafo 25 de esta Norma: identificadores 110, 120, 160, 210, 211, 212, 213 y 400) y los identificadores numéricos de los elementos de datos que deben incluirse en las circunstancias que especifica esta norma (véanse los párrafos 26, 27, 28, 29 y 30 de esta Norma: identificadores 130, 140, 141, 150 y 151, y 220 a 223) se señalan con la letra “O”.

Los identificadores numéricos de los elementos de datos facultativos (véase el párrafo 31 de esta Norma) se señalan con la letra “F”.

Identificador numérico	Descripción del identificador numérico	Obligatorio (O) o Facultativo (F)	Comentario
<110>	Nombre del solicitante	O	cuando el nombre del solicitante esté escrito en caracteres distintos del alfabeto latino, también deberá indicarse en caracteres del alfabeto latino sea como simple transliteración o mediante su traducción al inglés
<120>	Título de la invención	M	
<130>	Referencia del expediente	O, en los casos mencionados en el párrafo 26 de esta Norma	véase el párrafo 26 de esta Norma
<140>	Solicitud de patente en trámite	O, en los casos mencionados en el párrafo 27 de esta Norma	véase el párrafo 27 de esta Norma; la solicitud de patente en trámite deberá estar identificada, en el siguiente orden, por el código de dos letras indicado conforme a la Norma ST.3 de la OMPI y el número de solicitud (en el formato utilizado por la oficina de propiedad industrial en la que se presentó la solicitud de patente en trámite) o, si es una solicitud internacional, por el número de solicitud internacional
<141>	Fecha de presentación de la solicitud en trámite	O, en los casos mencionados en el párrafo 27 de esta Norma	véase el párrafo 27 de esta Norma; se indicará la fecha conforme a la Norma ST.2 de la OMPI (AACC MM DD)



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.14

[Anexo C, Apéndice 1, continuación]

Identificador numérico	Descripción del identificador numérico	Obligatorio (O) o Facultativo (F)	Comentario
<150>	Solicitud de patente anterior	O, en los casos mencionados en el párrafo 28 de esta Norma	véase el párrafo 28 de esta Norma; la solicitud de patente anterior deberá identificarse, en el orden siguiente, por el código de dos letras indicado conforme a la Norma ST.3 de la OMPI y por el número de solicitud (en el formato empleado por la oficina de propiedad industrial en la que se presentó la solicitud de patente anterior) o, si es una solicitud internacional, por el número de solicitud internacional
<151>	Fecha de presentación de la solicitud anterior	O, en los casos mencionados en el párrafo 28 de esta Norma	véase el párrafo 28 de esta Norma; se indicará la fecha conforme a la Norma ST.2 de la OMPI (AACC MM DD)
<160>	Número de SEQ ID NO	O	
<170>	Soporte lógico	F	
<210>	Información respecto de SEQ ID NO: x	O	la respuesta estará compuesta por un número entero que represente el SEQ ID NO anotado
<211>	Longitud	O	longitud de la secuencia expresada mediante el número de pares de bases o de aminoácidos
<212>	Tipo	O	tipo de molécula secuenciada en el SEQ ID NO: x, es decir, DNA (ADN), RNA (ARN) o PRT (proteína); si la secuencia de nucleótidos contiene fragmentos de ADN y de ARN, el valor será "DNA"; además, la molécula combinada de ADN/ARN también deberá ser objeto de descripción en la sección de características <220> a <223>
<213>	Organismo	O	Género y especie (es decir, nombre científico) o "Artificial Sequence" (Secuencia artificial) o "Unknown" (Desconocido)



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.15

[Anexo C, Apéndice 1, continuación]

Identificador numérico	Descripción del identificador numérico	Obligatorio (O) o Facultativo (F)	Comentario
<220>	Característica	O, en los casos mencionados en los párrafos 29 y 30 de esta Norma	déjese en blanco; véanse los párrafos 29 y 30 de esta Norma; la descripción de rasgos de importancia biológica para la secuencia en SEQ ID NO: x (puede repetirse en función del número de características indicadas)
<221>	Nombre/clave	O, en los casos mencionados en el párrafo 29 de esta Norma	véase el párrafo 29 de esta Norma; sólo se emplearán las claves descritas en el Cuadro 5 ó 6 del Apéndice 2
<222>	Localización	O, en los casos mencionados en el párrafo 29 de esta Norma	véase el párrafo 29 de esta Norma; – “FROM” (“desde”, número de la primera base/aminoácido en la característica) – “TO” (“hasta”, número de la última base/aminoácido en la característica) – base pairs (pares de bases, los números hacen referencia a las posiciones de los pares de bases en la secuencia de nucleótidos) – amino acids (aminoácidos, los números hacen referencia a las posiciones de los residuos de aminoácidos en una secuencia de aminoácidos) – si la característica se encuentra en la cadena complementaria de la que ha sido presentada en la lista de secuencia



