**NORMA ST. 26 - ANEXO VI**

DOCUMENTO DE ORIENTACIÓN

*Versión 1.~~1~~2*

*Proyecto final*

~~Adoptado por el Comité de Normas Técnicas de la OMPI (CWS)  
 en su quinta sesión, el 2 de junio de 2017~~*Propuesta presentada por el Equipo Técnico SEQL para su consideración y aprobación en la sexta sesión del CWS*

## INTRODUCCIÓN

En la presente Norma se indica que uno de sus objetivos es “permitir que el solicitante establezca una única lista de secuencias en una solicitud de patente que sea aceptable a los efectos tanto de los procedimientos internacionales como nacionales o regionales”. El propósito del presente documento de orientación es velar por que todos los solicitantes y las oficinas de propiedad intelectual (oficinas de PI) entiendan y acepten los requisitos relativos a la inclusión y la representación de las divulgaciones de secuencias, de manera de alcanzar el objetivo mencionado.

El presente documento de orientación consta de esta introducción, un índice de ejemplos, ejemplos de divulgaciones de secuencias y un apéndice que contiene una lista de secuencias en XML con secuencias que figuran en los ejemplos. En esta introducción se explican algunos conceptos y la terminología utilizada en el resto del presente documento. Los ejemplos ilustran los requisitos de párrafos específicos de la Norma, y cada ejemplo ha sido designado con el número de párrafo pertinente. Algunos ejemplos ilustran además otros párrafos y, al final de cada ejemplo, se indican las referencias a otros ejemplos correspondientes. En el índice se indican los números de página de los ejemplos, así como las referencias a otros ejemplos. A cada una de las secuencias de un ejemplo que debe o puede incluirse en una lista de secuencias se ha asignado un identificador de secuencia (SEQ ID NO) y dichas secuencias figuran en formato XML en el [Apéndice](#Appendix) del presente documento.

Para cada ejemplo, toda información explicativa que se presenta con una secuencia ha de ser considerada como la totalidad de la divulgación relativa a esa secuencia. En las respuestas ofrecidas se tiene en cuenta únicamente la información presentada expresamente en el ejemplo.

La orientación que brinda el presente documento está destinada a la preparación de una lista de secuencias que ha de suministrarse **en la fecha de presentación** de una solicitud de patente. Al preparar una lista de secuencias que haya de suministrarse **con posterioridad a la fecha de presentación** de una solicitud de patente debe tenerse en cuenta si la oficina de PI podría considerar que en la información suministrada se añade materia a la divulgación original. Por lo tanto, es posible que la orientación brindada en el presente documento no sea aplicable a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente.

### *Preparación de una lista de secuencias*

Al preparar una lista de secuencias para una solicitud de patente es preciso formularse las preguntas siguientes:

1. ¿Exige el párrafo 7 de la Norma ST.26 la inclusión de una secuencia divulgada en particular?

2. Si no se exige la inclusión de una secuencia divulgada en particular, ¿permite la Norma ST.26 la inclusión de esa secuencia?

3. Si la Norma ST.26 exige o permite la inclusión de una secuencia divulgada en particular, ¿cómo debería representarse esa secuencia en la lista de secuencias?

En cuanto a la primera pregunta, el párrafo 7 de la Norma ST.26 (con determinadas restricciones) exige la inclusión de una secuencia divulgada en una solicitud de patente mediante la **enumeración de sus residuos**, cuando la secuencia contiene diez o más nucleótidos **específicamente definidos** o cuatro o más **aminoácidos específicamente** definidos.

En cuanto a la segunda pregunta, el párrafo 8 de la Norma ST.26 prohíbe la inclusión cualquier secuencia que tenga menos de diez nucleótidos **específicamente definidos** o cuatro aminoácidos **específicamente definidos**.

Para responder a esas preguntas, es necesario entender claramente el significado de “enumeración de sus residuos” y “específicamente definidos”.

En cuanto a la tercera pregunta, el presente documento presenta divulgaciones de secuencias que sirven para ejemplificar distintas hipótesis, junto con un examen completo de los medios de representación preferidos para cada secuencia o, cuando una secuencia contiene múltiples variaciones, la **“secuencia más abarcadora”,** de conformidad con la presente Norma. Puesto que es imposible abordar todas las hipótesis posibles de secuencias inusuales, el presente documento de orientación intenta exponer el razonamiento en el que se funda el enfoque respecto de cada ejemplo, así como la forma de aplicar las disposiciones de la Norma ST.26, de manera que el mismo razonamiento pueda aplicarse a otras hipótesis de secuencias que no estén ejemplificadas.

#### “Enumeración de sus residuos”

En el párrafo 3.c) de la Norma ST.26 se define “**enumeración de sus residuos”** como la divulgación de una secuencia en una solicitud de patente en la que se enumera, por orden, cada residuo de la secuencia, donde i) el residuo se representa mediante un nombre, abreviatura, símbolo o estructura; o ii) varios residuos se representan mediante una fórmula abreviada. Una secuencia debería divulgarse en una solicitud de patente mediante la “enumeración de sus residuos” utilizando **símbolos convencionales**, que son los símbolos de nucleótidos expuestos en el Cuadro 1 de la Sección 1 del Anexo 1 de la Norma ST.26 (es decir, los símbolos en letra minúscula o sus equivalentes en letra mayúscula[[1]](#footnote-1)) y los símbolos de aminoácidos expuestos en el Cuadro 3 de la Sección 3 del Anexo 1 de la Norma ST.26 (es decir, los símbolos en letra mayúscula o sus equivalentes en letra minúscula1). Otros símbolos que no sean los que se exponen en dichos cuadros son “**no convencionales”**.

A veces, una secuencia se divulga de una manera que no es la preferida mediante “enumeración de sus residuos” utilizando **abreviaturas convencionales** o **nombres completos** (en lugar de símbolos convencionales), expuestos en los Cuadros A y B, más abajo, símbolos convencionales o abreviaturas utilizados de manera no convencional, símbolos o abreviaturas no convencionales, fórmulas/estructuras químicas, o fórmulas abreviadas. Debería procurarse divulgar las secuencias de la manera preferida; sin embargo, cuando las secuencias se divulgan de una manera no es la preferida, podrá ser necesario consultar la explicación de la secuencia en la divulgación para determinar el significado del símbolo o la abreviatura no preferidos que se han utilizado.

Aun cuando se utilizan un símbolo o una abreviatura convencionales, debe consultarse la explicación de la secuencia en la divulgación para confirmar que el símbolo se utiliza de manera convencional. O bien, si el símbolo se utiliza de manera no convencional, la explicación es necesaria para determinar si el párrafo 7 de la Norma ST.26 exige la inclusión en la lista de secuencias o si el párrafo 8 prohíbe la inclusión.

Cuando en la divulgación un símbolo o una abreviatura no convencionales equivalen a un símbolo o una abreviatura convencionales (por ejemplo, “Z1” significa “A”) o una secuencia específica de símbolos convencionales (por ejemplo, “Z1” significa “agga”), se interpreta la secuencia como si se divulgara utilizando el símbolo o símbolos o la abreviatura o abreviaturas convencionales equivalentes, para determinar si el párrafo 7 de la Norma ST.26 exige la inclusión en la lista de secuencias o si el párrafo 8 prohíbe la inclusión. Cuando se utiliza un símbolo no convencional de nucleótido como símbolo de ambigüedad (por ejemplo, X1 = inosina o pseudouridina), pero no es equivalente a uno de los símbolos convencionales de ambigüedad que figuran en el Cuadro 1 de la Sección 1 (es decir, “m”, “r”, “w”, “s”, “y”, “k”, “v”, “h”, “d”, “b”, o “n”), el residuo se interpreta como un residuo “n” para determinar si el párrafo 7 de la Norma ST.26 exige la inclusión de la secuencia en la lista de secuencias o si el párrafo 8 de la Norma ST.26 prohíbe la inclusión. De manera similar, cuando se utiliza un símbolo no convencional de aminoácido como símbolo de ambigüedad (por ejemplo, “Z1” significa “A”, “G”, “S” o “T”), pero no es equivalente a uno de los símbolos convencionales de ambigüedad que figuran en el Cuadro 3 de la Sección 3 (es decir, B, Z, J, o X), el residuo se interpreta como un residuo “X” para determinar si el párrafo 7 de la Norma ST.26 exige la inclusión de la secuencia en la lista de secuencias o si el párrafo 8 de la Norma ST.26 prohíbe la inclusión.

#### “Específicamente definido”

En el párrafo 3.k) de la Norma ST.26 se define “**específicamente definido**” como todo nucleótido distinto a los representados por el símbolo “n” y todo aminoácido distinto a los representados por el símbolo “X” que se enumeran en el Anexo I, utilizándose “n” y “X” de manera convencional, según se describe en el Cuadro 1 de la Sección 1 (es decir, “a o c o g o t/u; ‘unknown’ o ‘other’”) y en el Cuadro 3 de la Sección 3 (es decir, A o R o N o D o C o Q o E o G o H o I o L o K o M o F o P u O o S o U o T o W o Y o V, ‘unknown’ o ‘other’”), respectivamente. La explicación que figura más arriba acerca de los símbolos convencionales o símbolos o abreviaturas no convencionales y su utilización de manera convencional o no convencional se tendrá en cuenta para determinar si un nucleótido o un aminoácido son “específicamente definidos”.

#### “Secuencia más abarcadora”

Cuando una secuencia que cumple con los requisitos del párrafo 7 se divulga mediante enumeración de sus residuos una sola vez en una solicitud, pero se describe de forma diferente en múltiples realizaciones de la invención; por ejemplo, en una realización, “X” en uno o más lugares podría ser cualquier aminoácido, pero en otras realizaciones, “X” podría ser solamente un número limitado de aminoácidos, la Norma ST.26 exige la inclusión en una lista de secuencias tan solo de la secuencia única que ha sido enumerada mediante sus residuos. Conforme a los párrafos 15 y 27, cuando esa secuencia contiene múltiples símbolos de ambigüedad “n” o “X”, se interpreta que “n” o “X” representan cualquier nucleótido o aminoácido, respectivamente, en ausencia de anotaciones adicionales. En consecuencia, la secuencia única que se exige incluir es la secuencia más abarcadora divulgada. La **secuencia más abarcadora** es la secuencia única que tiene variantes de residuos que están representadas por los símbolos de ambigüedad más restrictivos que incluyen en las realizaciones más divulgadas. Sin embargo, se alienta *encarecidamente* a que se incluyan secuencias específicas adicionales, de ser factible, que representen, por ejemplo, realizaciones adicionales que constituyen una parte fundamental de la invención. La inclusión de las secuencias adicionales permite una búsqueda más minuciosa y ofrece información al público acerca de la materia respecto de la cual se solicita una patente.

*Utilización del símbolo de ambigüedad*

#### Utilización correcta del símbolo de ambigüedad “n” en una lista de secuencias

El símbolo “n”

* 1. no debe utilizarse para representar otras cosa que no sea un nucleótido simple;
  2. se interpretará como “a”, “c”, “g”, o “t/u”, excepto cuando se utilice en una descripción más detallada;
  3. deberá utilizarse para representar cualquiera de los nucleótidos siguientes, junto con una descripción más detallada:
     1. un nucleótido modificado, por ejemplo, natural, sintético o que no existe en estado natural, que no pueda representarse de otra manera por ningún otro símbolo descrito en el Anexo I (véase el Cuadro 1 de la Sección 1);
     2. un nucleótido “unknown”, es decir, no determinado, no divulgado o incierto;
     3. un sitio abásico; o
  4. puede utilizarse para representar variaciones de una secuencia, es decir, alternativas, supresiones, inserciones o sustituciones, siendo “n” el símbolo de ambigüedad más restrictivo.

#### Utilización correcta del símbolo de ambigüedad “X” en una lista de secuencias

El símbolo “X”

1. no debe utilizarse para representar otras cosa que no sea un aminoácido simple;
2. se interpretará como “A”, “R”, “N”, “D”, “C”, “Q”, “E”, “G”, “H”, “I”, “L”, “K”, “M”, “F”, “P”, “O”, “S”, “U”, “T”, “W”, “Y”, o “V”, excepto cuando se utilice en una descripción más detallada;
3. deberá utilizarse para representar cualquiera de los aminoácidos siguientes, junto con una descripción más detallada:
   * 1. un aminoácido modificado, por ejemplo, natural, sintético o que no existe en estado natural, que no pueda representarse de otra manera por ningún otro símbolo descrito en el Anexo I (véase el Cuadro 3 de la Sección 3);
     2. un aminoácido “unknown”, es decir, no determinado, no divulgado o incierto; o
4. puede utilizarse para representar variaciones de una secuencia, es decir, alternativas, supresiones, inserciones o sustituciones, siendo “X” el símbolo de ambigüedad más restrictivo.

|  |  |  |
| --- | --- | --- |
| **Símbolo** | **Abreviatura** | **Nombre del nucleótido** |
| a |  | Adenina |
| c |  | Citosina |
| g |  | Guanina |
| t |  | Timina en ADN  Uracilo en ARN (t/u) |
| m | a o c |
| r | a o g |
| w | a o t/u |
| s | c o g |
| y | c o t/u |
| k | g o t/u |
| v | a o c o g; no t/u |
| h | a o c o t/u; no g |
| d | a o g o t/u; no c |
| b | c o g o t/u; no a |
| n | a o c o g o t/u; “unknown” o “other” |

*Cuadro A – Símbolos, abreviaturas y nombres convencionales de nucleótidos*

|  |  |  |
| --- | --- | --- |
| **Símbolo** | **Abreviatura de 3 letras** | **Nombre del aminoácido** |
| A | Ala | Alanina |
| R | Arg | Arginina |
| N | Asn | Asparagina |
| D | Asp | Ácido aspártico (aspartato) |
| C | Cys | Cisteína |
| E | Glu | Ácido glutámico (glutamato) |
| Q | Gln | Glutamina |
| G | Gly | Glicina |
| H | His | Histidina |
| I | Ile | Isoleucina |
| L | Leu | Leucina |
| K | Lys | Lisina |
| M | Met | Metionina |
| F | Phe | Fenilalanina |
| P | Pro | Prolina |
| O | Pyl | Pirrolisina |
| S | Ser | Serina |
| U | Sec | Selenocisteína |
| T | Thr | Treonina |
| W | Trp | Triptófano |
| Y | Tyr | Tirosina |
| V | Val | Valina |
| B | Asx | Ácido aspártico o asparagina |
| Z | Glx | Glutamina o ácido glutámico |
| J | Xle | Leucina o isoleucina |
| X | Xaa | A o R o N o D o C o Q o E o G o H o I o L o K o M o F o P o O o S o U o T o W o Y o V, “unknown” o “other” |

*Cuadro B – Símbolos, abreviaturas y nombres convencionales de aminoácidos*

ÍNDICE DE EJEMPLOS

**Página**

**Párrafo 3.a) – Definición de “aminoácido”**

**Ejemplo 3.a)-1: D-aminoácidos** [**16**](#page16)

***Referencia a otros ejemplos***

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

Ejemplo 30-1: Clave de caracterización “CARBOHYD” [55](#page55)

**Párrafo 3.c) – Definición de “enumeración de sus residuos”**

**Ejemplo 3.c)-1: Enumeración de aminoácidos mediante la estructura química** [**17**](#page17)

**Ejemplo 3.c)-2: Fórmula abreviada de una secuencia de aminoácidos** [**18**](#page18)

***Referencia a otros ejemplos***

Ejemplo 27-1: Fórmula abreviada de una secuencia de nucleótidos [49](#page49)

Ejemplo 27-3: Fórmula abreviada – cuatro o más aminoácidos específicamente definidos [51](#page51)

**Párrafo 3.f) – Definición de “nucleótido modificado”**

***Referencia a otros ejemplos***

Ejemplo 3.g)-4: Análogos de ácido nucleico [22](#page22)

**Párrafo 3.g) – Definición de “nucleótido”**

**Ejemplo 3.g)-1: Secuencia de nucleótidos interrumpida mediante un espaciador C3** [**19**](#page19)

**Ejemplo 3.g)-2: Secuencia de nucleótidos con residuos alternativos, incluido un espaciador C3** [**20**](#page20)

**Ejemplo 3.g)-3: Sitio abásico** [**21**](#page21)

**Ejemplo 3.g)-4: Análogos de ácido nucleico** [**22**](#page22)

***Referencia a otros ejemplos***

Ejemplo 11.b)-1: Secuencia de nucleótidos de doble cadena – longitudes diferentes [45](#page45)