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.16

[Anexo C, Apéndice 1, continuación]

Identificador numérico	Descripción del identificador numérico	Obligatorio (O) o Facultativo (F)	Comentario
<223>	Otras informaciones:	O, en los casos mencionados en los párrafos 29 y 30 de esta Norma	véanse los párrafos 29 y 30 de esta Norma; cualquier otra información pertinente, utilizando un vocabulario lingüísticamente neutro, o en forma de texto libre (preferentemente en inglés); todo texto libre debe repetirse en la parte principal de la descripción en el mismo idioma (véase el párrafo 36 de esta Norma); cuando esté presente en la secuencia una base modificada o un L-aminoácido modificado o poco común que figure en los Cuadros 2 y 4 del Apéndice 2, se empleará el símbolo asociado con dicha base o aminoácido según los Cuadros 2 y 4 del Apéndice 2
<300>	Información sobre la publicación	F	déjese en blanco; repítase la sección respecto de cada publicación pertinente
<301>	Autores	F	
<302>	Título	F	título de la publicación
<303>	Revista	F	nombre de la revista en la que se publicaron los datos
<304>	Volumen	F	volumen de la revista en la que se publicaron los datos
<305>	Número	F	número de la revista en la que se publicaron los datos
<306>	Páginas	F	números de las páginas de la revista en la que se publicaron los datos
<307>	Fecha	F	fecha de la revista en la que se publicaron los datos; de ser posible, se indicará la fecha con arreglo a la Norma ST.2 de la OMPI (AACC MM DD)
<308>	Número de entrada en la base de datos	F	número de entrada asignado por la base de datos, incluido el nombre de la base de datos
<309>	Fecha de entrada en la base de datos	F	fecha de entrada en la base de datos; se indicará la fecha con arreglo a la Norma ST.2 de la OMPI (AACC MM DD)



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.17

[Anexo C, Apéndice 1, continuación]

Identificador numérico	Descripción del identificador numérico	Obligatorio (O) o Facultativo (F)	Comentario
<310>	Número de documento	F	número del documento, únicamente para las citas de patentes; se especificará el documento completo según el siguiente orden: el código de dos letras indicado con arreglo a la Norma ST.3 de la OMPI, el número de publicación indicado con arreglo a la Norma ST.6 de la OMPI, y el código de tipo de documento indicado con arreglo a la Norma ST.16 de la OMPI
<311>	Fecha de presentación	F	fecha de presentación del documento, únicamente para citas de patentes; se indicará la fecha con arreglo a la Norma ST.2 de la OMPI (AACC MM DD)
<312>	Fecha de publicación	F	fecha de publicación del documento; únicamente para citas de patentes; se indicará la fecha con arreglo a la Norma ST.2 de la OMPI (AACC MM DD)
<313>	Residuos relevantes en SEQ ID NO: x: from to	F	
<400>	Secuencia	O	el elemento SEQ ID NO: x debe ir después del identificador numérico y figurar en la línea anterior a la secuencia (véase el Apéndice 3)

[Anexo C, sigue el Apéndice 2]



Anexo C, Apéndice 2
Cuadro de símbolos de nucleótidos y aminoácidos
y cuadro de características

Cuadro 1: Lista de nucleótidos

Símbolo	Significado	Origen de la designación
a	a	adenina
g	g	guanina
c	c	citosa
t	t	timina
u	u	uracilo
r	g ó a	purina
y	t/u ó c	pirimidina (pyrimidine)
m	a ó c	amino
k	g ó t/u	ceto (keto)
s	g ó c	interacciones fuertes (strong interactions) 3 puentes de hidrógeno
w	a ó t/u	interacciones débiles (weak interactions) 2 puentes de hidrógeno
b	g ó c ó t/u	que no sea a
d	a ó g ó t/u	que no sea c
h	a ó c ó t/u	que no sea g
v	a ó g o c	que no sea t, ni u
n	a ó g ó c ó t/u, desconocido u otro	cualquiera (any)



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.19

[Anexo C, Apéndice 2, continuación]

Cuadro 2: Lista de nucleótidos modificados

Símbolo	Significado
ac4c	4-acetilcitidina
chm5u	5-(carboxihidroximetil)uridina
cm	2'-O-metilcitidina
cmnm5s2u	5-carboximetilaminometil-2-tiouridina
cmnm5u	5-carboximetilaminometiluridina
d	dihidrouridina
fm	2'-O-metilpseudouridina
gal q	beta, D-galactosilqueosina
gm	2'-O-metilguanosina
I	inosina
i6a	N6-isopenteniladenosina
m1a	1-metiladenosina
m1f	1-metilpseudouridina
m1g	1-metilguanosina
m1I	1-metilinosina
m22g	2,2-dimetilguanosina
m2a	2-metiladenosina
m2g	2-metilguanosina
m3c	3-metilcitidina
m5c	5-metilcitidina
m6a	N6-metiladenosina
m7g	7-metilguanosina
mam5u	5-metilaminometiluridina
mam5s2u	5-metoxiaminometil-2-tiouridina
man q	beta, D-manosilqueosina
mcm5s2u	5-metoxicarbonilmetil-2-tiouridina
mcm5u	5-metoxicarbonilmetiluridina
mo5u	5-metoxiuridina
ms2i6a	2-metiltio-N6-isopenteniladenosina
ms2t6a	N-((9-beta-D-ribofuranosil-2-metiltiopurin-6-ilo)carbamoil)treonina



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.20

[Anexo C, Apéndice 2, continuación]

mt6a	N-((9-beta-D-ribofuranosilpurin-6-ilo)N-metilcarbamoil)treonina
mv	5-metoxicarbonilmetoxiuridina
o5u	5-carboximetoxiuridina
osyw	Wybutoxosina
p	Pseudouridina
q	Queosina
s2c	2-tiocitidina
s2t	5-metil-2-tiouridina
s2u	2-tiouridina
s4u	4-tiouridina
t	5-metiluridina
t6a	N-((9-beta-D-ribofuranosilpurin-6-ilo)-carbamoil)treonina
tm	2'-O-metil-5-metiluridina
um	2'-O-metiluridina
yw	Wybutosina
x	3-(3-amino-3-carboxipropil)uridina, (acp3)u



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.21

[Anexo C, Apéndice 2, continuación]

Cuadro 3: Lista de aminoácidos

Símbolo	Significado
Ala	Alanina
Cys	Cisteína
Asp	Ácido aspártico
Glu	Ácido glutámico
Phe	Fenilalanina
Gly	Glicina
His	Histidina
Ile	Isoleucina
Lys	Lisina
Leu	Leucina
Met	Metionina
Asn	Asparagina
Pro	Prolina
Gln	Glutamina
Arg	Arginina
Ser	Serina
Thr	Treonina
Val	Valina
Trp	Triptófano
Tyr	Tirosina
Asx	Asp o Asn
Glx	Glu o Gln
Xaa	Desconocido u otro



[Anexo C, Apéndice 2, continuación]

Cuadro 4: Lista de aminoácidos modificados o poco comunes

Símbolo	Significado
Aad	Ácido 2-aminoadípico
Baad	Ácido 3-aminoadípico
bAla	Beta-Alanina, ácido beta-aminopropiónico
Abu	Ácido 2-aminobutírico
4Abu	Ácido 4-aminobutírico, ácido piperidínico
Acp	Ácido 6-aminocaproico
Ahe	Ácido 2-aminoheptanoico
Aib	Ácido 2-aminoisobutírico
bAib	Ácido 3-aminoisobutírico
Apm	Ácido 2-aminopimélico
Dbu	Ácido 2,4-diaminobutírico
Des	Desmosina
Dpm	Ácido 2,2'-diaminopimélico
Dpr	Ácido 2,3-diaminopropiónico
EtGly	N-etilglicina
EtAsn	N-etilasparagina
Hyl	Hidroxilisina
aHyl	alo-Hidroxilisina
3Hyp	3-Hidroxiprolina
4Hyp	4-Hidroxiprolina
Ide	Isodesmosina
alle	alo-Isoleucina
MeGly	N-metilglicina, sarcosina
MeIle	N-metilisoleucina
MeLys	6-N-metil-lisina
MeVal	N-metilvalina
Nva	Norvalina
Nle	Norleucina
Orn	Ornitina