Ejemplo 14-1: El símbolo “t” representa el uracilo en ARN [47](#page47)

**Párrafo 3.k) – Definición de “específicamente definido”**

**Ejemplo 3.k)-1: Símbolos de ambigüedad de nucleótidos** [**23**](#page23)

**Ejemplo 3.k)-2: Símbolo de ambigüedad “n” utilizado de manera tanto convencional como no convencional** [**24**](#page24)

**Ejemplo 3.k)-3: Símbolo de ambigüedad “n” utilizado de manera no convencional** [**25**](#page25)

**Ejemplo 3.k)-4: Los símbolos de ambigüedad distintos de “n” son “específicamente definidos”** [**26**](#page26)

**Ejemplo 3.k)-5: Abreviatura de ambigüedad “Xaa” utilizada de manera no convencional** [**27**](#page27)

**Párrafo 7 – Secuencias para las que se exige la inclusión en una lista de secuencias**

***Referencia a otros ejemplos***

Ejemplo 28-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [52](#page52)

Ejemplo 55-1: Molécula combinada ADN/ARN [61](#Example551dnarna)

Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada [63](#page63)

Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones [65](#page65)

**Párrafo 7.a) – Secuencia de nucleótidos exigida en una lista de secuencias**

**Ejemplo 7.a)-1: Secuencia de nucleótidos ramificada** [**28**](#page28)

**Ejemplo 7.a)-2: Secuencia de nucleótidos lineal que tiene una estructura secundaria** [**30**](#page30)

**Ejemplo 7.a)-3: Símbolos de ambigüedad de nucleótidos utilizados de manera no convencional** [**31**](#page31)

**Ejemplo 7.a)-4: Símbolos de ambigüedad de nucleótidos utilizados de manera no convencional** [**32**](#page32)

**Ejemplo 7.a)-5: Símbolos no convencionales de nucleótidos** [**33**](#page33)

**Ejemplo 7.a)-6: Símbolos no convencionales de nucleótidos** [**34**](#page34)

***Referencia a otros ejemplos***

Ejemplo 3.g)-1: Secuencia de nucleótidos interrumpida por un espaciador C3 [19](#page19)

Ejemplo 3.g)-2: Secuencia de nucleótidos con residuos alternativos, incluido un espaciador C3 [20](#page20)

Ejemplo 3.g)-3: Sitio abásico [21](#page21)

Ejemplo 3.g)-4: Análogos de ácido nucleico [22](#page22)

Ejemplo 3.k)-1: Símbolos de ambigüedad de nucleótidos [23](#page23)

Ejemplo 3.k)-2: Símbolo de ambigüedad “n” utilizado de manera tanto convencional   
como no convencional [24](#page24)

Ejemplo 3.k)-3: Símbolo de ambigüedad “n” utilizado de manera no convencional [25](#page25)

Ejemplo 3.k)-4: Los símbolos de ambigüedad distintos de “n” son “específicamente definidos” [26](#page26)

Ejemplo 11.a)-1: Secuencia de nucleótidos de doble cadena –longitudes iguales [44](#page44)

Ejemplo 11.b)-1: Secuencia de nucleótidos de doble cadena –longitudes diferentes [45](#page45)

Ejemplo 11.b)-2: Secuencia de nucleótidos de doble cadena – ningún segmento con   
apareamiento de bases [46](#page46)

Ejemplo 14-1: El símbolo “t” representa el uracilo en ARN [47](#page47)

Ejemplo 87-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [62](#page62)

Ejemplo 91-1: Representación de variantes enumeradas [67](#page67)

Ejemplo 93.b)-1: Representación de distintas variantes de secuencias con múltiples variaciones interdependientes [72](#page72)

**Párrafo 7.b) – Secuencias de aminoacidos exigidas en una lista de secuencias**

**Ejemplo 7.b)-1: Cuatro o más aminoácidos específicamente definidos** [**35**](#page35)

**Ejemplo 7.b)-2: Secuencia de aminoácidos ramificada** [**35**](#page35)

**Ejemplo 7.b)-3: Secuencia de aminoácidos ramificada** [**39**](#page39)

**Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada** [**40**](#page40)

**Ejemplo 7.b)-5: Péptido cíclico que contiene una secuencia de aminoácidos ramificada** [**43**](#page43)

***Referencia a otros ejemplos***

Ejemplo 3.a)-1: D-aminoácidos [16](#page16)

Ejemplo 3.c)-1: Enumeración de aminoácidos mediante la estructura química [17](#page17)

Ejemplo 3.c)-2: Fórmula abreviada de una secuencia de aminoácidos [18](#page18)

Ejemplo 3.k)-5: Abreviatura de ambigüedad “Xaa” utilizada de manera no convencional [27](#_Párrafo_27_–)

Ejemplo 27-1: Fórmula abreviada de una secuencia de nucleótidos [49](#page49)

Ejemplo 27-3: Fórmula abreviada - cuatro o más aminoácidos específicamente definidos [51](#page51)

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

Ejemplo 30-1: Clave de caracterización “CARBODHYD” [55](#page55)

Ejemplo 36-1: Secuencia con una región de un número conocido de residuos “X” representada como

secuencia única [56](#page56)

Ejemplo 37-1: Una secuencia con regiones de un número desconocido de residuos “X” no debe

representarse como secuencia única [59](#page59)

Ejemplo 37-2: Una secuencia con regiones de un número desconocido de residuos “X” no debe

representarse como secuencia única [60](#page60)

Ejemplo 87-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [55](#page55)

Ejemplo 91-2: Representación de variantes enumeradas [68](#page68)

Ejemplo 91-3: Representación de una secuencia de consenso [69](#page69)

Ejemplo 92-1: Representación de una secuencia única con aminoácidos alternativos enumerados [70](#page70)

Ejemplo 93.a)-1: Representación de una variante de secuencia mediante anotación de   
la secuencia principal [71](#page71)

**Párrafo 8 – Umbral para la inclusión de secuencias**

***Referencia a otros ejemplos***

Ejemplo 3.k)-1: Símbolos de ambigüedad de nucleótidos [23](#page23)

Ejemplo 3.k)-2: Símbolo de ambigüedad “n” utilizado de manera tanto convencional   
como no convencional [24](#page24)

Ejemplo 7.a)-1: Secuencia de nucleótidos ramificada [27](#page27)

Ejemplo 7.a)-6: Símbolos no convencionales de nucleótidos [34](#page34)

Ejemplo 7.b)-1: Cuatro o más aminoácidos específicamente definidos [35](#page35)

Ejemplo 7.b)-2: Secuencia de aminoácidos ramificada [36](#example7b2branched)

Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [40](#example7b4cyclicpeptide)

Ejemplo 14-1: El símbolo “t” representa el uracilo en ARN [47](#page47)

Ejemplo 37-1: Una secuencia con regiones de un número desconocido de residuos “X” no debe

representarse como secuencia única [59](#page59)

Ejemplo 37-2: Una secuencia con regiones de un número desconocido de residuos “X” no debe

representarse como secuencia única [60](#page60)

Ejemplo 92-1: Representación de una secuencia única con aminoácidos alternativos enumerados [70](#page70)

**Párrafo 11 – Representación de una secuencia de nucleótidos**

***Referencia a otros ejemplos***

Ejemplo 3.g)-4: Análogos de ácido nucleico [22](#page22)

Ejemplo 7.a)-1: Secuencia de nucleótidos ramificada [28](#page28)

**Párrafo 11.a) Secuencia de nucleótidos de doble cadena - totalmente complementarias**

**Ejemplo 11.a)-1: Secuencia de nucleótidos de doble cadena – longitudes iguales** [**44**](#page44)

**Párrafo 11.b) – Secuencia de nucleótidos de doble cadena – no totalmente complementarias**

**Ejemplo 11.b)-1: Secuencia de nucleótidos de doble cadena – longitudes diferentes** [**45**](#page45)

**Ejemplo 11.b)-2: Secuencia de nucleótidos de doble cadena – ningún segmento con apareamiento de bases** [**46**](#page46)

**Párrafo 13 – Representación de nucleótidos**

***Referencia a otros ejemplos***

Ejemplo 3.k)-2: Símbolo de ambigüedad “n” utilizado de manera tanto convencional   
como no convencional [24](#page24)

Ejemplo 7.a)-1: Secuencia de nucleótidos ramificada [28](#_Párrafo_28_–)

Ejemplo 14-1: El símbolo “t” representa el uracilo en ARN [47](#page47)

Ejemplo 91-1: Representación de variantes enumeradas [67](#page67)

**Párrafo 14 – Símbolo “t” interpretado como uracilo en ARN**

**Ejemplo 14-1: El símbolo “t” representa el uracilo en ARN** [**47**](#page47)

***Referencia a otros ejemplos***

Ejemplo 55-1: Molécula combinada ADN/ARN [61](#page61)

**Párrafo 15 – Debería utilizarse el símbolo de ambigüedad de nucleótidos más restrictivo**

***Referencia a otros ejemplos***

Ejemplo 3.g)-1: Secuencia de nucleótidos interrumpida por un espaciador C3 [19](#page19)

Ejemplo 3.g)-2: Secuencia de nucleótidos con residuos alternativos, incluido un espaciador C3 [20](#page20)

Ejemplo 3.k)-4: Los símbolos de ambigüedad distintos de “n” son “específicamente definidos” [26](#page26)

Ejemplo 93.b)-1: Representación de distintas variantes de secuencias con   
múltiples variaciones interdependientes [72](#page72)

**Párrafo 16 – Representación de un aminoácido modificado**

***Referencia a otros ejemplos***

Ejemplo 3.g)-1: Secuencia de nucleótidos interrumpida por un espaciador C3 [19](#page19)

Ejemplo 3.g)-4: Análogos de ácido nucléico [22](#page22)

**Párrafo 17 – Anotación de un aminoácido modificado**

***Referencia a otros ejemplos***

Ejemplo 3.g)-1: Secuencia de nucleótidos interrumpida por un espaciador C3 [19](#page19)

Ejemplo 3.g)-3: Sitio abásico [21](#page21)

Ejemplo 7.a)-1: Secuencia de nucleótidos ramificada [28](#page28)

Ejemplo 7.a)-2: Secuencia lineal de nucleótidos con una estructura secundaria [30](#page30)

Ejemplo 7.a)-6: Símbolos no convencionales de nucleótidos [34](#page34)

**Párrafo 18 – Anotación de regiones de nucleótidos moficados consecutivos**

***Referencia a otros ejemplos***

Ejemplo 3.g)-4: Análogos de ácido nucleico [22](#page22)

Ejemplo 11.b)-1: Secuencia de nucleótidos de doble cadena – longitudes diferentes [45](#page45)

**Párrafo 19 – Anotación de uracilo en ADN o timina en ARN**

***Referencia a otros ejemplos***

Ejemplo 14-1: El símbolo “t” representa el uracilo en ARN [47](#page47)

**Párrafo 25 – Residuo de la posición número 1 en la secuencia de aminoácidos**

***Referencia a otros ejemplos***

Ejemplo 3.a)-1: D-aminoácidos [16](#page16)

Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [40](#page40)

Ejemplo 7.b)-5: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [43](#page43)

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

**Párrafo 26 – Representación de aminoácidos**

***Referencia a otros ejemplos***

Ejemplo 7.b)-2: Secuencia de aminoácidos ramificada [36](#page36)

Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [40](#page40)

Ejemplo 7.b)-5: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [43](#page43)

Ejemplo 36-1: Secuencia con una región de un número conocido de residuos “X” representada como

secuencia única [56](#page56)

Ejemplo 87-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [62](#page62)

Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones [65](#page65)

Ejemplo 91-2: Representación de variantes enumeradas [68](#page68)

Ejemplo 91-3: Representación de una secuencia de consenso [69](#page69)

**Párrafo 27 – Debería utilizarse el símbolo de ambigüedad de aminoácidos más restrictivo**

**Ejemplo 27-1: Fórmula abreviada de una secuencia de nucleótidos** [**49**](#page49)

**Ejemplo 27-2: Fórmula abreviada – menos de cuatro aminoácidos específicamente definidos** [**50**](#page50)

**Ejemplo 27-3: Fórmula abreviada - cuatro o más aminoácidos específicamente definidos** [**51**](#page51)

***Referencia a otros ejemplos***

Ejemplo 3.c)-2: Fórmula abreviada de una secuencia de aminoácidos [18](#page18)

Ejemplo 7.b)-1: Cuatro o más aminoácidos específicamente definidos [35](#page35)

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

Ejemplo 36-1: Secuencia con una región de un número conocido de residuos “X” representada como

secuencia única [56](#page56)

Ejemplo 36-2: Secuencia con múltiples regiones de un número o una serie conocidos de residuos “X”

representada como secuencia única [57](#page57)

Ejemplo 36-3: Secuencia con múltiples regiones de un número o una serie conocidos de residuos “X”

representada como secuencia única [58](#page58)

Ejemplo 37-2: Una secuencia con regiones de un número desconocido de residuos "X" no debe representarse como secuencia única. [60](#page60)

Ejemplo 91-3: Representación de una secuencia de consenso [69](#page69)

Ejemplo 92-1: Representación de una secuencia única con aminoácidos alternativos enumerados [70](#page70)

Ejemplo 93.a)-1: Representación de una variante de secuencia mediante anotación de la secuencia   
principal [71](#page71)

**Párrafo 28 – Secuencias de aminoácidos separadas por símbolos internos de terminación**

**Ejemplo 28-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada** [**52**](#page52)

***Referencia a otros ejemplos***

Ejemplo 87-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [62](#page62)

Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones [65](#page65)

**Párrafo 29 – Representación de un aminoácido “other” modificado**

**Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other”** [**54**](#page54)

***Referencia a otros ejemplos***

Ejemplo 3.a)-1: D-aminoácidos [16](#page16)

Ejemplo 7.b)-2: Secuencia de aminoácidos ramificada [36](#_top)

Ejemplo 7.b)-3: Secuencia de aminoácidos ramificada [39](#page39)

Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [40](#page40)

Ejemplo 30-1: Clave de caracterización “CARBODHYD” [55](#page55)

**Párrafo 30 – Anotación de un aminoácido modificado**

**Ejemplo 30-1: Clave de caracterización “CARBODHYD”** [**55**](#page55)

***Referencia a otros ejemplos***

Ejemplo 3.a)-1: D-aminoácidos [16](#page16)

Ejemplo 3.c)-1: Enumeración de aminoácidos mediante la estructura química [17](#page17)

Ejemplo 7.b)-2: Secuencia de aminoácidos ramificada [36](#page36)

Ejemplo 7.b)-3: Secuencia de aminoácidos ramificada [39](#page39)

Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [40](#page40)

Ejemplo 7.b)-5: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [43](#page43)

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

**Párrafo 31 – Representación de un D-aminoácido**

***Referencia a otros ejemplos***

Ejemplo 3.a)-1: D-aminoácidos [16](#page16)

Ejemplo 3.c)-1: Enumeración de aminoácidos mediante la estructura química [17](#page17)

Ejemplo 7.b)-2: Secuencia de aminoácidos ramificada [36](#page36)

Ejemplo 7.b)-3: Secuencia de aminoácidos ramificada [39](#page39)

Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [40](#page40)

Ejemplo 7.b)-5: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [43](#page43)

**Párrafo 32 – Anotación de un aminoácido “unknown”**

***Referencia a otros ejemplos***

Ejemplo 3.c)-1: Enumeración de aminoácidos mediante la estructura química [17](#page17)

**Párrafo 34 – Anotación de una región contigua de resuduos “X”**

***Referencia a otros ejemplos***

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

**Párrafo 36 – Secuencias que contienen regiones de un número exacto de residuos “n” o “X” contiguos**

**Ejemplo 36-1: Secuencia con una región de un número conocido de residuos “X” representada como   
secuencia única** [**56**](#page56)

**Ejemplo 36-2: Secuencia con múltiples regiones de un número o una serie conocidos de residuos “X”**

**representada como secuencia única** [**57**](#page57)

**Ejemplo 36-3: Secuencia con múltiples regiones de un número o una serie conocidos de residuos “X”**

**representada como secuencia única** [**58**](#page58)

**Párrafo 37 – Secuencias que contienen regiones de un número desconocido de residuos “n” o “X” contiguos**

**Ejemplo 37-1: Una secuencia con regiones de un número desconocido de residuos “X” no debe**

**representarse como secuencia única** [**59**](#page59)

**Ejemplo 37-2: Una secuencia con regiones de un número desconocido de residuos “X” no debe**

**representarse como secuencia única** [**60**](#page60)

**Párrafo 41 – Caracteres reservados**

***Referencia a otros ejemplos***

Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada [63](#page63)