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.23

[Anexo C, Apéndice 2, continuación]

Cuadro 5: Lista de claves de caracterización de secuencias de nucleótidos

Clave	Descripción
allele	existen individuos o estirpes relacionadas que contienen formas estables y diferentes del mismo gen, y que difieren de la secuencia presentada en esta localización (y quizás en otras)
attenuator	1) región del ADN donde se produce una regulación de la terminación de la transcripción que controla la expresión de ciertos operones bacterianos; 2) segmento de secuencia situado entre el promotor y el primer gen estructural que causa la terminación parcial de la transcripción
C_region	región constante de las cadenas pesadas y ligeras de las inmunoglobulinas y de las cadenas alfa, beta y gamma del receptor del linfocito- T; incluye uno o más exones, según la cadena
CAAT_signal	caja CAAT; parte de una secuencia conservada que se sitúa 75 pares de bases en dirección 5' desde el sitio de iniciación de las unidades de transcripción eucarióticas, y que puede desempeñar un papel importante en la unión de la ARN polimerasa; consenso=GG (C o T) CAATCT
CDS	secuencia codificadora (" <i>coding sequence</i> "); secuencia de nucleótidos que se corresponde con la secuencia de aminoácidos en una proteína (la localización incluye el codón de terminación); contiene la traducción conceptual de los aminoácidos
conflict	las determinaciones independientes de la "misma" secuencia difieren en este sitio o en esta región
D-loop	lazo de desplazamiento (" <i>Displacement loop</i> "); región del ADN mitocondrial en la que una fibra corta de ARN se aparea con una cadena de ADN, desplazando en este sitio a la otra cadena de la doble hélice; utilizado también para describir el desplazamiento de una región de una cadena de ADN doble debido a un invasor monocatenario en la reacción catalizada por la proteína RecA
D-segment	segmento de diversidad (" <i>Diversity segment</i> ") de la cadena pesada de las inmunoglobulinas y de la cadena beta del receptor del linfocito- T
enhancer	secuencia que aumenta la utilización de (ciertos) promotores eucarióticos situados en la misma cadena de ADN (efecto en cis), y cuya acción puede efectuarse con independencia de la orientación y la localización (5' ó 3') respecto del promotor
exon	región del genoma que codifica una parte del ARN mensajero procesado; puede contener la región 5'UTR, todas las secuencias codificadoras (CDS) y la región 3'UTR
GC_signal	caja GC; región conservada rica en GC y localizada antes del punto de iniciación de las unidades de transcripción eucarióticas, y que puede adoptar la forma de copias múltiples y producirse en ambos sentidos; consenso=GGGCGG
gene	región de interés biológico identificada como gen y a la que se le ha asignado un nombre
iDNA	ADN intercalar; ADN que se elimina en varios tipos de recombinación
intron	segmento de ADN que se transcribe y luego se elimina por empalme de las secuencias (exones) situadas a ambos lados



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.24

[Anexo C, Apéndice 2, continuación]

Clave	Descripción
J_segment	segmento de ligazón (“ <i>Joining segment</i> ”) de las cadenas ligeras y pesadas de las inmunoglobulinas, así como de las cadenas alfa, beta y gamma del receptor del linfocito- T
LTR	repetición terminal larga (“ <i>long terminal repeat</i> ”); secuencia directamente repetida en los dos extremos de una secuencia definida, del tipo de las que se encuentran típicamente en los retrovirus
mat_peptide	secuencia codificadora de un péptido maduro o de una proteína madura; secuencia codificadora del péptido o de la proteína en su estado de madurez o final tras las modificaciones postraduccionales; la localización no incluye el codón de terminación (a diferencia de las CDS correspondientes)
misc_binding	sitio en un ácido nucleico que se une mediante enlace covalente o no con otra molécula, y que no puede describirse mediante ninguna otra clave de unión (primer_bind o protein_bind)
misc_difference	la secuencia caracterizada es diferente en esta posición de la que se ha presentado en la entrada y no puede describirse mediante ninguna otra clave de diferencia (conflict, unsure, old_sequence, mutation, variation, allele, o modified_base)
misc_feature	región de interés biológico que no puede describirse mediante ninguna otra clave de característica; una característica nueva o poco común
misc_recomb	sitio de cualquier recombinación generalizada, específica de un sitio o replicativa, donde se produce la escisión y ligamiento del ADN de doble cadena y que no puede describirse mediante ninguna otra clave de recombinación (iDNA o virion) ni por calificadores de clave de origen (/insertion_seq, /transposon, /proviral)
misc_RNA	cualquier porción transcrita o ARN que no puede definirse mediante ninguna otra clave de ARN (prim_transcript, precursor_RNA, mRNA, 5’clip, 3’clip, 5’UTR, 3’UTR, exon, CDS, sig_peptide, transit_peptide, mat_peptide, intron, polyA_site, rRNA, tRNA, scRNA, o snRNA)
misc_signal	cualquier región que contiene una señal que controla o modifica una función o la expresión de un gen, que no puede describirse mediante ninguna otra clave de señal (promoter, CAAT_signal, TATA_signal, -35_signal, -10_signal, GC_signal, RBS, polyA_signal, enhancer, attenuator, terminator o rep_origin)
misc_structure	cualquier conformación o estructura secundaria o terciaria que no puede ser descrita mediante ninguna otra clave de estructura (stem_loop o D-loop)
modified_base	el nucleótido indicado está modificado y debe ser sustituido por la molécula indicada (que figura en el valor calificador de mod_base)
mRNA	ARN mensajero; incluye la región no traducida en 5’ (5’UTR), las secuencias codificadoras (CDS, exon) y la región no traducida en 3’ (3’UTR)
mutation	una estirpe relacionada presenta un cambio brusco y transmisible en la secuencia, en esta localización
N_region	región de inserción de nucleótidos adicionales entre los segmentos reordenados de las inmunoglobulinas
old_sequence	la secuencia presentada es la versión modificada de una antigua secuencia en esta localización



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.25

[Anexo C, Apéndice 2, continuación]

Clave	Descripción
polyA_signal	sitio de reconocimiento indispensable para la escisión mediante una endonucleasa, seguida de poliadenilación, de una porción transcrita de ARN; consenso=AATAAA
polyA_site	sitio de un ARN transcrito en el que se añaden residuos de adenina por poliadenilación postranscripcional
precursor_RNA	ARN precursor, cualquier ARN inmaduro; puede abarcar la región cortada en 5' (5' clip), la región no traducida en 5' (5' UTR), las secuencias codificadoras (CDS, exon), las secuencias intercalares (intron), la región no traducida en 3' (3' UTR), y la región cortada en 3' (3' clip)
prim_transcript	cadena transcrita primaria (inicial, no procesada); incluye la región cortada en 5' (5' clip), la región no traducida en 5' (5' UTR), las secuencias codificadoras (CDS, exon), las secuencias intercalares (intron), la región no traducida en 3' (3' UTR), y la región cortada en 3' (3' clip)
primer_bind	sitio de unión no covalente de cebadores en la iniciación de la replicación, la transcripción o la transcripción inversa; incluye los sitios para cebadores sintéticos, por ejemplo, los usados en la reacción en cadena de la polimerasa (PCR)
promoter	región de una molécula de ADN a la que se une la ARN polimerasa para iniciar la transcripción
protein_bind	sitio de unión no covalente de proteínas en un ácido nucleico
RBS	sitio de unión del ribosoma (“ <i>ribosome binding site</i> ”)
repeat_region	región del genoma que contiene unidades de repetición
repeat_unit	Elemento que se repite en una repeat region
rep_origin	origen de la replicación; sitio donde comienza la duplicación de un ácido nucleico para obtener dos copias idénticas
rRNA	ARN ribosómico maduro; ARN que compone la partícula ribonucleoproteica (ribosoma) que sintetiza proteínas desde aminoácidos
S_region	región de cambio (“ <i>Switch region</i> ”) de las cadenas pesadas de inmunoglobulina; participa en la reordenación del ADN de la cadena pesada que conduce a la expresión de una clase diferente de inmunoglobulina por el mismo linfocito- B
satellite	múltiples repeticiones (idénticas o parecidas) en tándem de una unidad de repetición básica corta; muchas de ellas tienen una composición de bases u otra propiedad diferente de la del promedio del genoma, lo que permite separarlas del resto del ADN genómico (banda principal)
scRNA	ARN citoplásmico de pequeño tamaño (“ <i>small cytoplasmic RNA</i> ”); una de las diversas moléculas pequeñas de ARN citoplásmico presentes en el citoplasma y (a veces) en el núcleo de una célula eucariótica
sig_peptide	secuencia codificadora de un péptido señal; secuencia codificadora del dominio N-terminal de una proteína secretada; dicho dominio participa en la integración del polipéptido naciente en la membrana; secuencia líder
snRNA	ARN nuclear de pequeño tamaño (“ <i>small nuclear RNA</i> ”); cualquiera de las muchas especies de ARN de pequeño tamaño confinadas al núcleo; varios snRNA intervienen en el procesamiento del ARN