**Párrafo 54 – El elemento INSDSeq\_moltype**

***Referencia a otros ejemplos***

Ejemplo 14-1: El símbolo “t” representa el uracilo en ARN [47](#page47)

**Párrafo 55 – Secuencia de nucleótidos que contiene segmentos de ADN y ARN**

**Ejemplo 55-1: Molécula combinada ADN/ARN ……………………………..…………………………………………………….**[61](#Example551dnarna)

**Párrafo 54 – Ejemplo que ilustra una secuencia de nucleótidos que contiene segmentos de ADN y ARN**

***Referencia a otros ejemplos***

Ejemplo 55-1: Molécula combinada ADN/ARN [61](#page61)

**Párrafo 57 – El elemento INSDSeq\_*sequence***

***Referencia a otros ejemplos***

Ejemplo 28-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [52](#page52)

Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones [65](#page65)

**Párrafo 65 – Descriptor de localización**

***Referencia a otros ejemplos***

Ejemplo 3.g)-4: Análogos de ácido nucleico [22](#page22)

Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada [63](#page63)

**Párrafo 66 – Sintaxis del descriptor de localización**

***Referencia a otros ejemplos***

Ejemplo 3.g)-4: Análogos de ácido nucleico [22](#page22)

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada [63](#page63)

**Párrafo 67 – Operador de localización**

***Referencia a otros ejemplos***

Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [40](#page40)

**Párrafo 68 – Operadores de localización de unión (*join*) u orden (*order*)**

***Referencia a otros ejemplos***

Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada [40](#page40)

**Párrafo 70 – Localizaciones de características**

***Referencia a otros ejemplos***

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada [63](#page63)

**Párrafo 71 – Representación de los caracteres “<” and “>” en un descriptor de localización**

***Referencia a otros ejemplos***

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [48](#page48)

Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada [63](#page63)

**Párrafo 83 – Ejemplo que ilustra una secuencia de nucleótidos que no ocurre naturalmente**

***Referencia a otros ejemplos***

Ejemplo 55-1: Molécula combinada ADN/ARN [61](#page61)

**Párrafo 87 – Clave de caracterización “CDS”**

**Ejemplo 87-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada** [**62**](#page62)

**Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada** [63](#page63)

***Referencia a otros ejemplos***

Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones [65](#page65)

**Párrafo 88 – Los calificadores “transl\_table” y “translation”**

***Referencia a otros ejemplos***

Ejemplo 28-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [52](#page52)

Ejemplo 87-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [62](#page62)

Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones [65](#page65)

**Párrafo 90 – Inclusión de una secuencia de aminoácidos codificada en una lista de secuencias**

**Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones** [**65**](#page65)

***Referencia a otros ejemplos***

Ejemplo 28-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [52](#page52)

Ejemplo 87-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada [62](#page62)

Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada [63](#page63)

**Párrafo 91 – Secuencia principal y una variante, cada una de ellas enumerada mediante su residuo**

**Ejemplo 91-1: Representación de variantes enumeradas** [**67**](#page67)

**Ejemplo 91-2: Representación de variantes enumeradas** [**68**](#page68)

**Ejemplo 91-3: Representación de una secuencia de consenso** [**69**](#page69)

**Párrafo 92 – Variante de secuencia divulgada como secuencia única con residuos alternativos enumerados**

**Ejemplo 92-1: Representación de una secuencia única con aminoácidos alternativos enumerados** [**70**](#page70)

**Párrafo 93.a) – Variante de secuencia divulgada únicamente mediante referencia a una secuencia principal   
con múltiples variaciones independientes**

**Ejemplo 93.a)-1: Representación de una variante de secuencia mediante anotación de la secuencia principal** [**71**](#page713)

**Párrafo 93.b) – Variante de secuencia divulgada únicamente mediante referencia a una secuencia principal   
con múltiples variaciones interdependientes**

**Ejemplo 93.b)-1: Representación de distintas variantes de secuencias con múltiples variaciones   
interdependientes** [**72**](#page72)

**Párrafo 94 – Claves de caracterización y calificadores para una variante de secuencia**

***Referencia a otros ejemplos***

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

**Párrafo 95– Anotación de una variante de secuencia**

***Referencia a otros ejemplos***

Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other” [54](#page54)

Ejemplo 91-3: Representación de una secuencia de consenso [69](#page69)

Ejemplo 92-1: Representación de una secuencia única con aminoácidos alternativos enumerados [70](#page70)

**Ejemplos**

**Párrafo 3.a) Definición de “aminoácido”**

**Ejemplo 3.a)-1: D-aminoácidos**

En una solicitud de patente se describe la secuencia siguiente:

Cyclo (D-Ala-D-Glu-Lys-Nle-Gly-D-Met-D-Nle)

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la definición del párrafo 3.a) de la Norma se establece que “aminoácido” incluye los “D-aminoácidos” y los aminoácidos que contienen cadenas laterales modificadas o sintéticas. A la luz de esa definición, el péptido enumerado contiene cinco aminoácidos que están específicamente definidos (D-Ala, D-Glu, Lys, Gly y D‑Met). Por lo tanto, la secuencia debe incluirse en una lista de secuencias según se exige en el párrafo 7.b) de la Norma ST.26.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

En el párrafo 29 se exige que los D-aminoácidos estén representados en la secuencia como el L-aminoácido correspondiente no modificado. Además, todo aminoácido modificado que no pueda ser representado mediante ningún otro símbolo de los que figuran en el Cuadro 3 de la Sección 3 del Anexo I debe ser representado mediante el símbolo “X”.

En este ejemplo, la secuencia contiene tres D-aminoácidos que pueden ser representados por un L-aminoácido no modificado del Cuadro 3 de la Sección 3 del Anexo I, un L-aminoácido (Nle) y un D-aminoácido (D-Nle) que deben ser representados por el símbolo “X”.

En el párrafo 25 se indica que cuando las secuencias de aminoácidos tengan una configuración circular y el anillo esté compuesto solo por residuos de aminoácidos unidos por enlaces peptídicos, el solicitante deberá elegir el aminoácido del residuo de la posición número 1. Por lo tanto, la secuencia puede representarse como:

AEKXGMX (SEQ ID NO: 1)

o, de otra forma, con cualquier otro aminoácido en la secuencia del residuo de la posición número 1. Para cada D‑aminoácido deberá proporcionarse una clave de caracterización “SITE” y un calificador “NOTE” con el nombre completo, no abreviado, del D-aminoácido como valor calificador, por ejemplo, D-alanina y D-norleucina. Además, deberá proporcionarse una clave de caracterización “SITE” y un calificador “NOTE” con la abreviatura de L‑norleucina como valor calificador, es decir, “Nle”, según figura en el Cuadro 4 de la Sección 4 del Anexo I. Por último, deberá proporcionarse una clave de caracterización “REGION” y un calificador “NOTE” para indicar que el péptido es circular.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.a)**, 7.b), 25, 26, 29, 30 y 31.

**Párrafo 3.c) – Definición de “enumeración de sus residuos”**

**Ejemplo 3.c)-1: Enumeración de aminoácidos mediante la estructura química**



**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El péptido enumerado, ilustrado como estructura, contiene como mínimo cuatro aminoácidos específicamente definidos. Por lo tanto, la secuencia deberá incluirse en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia puede ser representada como:

VAFXGK (SEQ ID NO: 2)

donde “X” representa un aminoácido modificado “other”, que exige una clave de caracterización “SITE” junto con el calificador “NOTE”. El calificador “NOTE” proporciona el nombre completo, no abreviado, del triptófano modificado en la posición 4 del péptido enumerado, por ejemplo, “ácido 6-amino-7-(1H-indol-3-yl)-5- oxo-heptanóico”. Además, se necesitan claves de caracterización adicionales “SITE” y el calificador “NOTE” para indicar la acetilación del N-terminal y la metilación del C-terminal.

Como alternativa, la secuencia puede representarse como:

VAFW (SEQ ID NO: 3)

Se necesita una clave de caracterización “SITE” y un calificador “NOTE” para indicar la modificación del triptófano en la posición 4 del péptido enumerado, con el valor: “C-terminal unido por medio de un puente de glutaraldehído al dipéptido GK”. Además, se necesita una clave de caracterización adicional “SITE” en la localización 1 y el calificador “NOTE” para indicar la acetilación del N-terminal.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.c)**, 7.b), 29, 30 y 31.

**Ejemplo 3.c)-2: Fórmula abreviada de una secuencia de aminoácidos**

(G4z)n

donde G= Glicina, z = cualquier aminoácido y la variable n puede ser cualquier número entero.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**Sí**

En la divulgación se indica que “n” puede ser “cualquier número entero”; por lo tanto, la realización más abarcadora de “n” es indeterminada. Puesto que “n” es indeterminado, el péptido de la fórmula no puede ser expandido a una longitud definida y, por lo tanto, deberá considerarse la fórmula no expandida.

El péptido enumerado en la formula no expandida (“n” = 1) provee cuatro aminoácidos específicamente definidos, cada uno de los cuales es Gly, y el símbolo “z”. Convencionalmente, “Z” es el símbolo correspondiente a “glutamina o ácido glutámico”; sin embargo, en el ejemplo se define “z” como “cualquier aminoácido”. En la Norma ST.26, un aminoácido que no está específicamente definido se representa mediante “X”. Sobre la base de ese análisis, el péptido enumerado, es decir, GGGGX, contiene cuatro residuos glicina que están enumerados y específicamente definidos. Así pues, el párrafo 7.b) de la Norma ST.26 exige la inclusión de la secuencia en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia utiliza un símbolo no convencional “z”, cuya definición debe quedar determinada por la divulgación (*véase* la Introducción del presente documento). Puesto que “z” está definido como cualquier aminoácido, el símbolo convencional utilizado para representar ese aminoácido es “X.” Por lo tanto, la secuencia deberá representarse como una secuencia única:

GGGGX (SEQ ID NO: 4)

de preferencia anotada con la clave de caracterización REGION, la localización de característica “&gt;5” (corresponde a >5), con un calificador NOTE con el valor “La totalidad de la secuencia de aminoácidos 1-5 puede repetirse una vez o más.”

Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" deberá anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de PI podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.c)**, ~~y~~ 7.b) y 27.

### *Párrafo 3.g) Definición de “nucleótido”*

**Ejemplo 3.g)-1: Secuencia de nucleótidos interrumpida por un espaciador C3**

En una solicitud de patente se describe la secuencia siguiente:

atgcatgcatgcncggcatgcatgc

donde n = un espaciador C3 con la estructura siguiente:



**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

La secuencia enumerada contiene dos segmentos de nucleótidos específicamente definidos separados por un espaciador C3.

El espaciador C3 no es un nucleótido conforme al párrafo 3.g); el símbolo convencional “n” se utiliza de manera no convencional (*véase* la Introducción del presente documento). En consecuencia, cada segmento es una secuencia de nucleótidos distinta. Puesto que cada segmento contiene más de 10 nucleótidos específicamente definidos, ambos deben incluirse en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Cada segmento debe incluirse en una lista de secuencias como una secuencia distinta, cada una con su propio identificador de secuencia:

atgcatgcatgc (SEQ ID NO: 5)

cggcatgcatgc (SEQ ID NO: 6)

La citosina en cada segmento que está unida al espaciador C3 debería describirse en más detalle en un cuadro de características utilizando la clave de caracterización “misc\_feature” y el calificador “note”. El valor calificador “note”, que es “texto libre”, debería indicar la presencia del espaciador, que está unido a otro ácido nucleico e identificar el espaciador, ya sea mediante su nombre químico completo, no abreviado, ya sea mediante su nombre común, es decir, espaciador C3.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.g)**, 7.a) y 15.

**Ejemplo 3.g)-2: Secuencia de nucleótidos con residuos alternativos, incluido un espaciador C3**

En una solicitud de patente se describe la secuencia siguiente:

atgcatgcatgcncggcatgcatgc

donde n = c, a, g o un espaciador C3 con la estructura siguiente:



**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Hay 24 residuos específicamente definidos en la secuencia enumerada interrumpida por la variable “n.” Deberá consultarse la explicación de la secuencia en la divulgación para determinar si “n” se utiliza de manera convencional o no convencional (*véase* la Introducción del presente documento).

En la divulgación se indica que n = c, a, g o un espaciador C3. El símbolo “n” es un símbolo convencional utilizado de manera no convencional, puesto que se describe que puede incluir un espaciador C3, que no responde a la definición de nucleótido. También se describe que el símbolo “n” puede incluir “c”, “a” o “g”; por lo tanto, la Norma ST.26 exige la inclusión de la secuencia de 25 nucleótidos en una lista de secuencias. Puesto que los dos segmentos separados por el espaciador C3 son secuencias distintas de la secuencia de 25 nucleótidos, también pueden incluirse las dos secuencias de 12 nucleótidos.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

En el ejemplo se indica que “n = c, a, g o un espaciador C3”. Como se ha señalado más arriba, un espaciador C3 no es un nucleótido. Con arreglo al párrafo 15, el símbolo “n” no debe utilizarse para representar otra cosa que un nucleótido; por lo tanto, el símbolo “n” no puede representar un espaciador C3 en una lista de secuencias.

En el párrafo 15 también se establece que si correspondiera utilizar un símbolo de ambigüedad, deberá utilizarse el símbolo más restrictivo. El símbolo “v” representa “a or c o g” conforme al Cuadro 1 de la Sección 1 del Anexo I, y es más restrictivo que “n”.

Cuando la variable “n” en el ejemplo es c, a, o g, la secuencia única enumerada mediante sus residuos que incluye las realizaciones más divulgadas y, por lo tanto, es la secuencia más abarcadora (*véase* la Introducción del presente documento) que debe incluirse en una lista de secuencias, es:

atgcatgcatgcvcggcatgcatgc (SEQ ID NO: 7)

Se alienta encarecidamente a incluir cualesquiera secuencias adicionales indispensables para la divulgación o las reivindicaciones de la invención, como se ha señalado en la Introducción del presente documento.

Cuando la variable “n” en el ejemplo es un espaciador C3, la secuencia pude considerarse como dos segmentos distintos de nucleótidos específicamente definidos a cada lado de la variable “n”, es decir, atgcatgcatgc (SEQ ID NO: 8); y cggcatgcatgc (SEQ ID NO: 9). Si fuera indispensable para la divulgación o las reivindicaciones, esas dos secuencias deberían incluirse también en la lista de secuencias, cada una con su propio identificador de secuencia.

La citosina en cada segmento que está unida al espaciador C3 debería describirse en más detalle en un cuadro de características utilizando la clave de caracterización “misc\_feature” y el calificador “note”. El valor calificador “note”, que es “texto libre”, debería indicar la presencia del espaciador, que está unido a otro ácido nucleico e identificar el espaciador ya sea mediante su nombre químico completo, no abreviado, ya sea mediante su nombre común, es decir, espaciador C3.

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de PI podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.g)**, 7.a) y 15.

**Ejemplo 3.g)-3: Sitio abásico**

En una solicitud de patente se describe la secuencia siguiente:

gagcattgac-AP-taaggct

donde AP es un sitio abásico

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Los residuos específicamente definidos de la secuencia enumerada están interrumpidos por un sitio abásico. El lado 5’ del sitio abásico contiene 10 nucleótidos y el lado 3’ del sitio abásico contiene 7 nucleótidos. En el párrafo 3.g)ii)2) se define un sitio abásico como un “nucleótido” cuando forma parte de una secuencia de nucleótidos. En consecuencia, el sitio abásico de este ejemplo se considera como un “nucleótido” a los fines de determinar si se exige incluir la secuencia en una lista de secuencias, y de qué forma. Por lo tanto, los residuos a cada lado del sitio abásico forman parte de una secuencia única enumerada que contiene un total de 18 nucleótidos, 17 de los cuales son específicamente definidos. Por consiguiente, la secuencia debe ser incluida como secuencia única en una lista de secuencias, según lo exige el párrafo 7.a~~b~~) de la Norma ST.26.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia debe incluirse en una lista de secuencias como:

gagcattgacntaaggct (SEQ ID NO: 10)

El sitio abásico debe representarse mediante el símbolo “n” y debe describirse en más detalle en un cuadro de características. El medio preferido de anotación está dado por la clave de caracterización “modified\_base” y el calificador obligatorio “mod\_base” con el valor “OTHER”. Debe incluirse un calificador “note” que describa la base modificada como sitio abásico.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.g)**, 7.a) y 17.