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.26

[Anexo C, Apéndice 2, continuación]

Clave	Descripción
source	identifica la fuente biológica del intervalo de secuencia indicado; esta clave es obligatoria; cada entrada debe estar compuesta, como mínimo, por una clave única de fuente que abarque la totalidad de la secuencia; se puede utilizar más de una clave de fuente por secuencia
stem_loop	horquilla; región de doble hélice formada por el apareamiento de bases entre secuencias complementarias adyacentes (invertidas) que pertenecen a una sola cadena de ARN o ADN
STS	sitio marcado en la secuencia (“ <i>Sequence Tagged Site</i> ”); se trata de secuencias cortas de ADN presentes en una sola copia que identifican un hito del genoma y que pueden detectarse mediante PCR; el mapa de una región del genoma puede efectuarse determinando el orden de una serie de STS
TATA_signal	caja TATA; caja de Goldberg-Hogness; heptámero conservado, rico en A y T, situado a unos 25 pares de bases antes del punto de iniciación de cada unidad transcrita por la ARN polimerasa II de las células eucarióticas, y que puede participar en la colocación de la enzima para la correcta iniciación de la transcripción; consenso=TATA(A o T)A(A o T)
terminator	secuencia de ADN localizada en el extremo de la porción transcrita o adyacente a un promotor, y que causa que la ARN polimerasa termine la transcripción; también puede tratarse del sitio de unión de la proteína represora
transit_peptide	secuencia codificadora de un péptido de tránsito; secuencia codificadora del dominio N-terminal de un proteína de un orgánulo codificada en el núcleo; este elemento participa en la importación postraduccional de la proteína al orgánulo
tRNA	ARN de transferencia maduro, ARN de pequeño tamaño (75-85 bases), que media la traducción de una secuencia de ácido nucleico en una secuencia de aminoácidos
unsure	el autor no ofrece garantías en cuanto a la exactitud de la secuencia en esta región
V_region	región variable de las cadenas ligeras y pesadas de las inmunoglobulinas y de las cadenas alfa, beta y gamma del receptor del linfocito- T; codifica la región variable en el extremo amino terminal; puede estar compuesta por: V_segment, D_segment, N_region, y J_segment
V_segment	segmento variable de las cadenas ligeras y pesadas de las inmunoglobulinas y de las cadenas alfa, beta y gamma del receptor del linfocito- T; codifica la mayor parte de la región variable (V_region) y los últimos aminoácidos del péptido líder
variation	existe una estirpe relacionada que contiene mutaciones estables del mismo gen (por ejemplo: RFLP, polimorfismos, etc.) y que difiere de la secuencia presentada en esta localización (y quizás en otras)
3'clip	región en el extremo 3' de un ARN precursor, que se corta en el procesamiento
3'UTR	región al extremo 3' (posterior al codón de terminación) de un RNA maduro que no se traduce en proteína
5'clip	región en el extremo 5' de una RNA precursor, que se corta en el procesamiento
5'UTR	región en el extremo 5' (anterior al codón de iniciación) de un RNA maduro que no se traduce en proteína
-10_signal	caja de Pribnow; región conservada situada a unos 10 pares de bases antes del punto



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.27

[Anexo C, Apéndice 2, continuación]

Clave	Descripción
	de iniciación de las unidades de transcripción bacterianas, que puede participar en la unión de la ARN polimerasa; consenso=TAtAaT
-35_signal	héxamero conservado que se sitúa a unos 35 pares de bases antes del sitio de iniciación de las unidades de transcripción bacterianas; consenso=TTGACa[] o TGTTGACA []



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.28

[Anexo C, Apéndice 2, continuación]

Cuadro 6: Lista de claves de caracterización de secuencias de proteínas

Clave	Descripción
CONFLICT	secuencias diferentes según diversos documentos
VARIANT	los autores señalan que existen variaciones de la secuencia
VARSPPLIC	descripción de las variaciones de la secuencia producidas por un empalme diferencial
MUTAGEN	localización que ha sido alterada experimentalmente
MOD_RES	modificación postraduccional de un residuo
ACETYLATION	acetilación N-terminal u otra
AMIDATION	generalmente en el grupo C-terminal de un péptido maduro y activo
BLOCKED	grupo de bloqueo indeterminado, N- o C-terminal
FORMYLATION	formilación de la metionina N-terminal
GAMMA-CARBOXYGLUTAMIC ACID HYDROXYLATION	de la asparagina, el ácido aspártico, la prolina o la lisina
METHYLATION	en general, de la lisina o de la arginina
PHOSPHORYLATION	fosforilación de la serina, la treonina, la tirosina, el ácido aspártico o la histidina
PYRROLIDONE CARBOXYLIC ACID	glutamato N-terminal que ha formado una lactama cíclica interna
SULFATATION	en general, de la tirosina
LIPID	enlace covalente de un fragmento lipídico
MYRISTATE	grupo miristato unido por un enlace amida a un residuo de glicina N-terminal de la forma madura de una proteína o a un residuo interno de lisina
PALMITATE	grupo palmitato unido por un enlace tioéter a un residuo de cisteína o por un enlace éster a un residuo de serina o de treonina
FARNESYL	grupo farnesilo unido por un enlace tioéter a un residuo de cisteína
GERANYL-GERANYL	grupo geranilo-geranilo unido por un enlace tioéter a un residuo de cisteína
GPI-ANCHOR	grupo glicosil fosfatidilinositol (GPI) unido a un grupo alfa-carboxilo del residuo C-terminal de la forma madura de una proteína
N-ACYL DIGLYCERIDE	cisteína N-terminal de la forma madura de una lipoproteína procarionte con un ácido graso unido por una amida y un grupo glicerilo, al que se han unido dos ácidos grasos mediante enlaces éster
DISULFID	enlace disulfuro; los extremos "FROM" y "TO" representan los dos residuos que están unidos por un enlace disulfuro intracatenario; si los extremos 'FROM' a 'TO' son idénticos, el enlace disulfuro es un enlace



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.29

[Anexo C, Apéndice 2, continuación]

Clave	Descripción
	intercatenario y el campo de descripción indica la naturaleza del entrecruzamiento
THIOLEST	tioléster; los extremos 'FROM' y 'TO' representan los dos residuos unidos por el enlace tioléster
THIOETH	enlace tioéter; los extremos 'FROM' y 'TO' representan los dos residuos unidos por el enlace tioéter
CARBOHYD	sitio de glicosilación; la naturaleza del carbohidrato (cuando se conoce) viene indicada en el campo de descripción
METAL	sitio de unión de un ion metálico; en el campo de descripción se indica la naturaleza del metal
BINDING	sitio de unión de cualquier grupo químico (coenzima, grupo prostético, etc.); en el campo de descripción se indica la naturaleza química del grupo
SIGNAL	extensión de una secuencia señal (pre péptido)
TRANSIT	extensión de un péptido de tránsito (mitocondrial, cloroplástico o destinado a un microsoma)
PROPEP	extensión de un propéptido
CHAIN	extensión de la cadena polipeptídica en la proteína madura
PEPTIDE	extensión de un péptido activo liberado
DOMAIN	extensión de un dominio de interés en la secuencia; en el campo de descripción se indica la naturaleza de dicho dominio
CA_BIND	extensión de una región de unión de calcio
DNA_BIND	extensión de una región de unión de ADN
NP_BIND	extensión de una región de unión de un nucleótido fosfato; en el campo de descripción se indica la naturaleza del nucleótido fosfato
TRANSMEM	extensión de una región transmembrana
ZN_FING	extensión de una región de dedo de zinc (<i>zinc finger</i>)
SIMILAR	extensión de una región de similitud con otra secuencia proteica; en el campo de descripción figuran informaciones detalladas sobre esta secuencia
REPEAT	extensión de la repetición interna de la secuencia
HELIX	estructura secundaria: Hélices, por ejemplo, la hélice alfa, la hélice 3-10, o la hélice Pi
STRAND	estructura secundaria: cadena beta, por ejemplo, la cadena beta unida con puentes de hidrógeno, o residuo en un puente beta aislado
TURN	estructura secundaria: giros, por ejemplo, giros mantenidos por puentes de hidrógeno (giros de 3, 4 ó 5 residuos)
ACT_SITE	aminoácido(s) que participa(n) en la actividad de una enzima