**Ejemplo 3.g)-4: Análogos de ácido nucleico**

En una solicitud de patente se divulga la siguiente secuencia de ácido nucleico glicólico (ANG):

PO4-tagttcattgactaaggctccccattgact-OH

donde el extremo final de la secuencia imita el extremo 5’ de una secuencia de ADN.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ** – Los distintos residuos que constituyen una secuencia de ANG se consideran nucleótidos con arreglo al párrafo 3.g)i)2) de la Norma ST.26. Es decir que la secuencia tiene más de diez nucleótidos enumerados y “específicamente definidos” y por ello se exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Las secuencias de ANG no tienen un extremo 5’ y un extremo 3’, sino un extremo 3’ y un extremo 2’. El extremo 3’, que suele ilustrarse con un grupo fosfato terminal, corresponde al extremo 5’ del ADN o el ARN. (Cabe señalar que otros análogos de ácido nucleico pueden corresponder de manera diferente a los extremos 5’ y 3’ del ADN y el ARN). Conforme al párrafo ~~10~~11, debe incluirse en una lista de secuencias “en el sentido de izquierda a derecha que imite el sentido 5’ a 3”. Por lo tanto, debe incluirse en una lista de secuencias como:

tagttcattgactaaggctccccattgact (SEQ ID NO: 11)

La secuencia debe describirse en un cuadro de características utilizando la clave de caracterización “modified\_base” y el calificador obligatorio “mod\_base” con la abreviatura “OTHER”. Debe incluirse un calificador “note” con el nombre completo, no abreviado, de los nucleótidos modificados, como “ácidos nucleicos glicólicos” o “2,3- dihidroxipropil nucleósidos”. Un único elemento INSDFeature puede utilizarse para describir toda la secuencia como un ANG en la que el rango de la INSDFeature\_location es “1..30”.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~3.d), **3.g)**, 7.a), 11, 16, 18, 65 y 66.

### *Párrafo 3.k) Definición de “específicamente definido”*

**Ejemplo 3.k)-1: Símbolos de ambigüedad de nucleótidos**

5’ NNG KNG KNG K 3’

N y K son códigos de ambigüedad IUPAC-IUB

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

Los códigos de ambigüedad IUPAC-IUB corresponden a la lista de símbolos de nucleótidos que figuran en el Cuadro 1 de la Sección 1 del Anexo I. Con arreglo al párrafo 3.k), un nucleótido específicamente definido es todo nucleótido distinto a los representados por el símbolo “n” en el Anexo I. Por lo tanto, “K” y “G” son nucleótidos específicamente definidos y “N” no es un nucleótido específicamente definido.

La secuencia enumerada no tiene diez o más nucleótidos específicamente definidos y, por lo tanto, el párrafo 7.a) de la Norma ST.26 no exige que se incluya en una lista de secuencias.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

Con arreglo al párrafo 8, “Una lista de secuencias no deberá incluir ninguna secuencia que tenga menos de diez nucleótidos específicamente definidos […]”. La secuencia enumerada no tiene diez o más nucleótidos específicamente definidos; por lo tanto, no deberá incluirse en una lista de secuencias.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.k)**, 7.a), 8 y 13.

**Ejemplo 3.k)-2: Símbolo de ambigüedad “n” utilizado de manera tanto convencional como no convencional**

En una solicitud se divulga la secuencia artificial: 5’-AATGCCGGAN-3’. En la divulgación se establece además que:

i) en una realización, N es cualquier nucleótido;

ii) en una realización, N es opcional, pero, de preferencia, es G;

iii) en una realización, N es K;

iv) en una realización, N es C.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

La secuencia enumerada contiene 9 nucleótidos específicamente definidos y un símbolo “N.” Debe consultarse la explicación de la secuencia en la divulgación para determinar si el símbolo “N” se utiliza de manera convencional (*véase* la Introducción del presente documento).

Del examen de las realizaciones divulgadas i) a iv) de la secuencia enumerada se desprende que la realización más abarcadora del símbolo “N” es “cualquier nucleótido”. En la realización más abarcadora, “N” en la secuencia enumerada se utiliza de manera convencional.

En algunas realizaciones, se indica que “N” es un residuo específicamente definido (es decir, “N es C” en la parte iv)). Sin embargo, solo se considera la realización más abarcadora (es decir, “N es cualquier nucleótido”) para determinar si una secuencia debe incluirse en una lista de secuencias. Así pues, la secuencia enumerada que debe evaluarse es 5’-AATGCCGGAN-3’.

Sobre la base de ese análisis, la secuencia enumerada, es decir, AATGCCGGAN, no contiene diez nucleótidos específicamente definidos. Por lo tanto, el párrafo 7.a) de la Norma ST.26 no exige la inclusión de la secuencia en una lista de secuencias, pese al hecho de que “n” también se define como nucleótido específico en algunas realizaciones.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

La secuencia “AATGCCGGAN” no debe incluirse en una lista de secuencias.

Sin embargo, una secuencia alternativa descrita puede incluirse en una lista de secuencias si el símbolo “N” se sustituye por un nucleótido específicamente definido.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Se alienta **encarecidamente** a que se incluyan secuencias que representan realizaciones que constituyen una parte fundamental de la invención. La inclusión de esas secuencias permite una búsqueda más minuciosa y ofrece información al público acerca de la materia respecto de la cual se solicita una patente.

En el caso del ejemplo que figura más arriba, se recomienda encarecidamente que las tres secuencias adicionales siguientes se incluyan en la lista de secuencias, cada una con su propio identificador:

aatgccggag (SEQ ID NO: 12)

aatgccggak (SEQ ID NO: 13)

aatgccggac (SEQ ID NO: 14)

Si no se incluyen las tres secuencias que figuran más arriba, ni una menos, el nucleótido que sustituye al símbolo “n” debería anotarse para describir las alternativas. Por ejemplo, si de ellas solo se incluye en la lista de secuencias SEQ ID NO: 12, debería utilizarse la clave de caracterización “misc\_difference” con la localización de característica “10”, junto con dos calificadores “replace” en los que el valor de uno sería “g” y el del otro sería “c”.

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de P.I. podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.k)**, 7.a), 8 y 13.

**Ejemplo 3.k)-3: Símbolo de ambigüedad “n” utilizado de manera no convencional**

En una solicitud se divulga la secuencia: 5’-aatgttggan-3’

donde n es c

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Con arreglo al párrafo 3.k), un nucleótido “específicamente definido” es todo nucleótido distinto a los representados por el símbolo “n” en el Cuadro I de la Sección I del Anexo I.

En este ejemplo, “n” se utiliza de manera no convencional para representar únicamente “c”. La divulgación no indica que “n” se utilice de manera convencional para representar “todo nucleótido”. Por lo tanto, la secuencia debe interpretarse como si se hubiera utilizado en ella el símbolo convencional equivalente, es decir, “c” (*véase* la Introducción del presente documento). Por consiguiente, la secuencia enumerada que debe considerarse es:

5’-aatgttggac-3’

Esta secuencia tiene diez nucleótidos específicamente definidos y el párrafo 7.a) de la Norma ST.26 exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia debe incluirse en una lista de secuencias como: aatgttggac (SEQ ID NO: 15)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.k)** y 7.a).

**Ejemplo 3.k)-4: Los símbolos de ambigüedad distintos de “n” son “específicamente definidos”**

En una solicitud de patente se describe la secuencia siguiente:

5’ NNG KNG KNG KAG VCR 3’

donde N, K, V y R son códigos de ambigüedad IUPAC-IUB

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Los códigos de ambigüedad IUPAC-IUB corresponden a la lista de símbolos de nucleótidos definidos en el Cuadro 1 de la Sección 1 del Anexo I. Con arreglo al párrafo 3.k), un nucleótido “específicamente definido” es todo nucleótido distinto a los representados por el símbolo “n” en el Cuadro 1 de la Sección 1 del Anexo I. Por lo tanto, “K”, “V” y “R” son “nucleótidos específicamente definidos.

La secuencia tiene 11 nucleótidos enumerados y “específicamente definidos” y el párrafo 7.a) de la Norma ST.26 exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia debe incluirse en una lista de secuencias como:

nngkngkngkagvcr (SEQ ID NO: 16)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.k)**, 7.a) y 15.

**Ejemplo 3.k)-5: Abreviatura de ambigüedad “Xaa” utilizada de manera no convencional**

En una solicitud de patente se describe la secuencia siguiente:

Xaa-Tyr-Glu-Xaa-Xaa-Xaa-Leu

donde Xaa en la posición 1 es cualquier aminoácido, Xaa en la posición 4 es Lys, Xaa en la posición 5 es Gly y Xaa en la posición 6 es leucina o isoleucina.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El péptido enumerado en la fórmula tiene tres aminoácidos específicamente definidos en las posiciones 2, 3 y 7. El primer aminoácido está representado por una abreviatura convencional, es decir, Xaa, que representa cualquier aminoácido. Sin embargo, el 4º, el 5º y el 6º aminoácido están representados por una abreviatura convencional utilizada de manera no convencional (*véase* la Introducción del presente documento). Por lo tanto, se consulta la explicación de la secuencia en la divulgación para determinar la definición de “Xaa” en esas posiciones. Puesto que “Xaa” en las posiciones 4 a 6 está indicado como un aminoácido específico, debe interpretarse la secuencia como si en ella se hubieran utilizado las abreviaturas convencionales equivalentes, es decir, Lys, Gly y (Leu o Ile). Por consiguiente, la secuencia contiene cuatro o más aminoácidos específicamente definidos y debe incluirse en una lista de secuencias, según lo exige el párrafo 7.b) de la Norma ST.26.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

En la secuencia se utiliza la abreviatura convencional “Xaa” de manera no convencional. Por lo tanto, debe consultarse la explicación de la secuencia en la divulgación para determinar la definición de “Xaa” en las posiciones 4, 5 y 6. La explicación define “Xaa” como una lisina en la posición 4, una glicina en la posición 5 y una leucina o isoleucina en la posición 6. Los símbolos convencionales para esos aminoácidos son K, G y J respectivamente. Por lo tanto, la secuencia debería representarse en la lista de secuencias como:

XYEKGJL (SEQ ID NO: 17)

Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**3.k)**, 7.b), 26 y 27.

### *Párrafo 7.a) – Secuencia de nucleótidos exigida en una lista de secuencias*

**Ejemplo 7.a)-1: Secuencia de nucleótidos ramificada**

En la descripción se divulga la siguiente secuencia de nucleótidos ramificada:



donde "pnp" es un enlace o monómero que contiene una funcionalidad del bromoacetilamino;

3’-CA(pnp)CACACA(pnp)CACACA(pnp)CACACACA-(5’)NH—C(=O)CH2 3’ es el segmento A;

SP(O-)(=O)CACACAAAAAAAAAAAAAAAAAAAAAAAAA 3’ es los segmentos B, C y D; y

SP(O-)(=O)CACATAGGCATCTCCTAGTGCAGGAAGA 3’ es el segmento E.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ** – los cuatro segmentos verticales B a E deben incluirse en una lista de secuencias

**NO** – el segmento horizontal A no debe incluirse en una lista de secuencias

La figura que se reproduce más arriba es un ejemplo de una secuencia de ácido nucleico ramificada de “tipo peine” que contiene cinco segmentos lineales: el segmento horizontal A y los cuatro segmentos verticales B a E.

Con arreglo al párrafo 7.a), las ~~porciones~~regiones lineales de secuencias de nucleótidos ramificadas que contienen diez o más nucleótidos específicamente definidos, cuando los nucleótidos adyacentes están unidos en el sentido de 3’ a 5’, deben incluirse en una lista de secuencias.

Los cuatro segmentos verticales B a E contienen, cada uno, más de diez nucleótidos específicamente definidos, y en ellos los nucleótidos adyacentes están unidos en el sentido de 3’ a 5’; por lo tanto se exige que cada uno de ellos se incluya en una lista de secuencias.

En el segmento horizontal A, las ~~porciones~~regiones lineales de la secuencia de nucleótidos están enlazadas por la fracción no nucleótido “pnp” y cada una de esas ~~porciones~~regiones lineales enlazadas contiene menos de diez nucleótidos específicamente definidos. Por lo tanto, puesto que ninguna ~~porción~~región del segmento A contiene diez o más nucleótidos específicamente definidos y ningún nucleótido adyacente está unido en el sentido de 3’ a 5’, el párrafo 7.a) de la Norma ST.26 no exige que se incluyan en una lista de secuencias.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

Con arreglo al párrafo 8, “una lista de secuencias no deberá incluir ninguna secuencia que tenga menos de diez nucleótidos específicamente definidos […]”.

Ninguna ~~porción~~región del segmento A contiene diez o más nucleótidos específicamente definidos con nucleótidos adyacentes unidos en el sentido de 3’ a 5’; por lo tanto, no debe incluirse en una lista de secuencias como secuencia distinta con su propio identificador de secuencia.

Sin embargo, los segmentos B, C, D y E pueden anotarse indicando que están enlazados al segmento A.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Los segmentos B, C y D son idénticos y deben incluirse en una lista de secuencias como secuencia única:

cacacaaaaaaaaaaaaaaaaaaaaaaaaa. (SEQ ID NO: 18)

La primera “c” de la secuencia debería describirse en más detalle como un nucleótido modificado utilizando la clave de caracterización “misc\_feature” y el calificador “note” con el valor, por ejemplo, “Esta secuencia es una de cuatro ramas de un polinucleótido ramificado.”.

El segmento E debe incluirse en una lista de secuencias como secuencia única:

cacataggcatctcctagtgcaggaaga. (SEQ ID NO: 19)

La primera “c” de la secuencia debería describirse en más detalle como un nucleótido modificado utilizando la clave de caracterización “misc\_feature” y el calificador “note” con el valor, por ejemplo, “Esta secuencia es una de cuatro ramas de un polinucleótido ramificado.”

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.a)**, 8, 11, 13 y 17.

**Ejemplo 7.a)-2: Secuencia de nucleótidos lineal con una estructura secundaria**

En una solicitud de patente se describe la secuencia siguiente:



donde Ψ es pseudouridina.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

La secuencia de nucleótidos contiene 73 nucleótidos enumerados y específicamente definidos. Así pues, en el ejemplo hay 10 o más nucleótidos “específicamente definidos” y, según exige el párrafo 7)a) de la Norma ST.26, debe incluirse en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Al consultar la divulgación se observa que “Ψ” es equivalente a pseudouridina. El único símbolo convencional que puede utilizarse para representar la pseudouridina es “n”; por lo tanto, “Ψ” es un símbolo no convencional utilizado para representar el símbolo convencional “n” (*véase* la Introducción del presente documento). Por consiguiente, debe interpretarse que la secuencia tiene dos símbolos “n”, en lugar de los dos símbolos “Ψ”.

El símbolo “u” no debe utilizarse para representar el uracilo en una molécula de ARN en la lista de secuencias. Con arreglo al párrafo 14, el símbolo “t” se interpretará como uracilo en ARN. La secuencia debe incluirse como:

gcggatttagctcagctgggagagcgccagactgaatanctggagtcctgtgtncgatccacagaattcgcacca (SEQ ID NO: 20)

El valor del calificador obligatorio “mol\_type” de la clave de caracterización obligatoria “source” es “tRNA”. Puede proporcionarse información adicional con la clave de caracterización “tRNA” y cualquier calificador o calificadores apropiados.

Los residuos “n” deben describirse en más detalle en un cuadro de características utilizando la clave de caracterización “modified\_base” y el calificador obligatorio “mod\_base” con la abreviatura “p” para la pseudouridina como valor calificador (véase el Cuadro 2 del Anexo 1).

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.a)**, 11, 13, 14, 17, 62, 84 y secciones 2 y 5 del Anexo I, clave de caracterización 5.43.

**Ejemplo 7.a)-3: Símbolos de ambigüedad de nucleótidos utilizados de manera no convencional**

En una solicitud de patente se describe la secuencia siguiente:

5’ GATC-MDR-MDR-MDR-MDR-GTAC 3’

La explicación de la secuencia en la divulgación indica asimismo: “Un “DR Element” consiste en la secuencia 5’ ATCAGCCAT 3’. Un DR Element mutante, o MDR, es un elemento DR en el que, en el medio, 5 nucleótidos, CAGCC, han mutado a TTTTT.”

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la secuencia enumerada se utiliza el símbolo “MDR”. Cuando no queda claro si al utilizar un símbolo en una secuencia se pretende que sea un símbolo convencional, es decir, un símbolo que figura en el Cuadro 3 de la Sección 3 del Anexo 1, o un símbolo no convencional, debe consultarse la explicación de la secuencia en la divulgación para tomar una decisión a ese respecto (*véase* la Introducción del presente documento). Con arreglo al Cuadro 3, “MDR” podría interpretarse como tres símbolos convencionales (m = a o c, d = a o g o t/u, r = g o a) o como abreviatura que equivale a la notación abreviada para alguna otra estructura.

Al consultar la divulgación se observa que un elemento MDR es equivalente a 5’ ATTTTTTAT 3’. Se considera que las letras “MDR” son símbolos convencionales utilizados de manera no convencional; por lo tanto, debe interpretarse la secuencia como si hubiera sido divulgada utilizando los símbolos convencionales equivalentes. Por consiguiente, la secuencia enumerada que se considera para su inclusión en una lista de secuencias es:

5’ GATC ATTTTTTAT ATTTTTTAT ATTTTTTAT ATTTTTTAT GTAC 3’

La secuencia enumerada tiene 44 nucleótidos específicamente definidos y el párrafo 7.a) de la Norma ST.26 exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia debe incluirse en una lista de secuencias como:

gatcattttttatattttttatattttttatattttttatgtac (SEQ ID NO: 21)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.a)** y 13**.**

**Ejemplo 7.a)-4: Símbolos de ambigüedad de nucleótidos utilizados de manera no convencional**

En una solicitud de patente se describe la secuencia siguiente:

5’ ATTC-N-N-N-N-GTAC 3’

La explicación de la secuencia en la divulgación indica asimismo que “N” consiste en la secuencia 5’ ATACGCACT 3’.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la secuencia enumerada se utiliza el símbolo “N”. Debe consultarse la explicación de la secuencia en la divulgación para determinar si “N” se utiliza de manera convencional o no convencional (*véase* la Introducción del presente documento).

Al consultar la divulgación se observa que “N” es equivalente a 5’ ATACGCACT 3’. Así pues, “N” es un símbolo convencional utilizado de manera no convencional. Por consiguiente, la secuencia debe interpretarse como si se hubiera divulgado utilizando los símbolos convencionales equivalentes:

5’ ATTC-ATACGCACT-ATACGCACT-ATACGCACT-ATACGCACT-GTAC 3’

La secuencia enumerada tiene 44 nucleótidos específicamente definidos y el párrafo 7.a) de la Norma ST.26 exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia debe incluirse en una lista de secuencias como:

attcatacgcactatacgcactatacgcactatacgcactgtac (SEQ ID NO: 22)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.a)** y 13.

**Ejemplo 7.a)-5: Símbolos no convencionales de nucleótidos**

En una solicitud de patente se describe la secuencia siguiente:

5’ GATC-β-β-β-β-GTAC 3’

La explicación de la secuencia en la divulgación indica asimismo que “β” consiste en la secuencia 5’ ATACGCACT 3’.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la secuencia enumerada se utiliza el símbolo no convencional “β”. Debe consultarse la explicación de la secuencia en la divulgación para determinar el significado de “β” (*véase* la Introducción del presente documento).

Al consultar la divulgación se observa que “β” es equivalente a 5’ ATACGCACT 3’. Así pues, “β” es un símbolo no convencional utilizado para representar una secuencia de nueve símbolos convencionales específicamente definidos. Por consiguiente, la secuencia debe interpretarse como si se hubiera divulgado utilizando los símbolos convencionales equivalentes:

5’ GATC-ATACGCACT-ATACGCACT-ATACGCACT-ATACGCACT-GTAC 3’

La secuencia enumerada tiene 44 nucleótidos específicamente definidos y el párrafo 7.a) de la Norma ST.26 exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia debe incluirse en una lista de secuencias como:

gatcatacgcactatacgcactatacgcactatacgcactgtac (SEQ ID NO: 23)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.a)** y 13.

**Ejemplo 7.a)-6: Símbolos no convencionales de nucleótidos**

En una solicitud de patente se describe la secuencia siguiente:

5’ GATC-β-β-β-β-GTAC 3’

La explicación de la secuencia en la divulgación indica asimismo que “β” es igual a adenina, inosina, o pseudouridina.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

En la secuencia enumerada se utiliza el símbolo no convencional “β”. Debe consultarse la explicación de la secuencia en la divulgación para determinar el significado de “β” (*véase* la Introducción del presente documento).

Al consultar la divulgación se observa que “β” es equivalente a adenina, inosina, o pseudouridina. El único símbolo convencional que puede utilizarse para representar “adenina, inosina, o pseudouridina” es “n”; por lo tanto, “β” es un símbolo no convencional utilizado para representar el símbolo convencional “n”. Por consiguiente, debe interpretarse que la secuencia tiene cuatro símbolos “n” en lugar de los cuatro símbolos “β”:

5’ GATC-N-N-N-N-GTAC 3’

La secuencia enumerada tiene solo ocho nucleótidos específicamente definidos y el párrafo 7.a) de la Norma ST.26 no exige su inclusión en una lista de secuencias.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

La secuencia enumerada, 5’ GATC-N-N-N-N-GTAC 3’ no debe incluirse en una lista de secuencias.

Sin embargo, una secuencia alternativa divulgada puede incluirse en una lista de secuencias si al menos 2 de los símbolos “n” son sustituidos por adenina, dando como resultado una secuencia con al menos 10 o más nucleótidos específicamente definidos.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Una posible representación permitida es:

gatcaaaagtac (SEQ ID NO: 24)

En el ejemplo que figura más arriba, los cuatro nucleótidos de adenina que reemplazan los símbolos β deberían anotarse para señalar que esas posiciones podrían ser sustituidas por inosina o pseudouridina.

Debería utilizarse la clave de caracterización “misc\_difference” con una localización de característica 5-8 y un calificador “note” con el valor, por ejemplo, “Un nucleótido en cualquiera de las posiciones entre 5 y 8 puede sustituirse por inosina o pseudouridina”. Puesto que esas alternativas son nucleótidos modificados, será necesario utilizar la clave de caracterización “modified\_base” junto con el calificador “mod\_base”. El valor para el calificador “mod\_base” puede ser “OTHER” junto con un calificador “note” y el valor “i o p”.

Es posible realizar otras sustituciones.

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de P.I. podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.a)**, 8, 13 y 17.

### *Párrafo 7.b) – Secuencias de aminoacidos exigidas en una lista de secuencias*

**Ejemplo 7.b)-1: Cuatro o más aminoácidos específicamente definidos**

XXXXXXXXDXXXXXXXXXXFXXXXXXXXXXXXXXXXXXXXXXXXXXXXAXXXXXXXXXXXXXXXXXXXGXXXXX

donde X = cualquier aminoácido

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El péptido enumerado contiene cuatro aminoácidos específicamente definidos. El símbolo “X” se utiliza de manera convencional para representar los aminoácidos restantes como cualquier aminoácido (*véase* la Introducción del presente documento).

Habida cuenta de que hay cuatro aminoácidos específicamente definidos, es decir, Asp, Phe, Ala y Gly, el párrafo 7.b) de la Norma ST.26 exige que la secuencia se incluya en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia debe representarse como:

XXXXXXXXDXXXXXXXXXXFXXXXXXXXXXXXXXXXXXXXXXXXXXXXAXXXXXXXXXXXXXXXXXXXGXXXXX (SEQ ID NO: 25)

~~Puesto que “X” puede ser cualquier aminoácido, en virtud del párrafo 27 no se exige la anotación de los residuos “X”.~~Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.b)**, 8 y 27.

**Ejemplo 7.b)-2: Secuencia de aminoácidos ramificada**

En la solicitud se describe una secuencia ramificada en la cual los residuos lisina se utilizan como andamiaje para formar ocho ramas a las que están unidas múltiples cadenas lineales de péptidos. La lisina es un aminoácido dibásico, que provee dos sitios para la unión de péptidos. El péptido se ilustra de la manera siguiente:



En el péptido ramificado que figura más arriba, las uniones entre la lisina y otro aminoácido indicadas mediante representan un enlace amida entre la terminal amina de la lisina y el extremo carboxilo del aminoácido unido. Las uniones indicadas mediante representan un enlace amida entre el amino en la cadena lateral de la lisina y el extremo carboxilo del aminoácido unido.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En el ejemplo se divulga una secuencia ramificada en la que los residuos lisina se utilizan como andamiaje. El párrafo 7.b) exige que la ~~porción~~región no ramificada o lineal de la secuencia, que contiene cuatro o más aminoácidos específicamente definidos, se incluya en una lista de secuencias. Las ~~porciones~~regiones lineales del péptido ramificado en el ejemplo que figura más arriba, que tienen cuatro o más aminoácidos específicamente definidos, figuran más abajo rodeadas por un círculo:



El párrafo 7.b) de la Norma ST.26 exige la inclusión de los péptidos 1 a 6, más arriba, en una lista de secuencias.

Los péptidos cuya inclusión en la lista de secuencias no se exige ~~y, de hecho, está prohibida~~ son los siguientes:

YFA

LLK

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

Con arreglo al apartado 8, una lista de secuencias no deberá incluir ninguna secuencia que tenga menos de cuatro aminoácidos específicamente definidos.

Los péptidos YFA y LLK contienen, cada uno, solo tres aminoácidos específicamente definidos y, por lo tanto, no deberán incluirse en una lista de secuencias como secuencias distintas, cada una con su propio número de identificación.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Los péptidos 1 a 6 deben representarse con identificadores de secuencia distintos:

RISL (SEQ ID NO: 26)

LLKK (SEQ ID NO: 27)

IPACTA (SEQ ID NO: 28)

FRAGGK (SEQ ID NO: 29)

HQYFA (SEQ ID NO: 30)

ATFGKKKA (SEQ ID NO: 31)

El enlace cruzado se anota, de preferencia, utilizando la clave de caracterización “SITE” y el calificador obligatorio “NOTE” con el valor, por ejemplo, “Esta secuencia es una parte de ~~un péptido ramificado~~una secuencia de aminoácidos ramificada”. Con arreglo al párrafo 29 de la norma ST.26, las secuencias SEQ ID NO: 27, 29 y 31 deben incluir una anotación por cada lisina para indicar que es un aminoácido modificado, utilizando la clave de caracterización "SITE" junto con el calificativo "NOTE" que describa que la cadena lateral de la lisina está unida por medio de un enlace amida a otra secuencia. Preferiblemente, cada una de las secuencias SEQ ID NO: 26, 28 y 30 debería incluir una anotación para indicar que el aminoácido C-terminal está unido a otra secuencia, utilizando la clave de caracterización "SITE" junto con el calificativo "NOTE".

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.b)**, 8, 26, 29, 30 y 31.

**Ejemplo 7.b)-3: Secuencia de aminoácidos ramificada**

Péptido de la secuencia siguiente:



El enlace entre el residuo glicina terminal en la secuencia inferior y la lisina en la secuencia superior se hace mediante una unión amida entre el carboxi-terminal de la glicina y el amino en la cadena lateral de la lisina.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

La ~~porción~~región no ramificada o lineal de una secuencia que contiene cuatro o más aminoácidos específicamente definidos, debe incluirse en una lista de secuencias. En el ejemplo que figura más arriba, las ~~porciones~~regioneslineales del péptido ramificado que tienen más de cuatro aminoácidos son:



El párrafo 7.b) de la Norma ST.26 exige la inclusión de ~~los péptidos~~las secuencias 1 y 2 en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

~~Los péptidos~~Las secuencias 1 y 2 deben representarse con identificadores de secuencia distintos:

DGSAKKK (SEQ ID NO: 32)

AASHG (SEQ ID NO: 33)

~~De preferencia,~~ La secuencia DGSAKKKK ~~debería~~debe incluir una anotación para indicar que la ~~5ª~~ lisina en la posición 5 es un aminoácido modificado, utilizando la clave de caracterización “SITE” junto con el calificador “NOTE” que describa que la cadena lateral de la lisina ~~enlaza el péptido AASHG~~ está enlazada a otra secuencia por medio de una conexión amida. De preferencia, la secuencia AASHG debería incluir una anotación para indicar que la ~~5ª~~ glicina en la posición 5 está enlazada a ~~DGSAKKK~~otra secuencia, utilizando la clave de caracterización “SITE” junto con el calificador “NOTE”.

~~De preferencia,~~ La secuencia DSAKKKK debería incluir una anotación para indicar que la 5ª lisina es un aminoácido modificado, utilizando la clave de caracterización “SITE” junto con el calificador “NOTE” que describa que la lisina enlaza el péptido AASHG. De preferencia, la secuencia AASHG debería incluir una anotación para indicar que la 5ª glicina está enlazada a DGSAKKK, utilizando la clave de caracterización “SITE” junto con el calificador “NOTE”.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~**7.b)**, 26, 29, 30 y 31.

**Ejemplo 7.b)-4: Péptido cíclico que contiene una secuencia de aminoácidos ramificada**

En una solicitud de patente se describe la secuencia siguiente:



La cisteína y la leucina de la estructura cíclica están enlazadas por medio de la cadena lateral de la cisteína y del terminal carbonilo de la leucina.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

La estructura que se muestra a continuación corresponde a una secuencia cíclica ramificada de aminoácidos que contiene los siguientes aminoácidos:



Puesto que la cadena lateral de la cisteína y el carbonilo terminal de la leucina intervienen en la ciclación, el extreme N-terminal del péptido cíclico se encuentra en la posición Cys-1.

**SÍ - la porción cíclica del péptido**

El apartado b) del párrafo 7 de la Norma ST.26 exige que toda región lineal de una secuencia ramificada que contenga cuatro o más aminoácidos específicamente definidos, en la que los aminoácidos formen un único esqueleto peptídico, se incluya en una lista de secuencias. En el ejemplo anterior, la porción cíclica del péptido ramificado tiene más de cuatro aminoácidos y, por lo tanto, debe incluirse en una lista de secuencias.

**NO - la ramificación tripéptídica del péptido**

No se exige la inclusión de la ramificación tripteptídica Ala-Leu-Glu en la lista de secuencias.

**Pregunta 2: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

Con arreglo al párrafo 8, una lista de secuencias no deberá incluir ninguna secuencia que tenga menos de cuatro aminoácidos específicamente definidos.

La ramificación tripéptídica contiene solo tres aminoácidos específicamente definidos y, por lo tanto, no deberá incluirse en una lista de secuencias como una secuencia distinta, con su propio identificador de secuencia.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Si bien este ejemplo ilustra un péptido de configuración circular, el anillo no está compuesto exclusivamente de residuos de aminoácidos unidos por enlaces peptídicos, como se indica en el párrafo 25. Puesto que la ciclación de la secuencia de aminoácidos se produce por medio de la cadena lateral de la cisteína (Cys) y el carboxilo terminal de la leucina (Leu), a la cisteína se le debe asignar la posición 1 dentro de la porción cíclica del péptido. Por consiguiente, la secuencia debe representarse como:

CALRDKL (SEQ ID NO: 89).

Como se indica en la figura anterior, la secuencia de aminoácidos se cicla mediante un enlace tioéster entre la cadena lateral de la cisteína y el carboxilo terminal de la leucina. Debe utilizarse la clave de caracterización "SITE" para describir la cisteína modificada que forma el enlace entre cadenas con la leucina. El operador de localización “join” debería utilizarse con descriptores de localización para indicar qué residuos que intervienen en el acoplamiento, es decir, "join(1,7)". El calificativo obligatorio "NOTE" debería indicar la naturaleza del enlace, por ejemplo, "cysteine leucine thioester (Cys-Leu)", para especificar que las posiciones Cys-1 y Leu-7 están enlazadas por medio de un enlace tioéster. Además, debe anotarse la lisina en la posición 6 para indicar que se ha modificado, utilizando la clave de caracterización "SITE" junto con el calificativo obligatorio "NOTE", cuando el valor calificador describa que la cadena lateral de la lisina enlaza el tripéptido ALE.