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.30

[Anexo C, Apéndice 2, continuación]

Clave	Descripción
SITE	otro sitio de interés en la secuencia
INIT_MET	la secuencia comienza con una metionina de iniciación
NON_TER	el residuo en un extremo de la secuencia no es el residuo terminal; aplicado a la posición 1, ello significa que la primera posición no es la posición N-terminal de la molécula completa; aplicado a la última posición, ello significa que esta posición no es la posición C-terminal de la molécula completa; no hay campo de descripción para esta clave
NON_CONS	residuos no consecutivos; indica que dos residuos de una secuencia no son consecutivos y que existen varios residuos no secuenciados entre ellos
UNSURE	zonas de incertidumbre en la secuencia; se emplea para describir la o las regiones de una secuencia respecto de las cuales los autores no están seguros de su definición

[Anexo C, sigue el Apéndice 3]



Apéndice 3
Ejemplo de lista de secuencias

<110> Smith, John; Smithgene Inc.
<120> Example of a Sequence Listing
<130> 01-00001
<140> PCT/EP98/00001
<141> 1998-12-31
<150> US 08/999,999
<151> 1997-10-15
<160> 4
<170> PatentIn version 2.0
<210> 1
<211> 389
<212> DNA
<213> Paramecium sp.
<220>
<221> CDS
<222> (279)...(389)
<300>
<301> Doe, Richard
<302> Isolation and Characterization of a Gene Encoding a Protease from Paramecium sp.
<303> Journal of Genes
<304> 1
<305> 4
<306> 1-7
<307> 1988-06-31
<308> 123456
<309> 1988-06-31
<400> 1
agctgtagtc attcctgtgt cctcttctct ctgggcttct Caccctgcta atcagatctc 60
agggagagtg tcttgaccct cctctgcctt tgcagcttca Caggcaggca ggcaggcagc 120
tgatgtggca attgctggca gtgccacagg cttttcagcc Aggcttaggg tgggttccgc 180
cgcggcgcgg cggcccctct cgcgctctc tcgcgcctct Ctctcgtctt cctctcgtctc 240
ggacctgatt aggtgagcag gaggaggggg cagttagc atg Gtt tca atg ttc agc 296
Met Val Ser Met Phe Ser
1 5
ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt Gtt tgt ttg ttc caa 344
Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val Cys Leu Phe Gln
10 15 20
tgt ccc aaa gtc ctc ccc tgt cac tca tca ctg Cag ccg aat ctt 389
Cys Pro Lys Val Leu Pro Cys His Ser Ser Leu Gln Pro Asn Leu
25 30 35



MANUAL DE INFORMACIÓN Y DOCUMENTACIÓN EN MATERIA DE PROPIEDAD INDUSTRIAL

Ref.: Normas – ST.25

página: 3.25.32

[Anexo C, Apéndice 3, continuación]

<210> 2
<211> 37
<212> PRT
<213> Paramecium sp.

<400> 2
Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu
1 5 10 15
Phe Val Cys Leu Phe Gln Cys Pro Lys Val Leu Pro Cys His Ser Ser
20 25 30
Leu Gln Pro Asn Leu
35

<210> 3
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Designed peptide based on size and polarity to act as a linker
between the alpha and beta chains of Protein XYZ.

<400> 3
Met Val Asn Leu Glu Pro Met His Thr Glu Ile
1 5 10

<210> 4
<400> 4
000

[Fin del Apéndice 3 y de la Norma]