**Párrafos de la Norma ST.26:** ~~párrafos~~**7 b)**, 8, 25, 26, 29, 30, 31, 67 y 68

**Ejemplo 7.b)-5: Péptido cíclico que contiene una secuencia de aminoácidos ramificada**

En una solicitud de patente se divulga el siguiente péptido cíclico ramificado:



La serina y la lisina están enlazadas por medio de un enlace amídico entre el carboxi-terminal de la serina y la amina de la cadena lateral de la lisina.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El párrafo 7.b) exige la inclusion de cualquier secuencia que contenga cuatro o más aminoácidos específicamente definidos y que pueda representarse como una región lineal de una secuencia ramificada en una lista de secuencias. En el ejemplo anterior, el péptido contiene una región cíclica en la que los aminoácidos están unidos por enlaces peptídicos, y una región ramificada que se une a una cadena lateral del lisina en la región cíclica. Las regiones de este péptido ramificado que pueden representarse como lineales y que contienen cuatro o más aminoácidos específicamente definidos son:



La Norma ST.26 exige la inclusión de las secuencias 1 y 2 de este péptido cíclico ramificado en una lista de secuencias, cada una con su propio identificador de secuencia.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia 1 debe representarse como:

LRDQS (SEQ. ID. NO: 90)

De preferencia, la secuencia se anota utilizando la clave de caracterización "SITE" junto con el calificativo "NOTE" para describir que la serina en la posición 5 está unida a otra secuencia por medio de un enlace amídico entre una serina y la cadena lateral de una lisina de la otra secuencia.

La secuencia 2 es un péptido cíclico. El párrafo 25 indica que, cuando una secuencia de aminoácidos tiene una configuración circular y carece de amino y carboxiterminales, el solicitante debe elegir el residuo de aminoácidos para la posición 1. Por consiguiente, la secuencia puede representarse como:

ALFKNG (SEQ. ID. NO: 91)

Alternativamente, cualquier otro aminoácido de la secuencia podría designarse como residuo de la posición 1. La secuencia ALFKNG debe describirse más detalladamente utilizando la clave de caracterización "SITE" junto con el calificativo "NOTE" para describir que la cadena lateral de la lisina del residuo de la posición 4 está unida mediante un enlace amídico a otra secuencia. Este enlace de la cadena lateral modifica la lisina y, conforme con el párrafo 30 de la Norma ST.26, un aminoácido modificado debe describirse más detalladamente en el cuadro de características. Además, debe proporcionarse una clave de caracterización "REGION" y un calificativo "NOTE" para indicar que el péptido ALFKNG es circular.

**Párrafos de la Norma ST.26:** ~~párrafos~~**7.b),** 25, 26, 30 y 31

*Párrafo 11.a) – Secuencia de nucleótidos de doble cadena - totalmente complementarias*

**Ejemplo 11.a)-1: Secuencia de nucleótidos de doble cadena – longitudes iguales**

En una solicitud de patente se describe la siguiente secuencia de ADN de doble cadena:

3’-CCGGTTAACGCTA-5’

5’-GGCCAATTGCGAT-3’

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Cada una de las secuencias de nucleótidos enumeradas tiene más de 10 nucleótidos específicamente definidos. Al menos una cadena debe incluirse en la lista de secuencias, porque las dos cadenas de esta secuencia de nucleótidos de doble cadena son totalmente complementarias entre sí.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Si bien debe incluirse en la lista de secuencias la secuencia de una sola cadena, pueden incluirse las secuencias de ambas cadenas, cada una con su identificador de secuencia.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia de ADN de doble cadena debe representarse o bien como secuencia única o bien como dos secuencias distintas. Cada secuencia incluida en la lista de secuencias debe estar representada en sentido 5’ a 3’ y debe atribuírsele su propio identificador de secuencia.

atcgcaattggcc (cadena superior) (SEQ ID NO: 34)

y/o

ggccaattgcgat (cadena inferior) (SEQ ID NO: 35)

**Párrafos de la Norma ST.26:** ~~párrafos~~7.a), **11.a)** y 13.

### *Párrafo 11.b) – Secuencia de nucleótidos de doble cadena – no totalmente complementarias*

**Ejemplo 11.b)-1: Secuencia de nucleótidos de doble cadena – longitudes diferentes**

Una solicitud de patente contiene el dibujo y la leyenda siguientes:

5’-tagttcattgactaaggctccccattgactaaggcgactagcattgactaaggcaagc-3’

||||||||||||||||

gggtaactgantccgc

La región del promotor del gen humano ABC1 (cadena superior) unida por una sonda de APN (cadena inferior), donde “n” en la sonda de APN es una base universal de APN seleccionada del grupo compuesto por 5-nitroindol y 3-nitroindol.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ** – la región promotoraABC1 (cadena superior)

La cadena superior tiene más de 10 nucleótidos enumerados y “específicamente definidos” y se exige su inclusión en una lista de secuencias.

**SÍ** – la sonda deAPN (cadena inferior)

La cadena inferior también debe incluirse en la lista de secuencias, con su propio identificador de secuencia, porque las dos cadenas no son totalmente complementarias entre sí. Los distintos residuos que conforman un APN o “ácido peptidonucleico” se consideran nucleótidos con arreglo al párrafo 3.g) de la Norma ST.26. Por lo tanto, la cadena inferior tiene más de 10 nucleótidos enumerados y “específicamente definidos” y se exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La cadena superior debe incluirse en una lista de secuencias como:

tagttcattgactaaggctccccattgactaaggcgactagcattgactaaggcaagc (SEQ ID NO: 36)

La cadena inferior es un ácido peptidonucleico y por lo tanto no tiene extremo 3’ y 5’. Con arreglo al párrafo 11, debe incluirse en una lista de secuencias “en el sentido 5’ a 3’ y de izquierda a derecha, o en el sentido de izquierda a derecha que imite el sentido 5’ a 3’.” Por lo tanto, debe incluirse en una lista de secuencias como:

cgcctnagtcaatggg (SEQ ID NO: 37)

El calificador “organism” de la clave de caracterización “source” debe tener el valor “synthetic construct” y el calificador obligatorio “mol\_type” con el valor “other DNA”. La cadena inferior debe describirse en un cuadro de características utilizando la clave de caracterización “modified\_base” y el calificador obligatorio “mod\_base” con la abreviatura “OTHER”. Debe incluirse un calificador “note” con el nombre completo, no abreviado, de los nucleótidos modificados, por ejemplo, “nucleósidos N-(2-aminoetil) glicina”.

El residuo “n” debe describirse en más detalle en un cuadro de características utilizando la clave de caracterización “modified\_base” y el calificador obligatorio “mod\_base” con la abreviatura “OTHER”. Debe incluirse un calificador “note” con el nombre completo, no abreviado, del nucleótido modificado: “N-(2-aminoetil) glicina 5-nitroindol o N-(2-aminoetil) glicina 3-nitroindol”.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~3.g), 7.a), **11.b)**, 17 y 18.

**Ejemplo 11.b)-2: Secuencia de nucleótidos de doble cadena – ningún segmento con apareamiento de bases**

En una solicitud de patente se describe la siguiente secuencia de ADN de doble cadena:

3’-CCGGTTAGCTTATACGCTAGGGCTA-5’

||||||| ||||||||||||

5’-GGCCAATATGGCTTGCGATCCCGAT-3’

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Cada cadena de la secuencia de nucleótidos de doble cadena enumerada tiene más de 10 nucleótidos específicamente definidos. Ambas cadenas deben incluirse en la lista de secuencias, cada una con su propio identificador de secuencia, porque las dos cadenas no son totalmente complementarias entre sí.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia de cada cadena debe representarse en el sentido 5’ a 3’ y debe asignársele su propio identificador de secuencia:

atcgggatcgcatattcgattggcc (cadena superior) (SEQ ID NO: 38)

y

ggccaatatggcttgcgatcccgat (cadena inferior) (SEQ ID NO: 39)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.a), **11.b)** y 13.

### *Párrafo 14 – Símbolo “t” interpretado como uracilo en ARN*

**Ejemplo 14-1: El símbolo “t” representa el uracilo en ARN**

En una solicitud de patente se describe el compuesto siguiente:

segmento A: ccugucgt-3’

segmento B: uaguuguagaggccugucct-5’



donde el segmento A y el segmento B son secuencias de ARN.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ – segmento B**

**NO – segmento A**

La secuencia enumerada contiene dos segmentos de nucleótidos específicamente definidos separados por la siguiente estructura “de unión”:



La estructura de unión no es un nucleótido conforme al párrafo 3.g); por lo tanto, cada segmento debe considerarse como una secuencia distinta. El segmento B contiene más de 10 nucleótidos específicamente definidos y el párrafo 7.a) de la Norma ST.26 exige su inclusión en una lista de secuencias. El segmento A contiene tan solo 8 nucleótidos específicamente definidos y por lo tanto no se exige su inclusión en una lista de secuencias.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

El segmento A contiene menos de 10 nucleótidos específicamente definidos y, ~~por lo tanto,~~ conforme al párrafo 8 de la Norma ST.26, no debe incluirse en una lista de secuencias.

**Pregunta 3:¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

El segmento B es una molécula de ARN; por lo tanto, el elemento “INSDSeq\_moltype” debe ser “ARN.” El símbolo “u” no debe utilizarse para representar uracilo en una molécula de ARN en una lista de secuencias. Con arreglo al párrafo 14, el símbolo “t” se interpretará como uracilo en ARN. Por consiguiente, el segmento B debe incluirse en la lista de secuencias como:

tcctgtccggagatgttgat (SEQ ID NO: 40)

La timina en ARN se considera como un nucleótido modificado, es decir, uracilo modificado, y debe representarse en la secuencia como “t” y describirse en más detalle en un cuadro de características. En consecuencia, la timina en posición 1 debe describirse en más detalle utilizando la clave de caracterización “modified\_base”, el calificador “mod\_base” con “OTHER” como valor calificador y un calificador “note” con “timina” como valor calificador.

La timina, es decir, el uracilo modificado, en posición 1 también debería describirse en más detalle en un cuadro de características utilizando la clave de caracterización “misc\_feature” y un calificador “note”. ~~con el valor, por ejemplo, “ccugucgt (el segmento A) está unido en su extremo 3' a un enlazador que está unido al~~ El oxígeno 5’ de la timidina está unido por medio del enlazador ácido (4-(3-hidroxibenzamida) butil) fosfínico a otra secuencia de nucleótidos. Cuando sea posible, la otra secuencia podrá indicarse directamente como un valor del calificador "note".

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~3.g), 7.a), 8, 13, **14**, 19 y 54

### *Párrafo 27 – Debería utilizarse el símbolo de ambigüedad de aminoácidos más restrictivo*

**Ejemplo 27-1: Fórmula abreviada de una secuencia de nucleótidos**

(GGGz)2

donde z es cualquier aminoácido.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

La secuencia está divulgada como fórmula. (GGGz)2 es simplemente una forma abreviada de representar la secuencia GGGzGGGz. Convencionalmente, una secuencia se expande, en primer lugar, y luego se determina la definición de cualquier variable, es decir, “z”.

En la secuencia se utiliza el símbolo no convencional “z”. La definición de “z” debe determinarse a partir de la explicación de la secuencia en la divulgación, que define ese símbolo como cualquier aminoácido (*véase* la Introducción del presente documento). En el ejemplo no se prevén limitaciones respecto de “z”, por ejemplo, que sea igual cada vez que aparece.

Por lo tanto, “z” es equivalente al símbolo convencional “X”, y el péptido del ejemplo tiene ocho aminoácidos enumerados, seis de los cuales son residuos glicina específicamente definidos. El párrafo 7.b) de la Norma ST.26 exige la exige la inclusión de la secuencia en una lista de secuencias como secuencia única con un único identificador de secuencia.

Cabe señalar que la secuencia sigue estando comprendida en el párrafo 7.b) a pesar del hecho de que los residuos enumerados y específicamente definidos no sean contiguos.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

En la secuencia se utiliza el símbolo no convencional “z” que, conforme a la divulgación, es un aminoácido. El símbolo convencional utilizado para representar “todo aminoácido” es “X”. Por lo tanto, la secuencia debe representarse como la secuencia única expandida:

GGGXGGGX (SEQ ID NO: 41)

Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Además, en el ejemplo no se divulga que “z” sea el mismo aminoácido en ambas posiciones en la secuencia expandida. Sin embargo, si “z” se divulga como el mismo aminoácido en ambas posiciones, deberá proporcionarse la clave de caracterización “VARIANT” y un calificador “NOTE” declarando que “X” en las posiciones 4 y 8 puede ser cualquier aminoácido, siempre y cuando se trate del mismo en ambas posiciones.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~3.c), 7.b) y **27**.

**Ejemplo 27-2: Fórmula abreviada – menos de cuatro aminoácidos específicamente definidos**

Un péptido de la fórmula (Gly-Gly-Gly-z)n

En la divulgación se declara asimismo que z es cualquier aminoácido y

i) la variable n tiene cualquier longitud; o

ii) la variable n es 2-100, de preferencia 3

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

Del examen de las dos realizaciones divulgadas, i) y ii), del péptido enumerado de la fórmula, se desprende que “n” puede ser de “cualquier longitud”; por lo tanto, la realización más abarcadora de “n” es indeterminada. Puesto que “n” es indeterminado, el péptido de la fórmula no puede expandirse hasta una longitud determinada y, por lo tanto, debe considerarse la fórmula no expandida.

El péptido enumerado en la fórmula no expandida (“n” = 1) tiene tres aminoácidos específicamente definidos, cada uno de los cuales es Gly, y el símbolo “z”. Convencionalmente “Z” es el símbolo correspondiente a “glutamina o ácido glutámico”; sin embargo, en el ejemplo se define “z” como “cualquier aminoácido” (*véase* la Introducción del presente documento). En virtud de la Norma ST.26, un aminoácido que no esté específicamente definido está representado mediante “X”. A la luz de este análisis, el péptido enumerado, es decir, GGGX, no contiene cuatro aminoácidos específicamente definidos. Por lo tanto, el párrafo 7.b) de la Norma ST.26 no exige su inclusión, a pesar del hecho de que “n” también está definido como valor numérico específico en algunas realizaciones.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En el ejemplo se proporciona un valor numérico específico para la variable “n,” es decir, un límite inferior de 2, un límite superior de 100 y un valor exacto de 3. Toda secuencia que contenga como mínimo cuatro aminoácidos específicamente definidos puede incluirse en la lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Se prefiere una secuencia que contenga 100 copias de GGGX (SEQ ID NO: 42). En una anotación adicional debería indicarse que hasta 98 copias de GGGX podrían suprimirse. Se alienta encarecidamente a incluir realizaciones específicas adicionales que constituyan una parte fundamental de la invención.

Con arreglo al párrafo 27, el ´simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción más detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización VARIANT y un calificador NOTE con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, pueden describirse conjuntamente con la clave de característica VARIANT utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador NOTE con el valor "X puede ser cualquier aminoácido".

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de P.I. podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~3.c), 7.b), 26 y **27**.

**Ejemplo 27-3: Fórmula abreviada - cuatro o más aminoácidos específicamente definidos**

Un péptido de la fórmula (Gly-Gly-Gly-z)n

donde z es cualquier aminoácido y la variable n es 2-100, de preferencia, 3.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El péptido enumerado de la fórmula tiene tres aminoácidos específicamente definidos, cada uno de los cuales es Gly, y el símbolo “z”. Convencionalmente “Z” es el símbolo correspondiente a “glutamina o ácido glutámico”; sin embargo, en la descripción de este ejemplo, “z” se define como “cualquier aminoácido” (*véase* la Introducción del presente documento). En virtud de la Norma ST.26, un aminoácido que no esté específicamente definido está representado mediante “X”. A la luz de este análisis, el péptido repetido enumerado no contiene cuatro aminoácidos específicamente definidos. Sin embargo, en la descripción se prevé un valor numérico específico para la variable “n,” es decir, un límite inferior de 2 y un límite superior de 100. Por lo tanto, en el ejemplo se divulga un péptido que tiene al menos seis aminoácidos específicamente definidos en la secuencia GGGzGGGz, y la Norma ST.26 exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Puesto que “z” representa cualquier aminoácido, el símbolo convencional utilizado para representar el cuarto y el octavo aminoácidos es “X.”

La norma ST.26 exige la inclusión en una lista de secuencias tan solo de la secuencia única que ha sido enumerada mediante sus residuos. Por lo tanto, al menos una secuencia que contenga 2, 3 o 100 copias de GGGX debe incluirse en la lista de secuencias; sin embargo, se prefiere la secuencia más abarcadora que contenga 100 copias de GGGX (SEQ ID NO: 42) (*véase* la Introducción del presente documento). En este último caso, podría indicarse en una anotación adicional que hasta 98 copias de GGGX podrían eliminarse. Se alienta encarecidamente a incluir dos secuencias adicionales que contengan 2 y 3 copias de GGGX, respectivamente (SEQ ID NO: 44-45).

Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Por otra parte, en el ejemplo no se divulga que la variable “z” sea la misma las dos veces que aparece, en la secuencia expandida. Sin embargo, si “z” se divulga como el mismo aminoácido en todas las localizaciones, una clave de caracterización VARIANT y un calificador NOTE deberían indicar que “X” en todas las posiciones puede ser cualquier aminoácido, en la medida en que se trate del mismo en todas las localizaciones.

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representativo no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de P.I. podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~3.c), 7.b), 26 y **27**.

### *Párrafo 28 – Secuencias de aminoácidos separadas por símbolos internos de terminación*

**Ejemplo 28-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada**

En una solicitud de patente se describen las secuencias siguientes:



**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la solicitud se describe una secuencia de nucleótidos, que contiene codones de terminación, que codifica tres secuencias de aminoácidos distintas.

La secuencia de nucleótidos enumerada contiene más de 10 nucleótidos específicamente definidos y debe incluirse en una lista de secuencias como secuencia única.

Por lo que respecta a las secuencias de aminoácidos codificadas, el párrafo 28 exige que las secuencias de aminoácidos separadas por un por un símbolo interno de terminación, por ejemplo, un espacio en blanco, deben incluirse como secuencias distintas. Puesto que la “Protein A”, la “Protein B” y la “Protein C” contienen, cada una de ellas, cuatro o más aminoácidos específicamente definidos, el párrafo 7.b) de la Norma ST.26 exige que cada una de ellas se incluya en una lista de secuencias y se le asigne su propio identificador de secuencia.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia de nucleótidos debe incluirse en una lista de secuencias como:

caattcagggtggtgaatatggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatggaaagcgggcagtgaatgaccatgattacggattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacattggtgtcaaaaataataataaccggatgtactatttatccctgatgctgcgtcgtcaggtgaatgaagtcgcttaagcaatcaatgtcggatgcggcgcgacgcttatccgaccaacatatcataa. (SEQ ID NO: 46)

La secuencia de nucleótidos debería describirse en más detalles utilizando una clave de caracterización “CDS” para cada una de las tres proteínas, y el elemento INSDFeature\_location ~~debería~~debe identificar la localización de cada secuencia codificante, incluyendo el codón de terminación. Además, para cada clave de caracterización “CDS”, debería incluirse el calificador “translation” con la secuencia de aminoácidos de la proteína como valor calificador. En la solicitud no se divulga la tabla del código genético que corresponde a la traducción (véase el Cuadro 5 de la Sección 9 del anexo 1). Si corresponde aplicar la tabla de códigos estándar, el calificador “transl\_table” no será necesario; sin embargo, si corresponde aplicar una tabla del código genético diferente, deberá indicarse el valor calificador adecuado del Cuadro 5 para el calificador “transl\_table”. Por último, debe incluirse el calificador “protein \_id” con el valor calificador que indique el identificador de secuencia de cada una de las secuencias de aminoácidos traducidas.

Las secuencias de aminoácidos deben incluirse como secuencias distintas, atribuyendo a cada una su propio identificador de secuencia:

MAPNTQTASPRALADSLMQLARQVSRLESGQ (SEQ ID NO: 47)

MTMITDSLAVVLQRRDWENPGVTQLNRLAAHWCQK (SEQ ID NO: 48)

MLRRQVNEVA (SEQ ID NO: 49)

NOTA: Véase el “Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones” en la que figura una ilustración de una secuencia de aminoácidos traducida representada como secuencia única.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7, 26, **28**, 57, 87 a 90.

### *Párrafo 29 – Representación de un aminoácido “other”*

**Ejemplo 29-1: Símbolo de ambigüedad más restrictivo para un aminoácido “other”**

En una solicitud de patente se describe la secuencia siguiente:

Ala-Hse-X1-X2-X3-X4-Tyr-Leu-Gly-Ser

donde X1= Ala o Gly,

X2= Ala o Gly,

X3= Ala o Gly,

X4= Ala o Gly y

Hse = Homoserina

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El péptido enumerado contiene cinco aminoácidos específicamente definidos. El símbolo “X” se utiliza convencionalmente para representar dos aminoácidos como alternativa (*véase* la Introducción del presente documento).

Puesto que hay cinco aminoácidos específicamente definidos, es decir, Ala, Tyr, Leu, Gly y Ser, el párrafo 7.b) de la Norma ST.26 exige que la secuencia se incluya en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

El párrafo 29 exige que cualquier aminoácido “other” esté representado por el símbolo “X”. En el ejemplo, la secuencia contiene el aminoácido Hse en la segunda posición, que no figura en el Cuadro 3 de la Sección 3 del Anexo I. En consecuencia, Hse es un aminoácido “other” y debe representarse mediante el símbolo “X”.

X1-X4 son posiciones variantes, y cada una de ellas puede ser A o G. El símbolo de ambigüedad más restrictivo para las alternativas A o G es “X”. Por lo tanto, la secuencia puede ser representada como:

AXXXXXYLGS (SEQ ID NO: 50)

Se alienta encarecidamente a incluir cualesquiera secuencias específicas fundamentales para la divulgación o las reivindicaciones de la invención, tal como se expone en la Introducción del presente documento.

Puesto que el aminoácido Hse no figura en el Cuadro 4 de la Sección 4 del Anexo I, deberán proporcionarse una clave de caracterización “SITE” y un calificador “NOTE” con el nombre completo, no abreviado, de la homoserina en la forma prevista en el párrafo 30 de la Norma ST.26.

Con arreglo al párrafo 27, puesto que X1-X4 representa una alternativa de tan solo dos aminoácidos, es necesaria una descripción adicional. El párrafo 94 indica que debería utilizarse la clave de caracterización “VARIANT” con el calificador “NOTE” y el valor calificador “A o G”. Con arreglo al párrafo 34 de la Norma ST.26, habida cuenta de que esas posiciones son adyacentes y tienen la misma descripción, pueden describirse conjuntamente utilizando la sintaxis “3..6” como descriptor de localización en el elemento INSDFeature\_location.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~3.a), 7.b), 25-27, **29**, 30, 34, 66, 70, 71 y 94 a 95.

### *Párrafo 30 – Anotación de un aminoácido modificado*

**Ejemplo 30-1: Clave de caracterización “CARBOHYD”**

En una solicitud de patente se describe un polipéptido con un aminoácido específicamente modificado, que contiene una cadena lateral glicosilada, que se caracteriza por el hecho de que la Cys correspondiente a las posiciones 4 y 15 del polipéptido forma una unión disulfuro, según la secuencia siguiente:

Leu-Glu-Tyr-Cys-Leu-Lys-Arg-Trp-Asn(asialyloligosaccharide)-Glu-Thr-Ile-Ser-His-Cys-Ala-Trp

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El péptido enumerado prevé 17 aminoácidos específicamente definidos. Hay 16 aminoácidos naturales, siendo el noveno (asparagina) glicosilado. Por lo tanto, la secuencia debe incluirse en una lista de secuencias, según lo exige el párrafo 7.b) de la Norma ST.26.

1. **Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Con arreglo al párrafo 29 de la Norma ST.26, un aminoácido modificado debería representarse en la secuencia como el correspondiente aminoácido no modificado, cuando sea posible.

Por lo tanto, la secuencia debe incluirse en una lista de secuencias como:

LEYCLKRWNETISHCAW (SEQ ID NO: 51)

Es necesaria una descripción más detallada del aminoácido modificado. Debería utilizarse la clave de caracterización “CARBOHYD” junto con el calificador (obligatorio) “NOTE” para indicar que en la posición 9 se produce la unión de una cadena de azúcar (un sialiloligosacárido) con asparagina. El calificador “NOTE” describe el tipo de enlace, por ejemplo enlace N. El descriptor de localización en el elemento localización de característica es el número de posición del residuo de la asparagina modificada.

Además, hay una unión disulfuro entre los dos residuos Cys. Por lo tanto, se utiliza la clave de caracterización “DISULFID” para describir un enlace cruzado entre cadenas. Los descriptores de localización en el elemento localización de característica son los números de posición de los residuos Cys enlazados, conjuntamente con el operador de localización “join”, “join(4,15)”. El calificador NOTE no es obligatorio.

1. **Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~3.a), 7.b), 26, 29, **30** y Sección 7 del Anexo I, clave de caracterización 7.4.

### *Párrafo 36 – Secuencias que contienen regiones de un número exacto de residuos “n” o “X” contiguos*

**Ejemplo 36-1: Secuencia con una región de un número conocido de residuos “X” representada como   
secuencia única**

LL-100-KYMR

donde “-100-“ entre los aminoácidos leucina y lisina refleja una región de 100 aminoácidos en la secuencia.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El párrafo 36 de la Norma ST.26 exige la inclusión de una secuencia que contiene al menos cuatro aminoácidos específicamente definidos separados por una o más regiones de un número definido de residuos “X”.

En la secuencia divulgada se utiliza un símbolo no convencional, es decir “-100-.” La definición de “-100-” debe determinarse a partir de la explicación de la secuencia en la divulgación, que define este símbolo como 100 aminoácidos entre la leucina y la lisina (*véase* la Introducción del presente documento). Por lo tanto, “-100-” es una región definida de residuos “X”. Puesto que seis de los 106 aminoácidos en la secuencia son específicamente definidos, el párrafo 7.b) de la Norma ST.26 exige que la secuencia se incluya en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

El símbolo no convencional “-100-” se representa como 100 residuos “X” (puesto que cualquier símbolo utilizado para representar un aminoácido es equivalente tan solo a un residuo). Por lo tanto, una secuencia única de 106 aminoácidos de longitud, que contiene 100 residuos “X” entre LL y KYMR, debe incluirse en una lista de secuencias (SEQ ID NO: 52).

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.b), 26, 27 y **36**.

**Ejemplo 36-2: Secuencia con múltiples regiones de un número o una serie conocidos de residuos “X”  
representada como secuencia única**

Lys-z2-Lys-zm-Lys-z3-Lys-zn-Lys-z2-Lys

donde z es cualquier aminoácido, m=20, n=19-20, z2 significa que los pares de lisina están separados por cualesquiera dos aminoácidos y z3 significa que los pares de lisina están separados por cualesquiera tres aminoácidos.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la secuencia divulgada se utiliza un símbolo no convencional, es decir “z.” Por lo tanto, debe consultarse la divulgación para determinar la definición; “z” se define como cualquier aminoácido (*véase* la Introducción del presente documento). El símbolo convencional utilizado para representar cualquier aminoácido es “X”. Considerando la presencia de variables “X”, el péptido contiene seis residuos lisina que están enumerados y específicamente definidos, y se exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

En la secuencia se utiliza un símbolo no convencional “z”, cuya definición debe determinarse a partir de la divulgación. Puesto que “z” se define cono cualquier aminoácido, el símbolo convencional es “X.”

El medio de representación preferido y más abarcador es (*véase* la Introducción del presente documento):

KXXKXXXXXXXXXXXXXXXXXXXXKXXXKXXXXXXXXXXXXXXXXXXXXKXXK (SEQ ID NO: 53)

donde zn es igual a 20 “X”, con una descripción adicional de que la variable “X” correspondiente a la posición 30 puede suprimirse.

Como alternativa, o además de lo expuesto más arriba, la secuencia puede ser representada como:

KXXKXXXXXXXXXXXXXXXXXXXXKXXXKXXXXXXXXXXXXXXXXXXXKXXK (SEQ ID NO: 54)

donde zn es igual a 19 “X”, con una descripción adicional de que puede insertarse una variable “X” entre las posiciones 29 y 30.

Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~26, 27 y **36**.

**Ejemplo 36-3: Secuencia con múltiples regiones de un número o una serie conocidos de residuos “X”  
representada como secuencia única**

K-z2-K-zm-K-z3-K-zn-K-z2-K

donde z es cualquier aminoácido, siendo m=15-25, de preferencia 20-22, n=15-25, de preferencia 19-20; z2 significa que los pares de lisina están separados por cualesquiera dos aminoácidos y z3 significa que los pares de lisina están separados por cualesquiera tres aminoácidos.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la secuencia del ejemplo se utiliza un símbolo no convencional, es decir, “z.” Por lo tanto, se consulta la divulgación correspondiente para determinar la definición de “z” (*véase* la Introducción del presente documento). En la divulgación se define este símbolo como cualquier aminoácido. El símbolo convencional utilizado para representar este aminoácido es “X.” Tomando en consideración la presencia de variables “X”, el péptido contiene 6 residuos lisina que están enumerados y específicamente definidos, lo que se exige incluir en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

En la secuencia se utiliza un símbolo no convencional “z”, cuya definición debe determinarse a partir de la divulgación. Puesto que “z” se define como cualquier aminoácido, el símbolo convencional es “X”. El medio de representación preferido y más abarcador es:

KXXKXXXXXXXXXXXXXXXXXXXXXXXXXKXXXKXXXXXXXXXXXXXXXXXXXXXXXXXKXXK (SEQ ID NO: 55)

(donde m=25 y n=25), con una descripción adicional de que hasta 10 residuos “X” en cada una de las regiones “zm” o “zn” pueden suprimirse.

Se alienta encarecidamente a incluir cualesquiera secuencias específicas fundamentales para la divulgación o las reivindicaciones de la invención, según lo expuesto en la Introducción del presente documento.

Como alternativa, la secuencia puede ser representada como:

KXXKXXXXXXXXXXXXXXXKXXXKXXXXXXXXXXXXXXXKXXK (SEQ ID NO: 56)

(donde m=15 y n=15), con una descripción adicional de que pueden insertarse hasta 10 residuos “X” en cada una de las regiones “zm” or “zn”.

Como alternativas adicionales, pueden incluirse cualesquiera o todas las variaciones posibles.

Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de P.I. podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~27 y **36**.

### *Párrafo 37 – Secuencias que contienen regiones de un número desconocido de residuos “n” o “X”*

**Ejemplo 37-1: Una secuencia con regiones de un número desconocido de residuos “X” no debe   
representarse como secuencia única**

Gly-Gly----Gly-Gly-Xaa-Xaa

donde el símbolo ---- es una región indefinida en la secuencia, Xaa es cualquier aminoácido y los residuos glicina y Xaa están conectados entre sí mediante enlaces peptídicos.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO**

El párrafo 37 de la Norma ST.26 prohíbe la inclusión de toda secuencia que contenga una región indefinida; por lo tanto, no se exige la inclusión de la totalidad de la secuencia.

El párrafo 37 de la Norma ST.26 exige la inclusión de toda porción de una secuencia adyacente a una región indefinida que contenga cuatro o más aminoácidos específicamente definidos. En el ejemplo que figura más arriba, no se exige la inclusión de las porciones adyacentes a la región indefinida, puesto que cada porción contiene solo dos aminoácidos específicamente definidos.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO –** la totalidad de la secuencia, no

**NO –** cualquier porción de la secuencia, no

El párrafo 37 de la Norma ST.26 no permite la inclusión de la totalidad de la secuencia.

El párrafo 8 de la Norma ST.26 no permite la inclusión de las porciones adyacentes a la región indefinida, puesto que cada porción contiene solo dos aminoácidos específicamente definidos.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.b), 8, 26 y **37**.

**Ejemplo 37-2: Una secuencia con regiones de un número desconocido de residuos “X” no debe  
representarse como secuencia única**

Gly-Gly----Gly-Gly-Ala-Gly-Xaa-Xaa

donde el símbolo ---- es una región indefinida en la secuencia, Xaa es cualquier aminoácido y los residuos glicina y Xaa están conectados entre sí mediante enlaces peptídicos.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO –** la totalidad de la secuencia, no

**SÍ –** una porción de la secuencia

El párrafo 37 de la Norma ST.26 prohíbe la inclusión de toda secuencia que contenga una región indefinida, pero exige la inclusión de toda porción de una secuencia adyacente a una región indefinida que contenga cuatro o más aminoácidos específicamente definidos.

En el ejemplo que figura más arriba, la Norma ST.26 no exige (y prohíbe) la inclusión tanto de la totalidad de la secuencia que contiene una región indefinida como de la porción Gly-Gly adyacente a la región indefinida, que contiene solo dos aminoácidos específicamente definidos. Sin embargo, la Norma ST.26 exige la inclusión de la porción Gly-Gly-Ala-Gly- Xaa-Xaa adyacente a la región indefinida, puesto que contiene al menos cuatro aminoácidos específicamente definidos.

**Pregunta 2: ¿Permite la Norma ST.26 la inclusión de la secuencia o secuencias?**

**NO –** la totalidad de la secuencia y la porción Gly-Gly, no

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La porción de la secuencia adyacente a la región indefinida que contiene cuatro aminoácidos específicamente definidos debe representarse como:

GGAGXX (SEQ ID NO: 57)

De preferencia, debería anotarse la secuencia de manera que indique que la secuencia representada forma parte de una secuencia mayor que contiene una región indefinida, utilizando la clave de caracterización “SITE”, la localización de característica “1” y el calificador “NOTE” con el valor, por ejemplo, “Este residuo presenta un enlace N-terminal a un péptido que tiene un N-terminal Gly-Gly y una región de longitud indefinida”.

Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.b), 8, 26, 27 y **37**.

### *Párrafo 55 – Una secuencia de nucleótidos que contiene segmentos de ADN y ARN*

**Ejemplo 55-1: Molécula combinada ADN/ARN**

En una solicitud de patente se divulga la secuencia de oligonucleótidos siguiente:

AGACCTTcggagucuccuguugaacagauagucaaaguagauC

En la que las letras mayúsculas representan residuos de ADN y las minúsculas, residuos de ARN.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

La secuencia divulgada tiene más de diez nucleótidos enumerados y específicamente definidos; por lo tanto, se exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia de nucleótidos debe incluirse en una lista de secuencias como:

agaccttcggagtctcctgttgaacagatagtcaaagtagatc (SEQ ID NO: 92)

Obsérvese que los nucleótidos de uracilo deben representarse mediante el símbolo "t" en la lista de secuencias.

El párrafo 55 de la Norma ST.26 establece que para toda secuencia de nucleótidos que contenga segmentos de ADN y ARN el tipo de molécula deberá indicarse como "DNA” y describirse con detalle utilizando la clave de caracterización "fuente", el calificador obligatorio "organism" con el valor "synthetic construct" y el calificador obligatorio "mol\_type" con el valor "other DNA". Además, cada segmento de la secuencia debe describirse con detalle utilizando la clave de caracterización "misc\_feature", que incluye la localización del segmento y el calificador "note", que indica si el segmento es ADN o ARN. La secuencia divulgada contiene dos segmentos de ADN (posiciones de nucleótidos 1 a 7 y 43) y un segmento de ARN (posiciones de nucleótidos 8 a 42).

### **Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7, 14, **55**-56, and 83

### *Párrafo 87 – Clave de caracterización “CDS”*

**Ejemplo 87-1: Secuencia de nucleótidos codificante y secuencia de aminoácidos codificada**

En una solicitud de patente se describe la siguiente secuencia de nucleótidos y su traducción:

atg acc gga aat aaa cct gaa acc gat gtt tac gaa att tta tga

Met Thr Gly Asn Lys Pro Glu Thr Asp Val Tyr Glu Ile Leu STOP

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ ~~– la secuencia de nucleótidos.~~**

La secuencia de nucleótidos enumerada tiene más de diez nucleótidos específicamente definidos.

**~~SÍ – la secuencia de nucleótidos.~~**

La secuencia de ~~péptidos~~aminoácidos enumerada tiene más de cuatro aminoácidos específicamente definidos.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia de nucleótidos debe representarse como:

atgaccggaaataaacctgaaaccgatgtttacgaaattttatga (SEQ ID NO: 58)

La secuencia de nucleótidos debería describirse además utilizando la clave de caracterización “CDS” y el elemento INSDFeature\_location ~~debería~~deberá identificar la totalidad de la secuencia, incluyendo el codón de terminación (es decir, de la posición 1 a la 45). Asimismo, debería incluirse el calificador “translation” con el valor calificador “MTGNKPETDVYEIL”. La solicitud no divulga la tabla del código genético que corresponde aplicar a la traducción (véase el Cuadro 5 de la Sección 9 del Anexo 1). Si corresponde aplicar la tabla del código genético estándar, el calificador “transl\_table” no será necesario; sin embargo, si corresponde aplicar una tabla del código genético diferente, deberá indicarse el valor calificador adecuado de la Tabla 5 para el calificador “transl\_table”. Por último, deberá incluirse el calificador “protein \_id” con el valor calificador que indique el identificador de secuencia de la secuencia de ~~péptidos~~aminoácidos traducida.

La secuencia de ~~péptidos~~aminoácidos debe presentarse por separado con su propio identificador de secuencia utilizando códigos de una letra, según se indica a continuación:

MTGNKPETDVYEIL (SEQ ID NO: 59)

El “STOP” después de la secuencia de ~~péptidos~~aminoácidos enumerada no debe incluirse en la secuencia de ~~péptidos~~aminoácidos en la lista de secuencias.

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de P.I. podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.a), 7.b), 26, 28, **87**, 88 y 90.

**Ejemplo 87-2: Una localización de característica se extiende más allá de la secuencia divulgada**

Una solicitud de patente contiene la figura siguiente, en la que se divulga una secuencia codificadora parcial y su secuencia de aminoácidos traducida:

cat cac gca gca gaa tgt gga ttt tgt cct caa caa tgg caa gtt cta 48

His His Ala Ala Glu Cys Gly Phe Cys Pro Gln Gln Trp Gln Val Leu

1 5 10 15

cgt ggg agt ctg tgc att tgt gag ggt cca gct gaa gga tgg ttc ata 96

Arg Gly Ser Leu Cys Ile Cys Glu Gly Pro Ala Glu Gly Trp Phe Ile

20 25 30

tca aga tgt tgg tta tgg tgt ggg cct caa gtc caa ggc ttt atc ttt 144

Ser Arg Cys Trp Leu Trp Cys Gly Pro Gln Val Gln Gly Phe Ile Phe

35 40 45

gga gaa ggc aag gaa gga ggc ggt gac aga cgg gct gaa gcg agc cct 192

Gly Glu Gly Lys Glu Gly Gly Gly Asp Arg Arg Ala Glu Ala Ser Pro

50 55 60

cag gag ttt tgg gaa tgc act tgg 216

Gln Glu Phe Trp Glu Cys Thr Trp

65 70

Figure 1 – partial coding sequence of the *Homo sapi*ens *ITCH1* gene, which encodes amino acids 20 through 91 of the 442 amino acid long ITCH1 protein.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la solicitud se divulga una secuencia de nucleótidos y su secuencia de aminoácidos traducida. La secuencia de nucleótidos enumerada contiene más de 10 nucleótidos específicamente definidos y deberá incluirse en una lista de secuencias.

La secuencia de aminoácidos contiene más de 4 aminoácidos específicamente definidos y también deberá incluirse en una lista de secuencias como una secuencia distinta con su propio identificador de secuencia.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia de nucleótidos debe incluirse en una lista de secuencias como:

catcacgcagcagaatgtggattttgtcctcaacaatggcaagttctacgtgggagtctgtgcatttgtgagggtccagctgaaggatggttcatatcaagatgttggttatggtgtgggcctcaagtccaaggctttatctttggagaaggcaaggaaggaggcggtgacagacgggctgaagcgagccctcaggagttttgggaatgcacttgg (SEQ ID NO: 93)

La secuencia de nucleótidos debería describirse con detalle utilizando una clave de caracterización "CDS". El elemento INSDFeature\_location debe identificar la ubicación la caracterización "CDS" en la secuencia e incluir el codón de terminación.

En la figura se describe una secuencia de codificación parcial que no incluye el codón de inicio ni el codón de terminación. Sin embargo, la descripción de la secuencia indica que el codón de inicio se sitúa antes del nucleótido en la posición 1 y el codón de terminación se sitúa después del último nucleótido en la posición 216.

La Norma ST.26 establece que el descriptor de localización no deberá incluir números de residuos fuera de la serie de la secuencia indicada en el elemento INSDSeq\_sequence. Por consiguiente, en el ejemplo anterior el descriptor de localización para la clave de caracterízación CDS no puede incluir números de posición fuera del intervalo entre las posiciones 1 y 216. La posición del codón de terminación en el elemento INSDFeature\_location debe representarse utilizando el símbolo ">" para indicar que el codón de terminación se sitúa después de la posición 216. Del mismo modo, el símbolo "<" puede utilizarse para indicar que el codón de inicio se situá después de la posición 1. Por lo tanto, el descriptor de localización para la clave de caracterización CDS debería aparecer como se indica a continuación:

<1..>216

Debe tenerse presente que "<" y ">" son caracteres reservados y serán sustituidos por "&lt;" y "&gt;", respectivamente, en la instancia XML de la lista de secuencias.

Debería incluirse el calificador “translation” con la secuencia de aminoácidos de la proteína como valor calificador.

En la figura no se divulga la tabla del código genético que corresponde aplicar a la traducción (véase la Tabla 5 de la Sección 9 del Anexo 1). Si corresponde aplicar la tabla del “código estándar”, el calificador “transl\_table” no será necesario; sin embargo, si corresponde aplicar una tabla del código genético diferente, deberá indicarse el valor calificador adecuado del Cuadro 5 que figura en el Anexo I de la Norma ST. 26 para el calificador “transl\_table”. Por último, deberá incluirse el calificador “protein \_id” con el valor calificador que indique el identificador de secuencia de la secuencia de aminoácidos traducida.

La secuencia de aminoácidos traducida deberá incluirse como una secuencia distinta con su propio identificador de secuencia:

HHAAECGFCPQQWQVLRGSLCICEGPAEGWFISRCWLWCGPQVQGFIFGEGKEGGGDRRAEASPQEFWECTW (SEQ ID NO: 94)

ATENCIÓN: La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de PI podría considerar que la información suministrada añade materia a la divulgación original.

Párrafos pertinentes de la Norma ST.26: ~~párrafos~~7, 41, 65, 66, 70, 71, 87 y 90

### *Párrafo 90 –Secuencia de aminoácidos codificada por una secuencia codificadora*

**Ejemplo 90-1: Secuencia de aminoácidos codificada por una secuencia codificadora con intrones**

Una solicitud de patente contiene la figura siguiente, en la que se divulga una secuencia codificadora y su traducción: 

Figura 1 – los nucleótidos que figuran en letra negrita son regiones de intrones.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En la solicitud se divulga una secuencia de nucleótidos y su traducción a aminoácidos. La secuencia de nucleótidos enumerada contiene más de 10 nucleótidos específicamente definidos y debe incluirse en una lista de secuencias como secuencia única.

La secuencia de nucleótidos contiene secuencia codificante (exones) separada por secuencia no codificante (intrones). La figura ilustra la traducción de la secuencia de nucleótidos como tres secuencias de aminoácidos no contiguas. Conforme a la leyenda de la figura, las regiones de nucleótidos que figuran en letra negrita son secuencias de intrones que serán eliminadas de un transcripto de ARN antes de la traducción a proteína. En consecuencia, de hecho, las tres secuencias de aminoácidos son una única secuencia enumerada, contigua, que contiene más de cuatro aminoácidos específicamente definidos y debe incluirse en una lista de secuencias como secuencia única.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia de nucleótidos debe incluirse en una lista de secuencias como:

atgaagactttcgcagccttgctttccgctgtcactctcgcgctctcggtgcgcgcccaggcggctgtctggagtcaatgtaagtgccgctgcttttcattgatacgagactctacgccgagctgacgtgctaccgtataggtggcggtacaccgggttggacgggcgagaccacttgcgttgctggttcggtttgtacctccttgagctcagtgagcgactttcaatccgtcgtcattgctcctcatgtattgacgattggccttcatagtcatactctcaatgcgttccgggctccgcaacgtccagcgctccggcggccccctcagcgacaacttcaggccccgcacctacggacggaacgtgctcggccagcggggcatggccgccattgacctga (SEQ ID NO: 74)

La secuencia de nucleótidos debería describirse en más detalle utilizando una clave de caracterización “CDS” y el elemento INSDFeature\_location ~~debería~~deberá identificar la localización de la secuencia codificante, incluyendo el codón de terminación indicado mediante “Ter”. Además, debería incluirse el calificador “translation”, con la secuencia de aminoácidos de la proteína como valor calificador. (Cabe señalar que el símbolo de terminación “Ter” en la última posición de la secuencia no debe incluirse en la secuencia de aminoácidos.) En la solicitud no se divulga la tabla del código genético que corresponde aplicar a la traducción (véase el Cuadro 5 de la Sección 9 del Anexo 1). Si corresponde aplicar la tabla del “código estándar”, el calificador “transl\_table” no será necesario; sin embargo, si corresponde aplicar una tabla del código genético diferente, deberá indicarse el valor calificador adecuado del Cuadro 5 para el calificador “transl\_table”. Por último, debe incluirse el calificador “protein \_id” con el valor calificador que indique el identificador de secuencia de la secuencia de aminoácidos traducida.

La secuencia de aminoácidos debe incluirse como secuencia única:

MKTFAALLSAVTLALSVRAQAAVWSQCGGTPGWTGETTCVAGSVCTSLSSSYSQCVPGSATSSAPAAPSATTSGPAPTDGTCSASGAWPPLT (SEQ ID NO: 75)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7, 26, 28, 57, 87 a **90**.

### *Párrafo 91 – Secuencia principal y una variante, cada una de ellas enumerada mediante sus residuos*

**Ejemplo 91-1: Representación de variantes enumeradas**

La descripción incluye el siguiente alineamiento de secuencias.

*D. melanogaster ACATTGAATCTCATACCACTTT*

*D. virilis ...-..G...C..--.G.....*

*D. simulans GT..G.CG..GT..SGT.G...*

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

En esta ámbito de la técnica, es común incluir “puntos” en un alineamiento de secuencias para indicar: “esta posición es la misma que la posición que se encuentra por encima de ella”. Por lo tanto, los “puntos” en las secuencias de las especies ~~2 y 3~~*D. virilis* y *D. simulans* se consideran nucleótidos enumerados y específicamente definidos, pues son simplemente una forma abreviada de indicar que una posición determinada es el mismo nucleótido que en ~~la especie 1~~*D. melanogaster.* Además, con el fin de maximizar el alineamiento, en los alineamientos de secuencias figura con frecuencia el símbolo “-“ para indicar la ausencia de un residuo.

Por consiguiente, la secuencia de nucleótidos de las especies ~~1 y 3~~*D. melanogaster* y *D. simulans* contiene 22 nucleótidos enumerados y específicamente definidos, y la secuencia de nucleótidos de la especie ~~2~~*D. virilis* contiene 19. Así pues, el párrafo 7.a) de la Norma ST.26 exige que cada una de las secuencias se incluya en una lista de secuencias con un identificador de secuencia distinto.

En esta ámbito de la técnica, es común incluir “puntos” en un alineamiento de secuencias para indicar: “esta posición es la misma que la posición que se encuentra por encima de ella”. Por lo tanto, los “puntos” en las especies 2 y 3 se consideran nucleótidos enumerados y específicamente definidos, pues son simplemente una forma abreviada de indicar que una posición determinada es el mismo nucleótido que en la especie 1. Además, con el fin de maximizar el alineamiento, en los alineamientos de secuencias figura con frecuencia el símbolo “-“ para indicar la ausencia de un residuo.

Por consiguiente, la secuencia de nucleótidos de las especies 1 y 3 contiene 22 nucleótidos enumerados y específicamente definidos, y la secuencia de nucleótidos de la especie 2 contiene 19. Así pues, el párrafo 7.a) de la Norma ST.26 exige que cada una de las secuencias se incluya en una lista de secuencias con un identificador de secuencia distinto.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia *Drosophila melanogaster* debe incluirse en una lista de secuencias como:

acattgaatctcataccacttt (SEQ ID NO: 60)

La secuencia *Drosophila virilis* debe incluirse en una lista de secuencias como:

acatggatcccacgacttt (SEQ ID NO: 61)

La secuencia *Drosophila simulans* debe incluirse en una lista de secuencias como:

gtatggcgtcgtatsgtagttt (SEQ ID NO: 62)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.a), 13 y **91**.

**Ejemplo 91-2: Representación de variantes enumeradas**

La descripción incluye la siguiente tabla de un péptido y sus variantes funcionales. Un espacio en blanco en la tabla indica que un aminoácido en la variante es el mismo que el aminoácido correspondiente en la “Secuencia” y un “-“ indica la eliminación del aminoácido correspondiente en la “Secuencia”.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Posición | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Secuencia | A | V | L | T | Y | L | R | G | E |
| Variante 1 |  |  |  |  |  |  |  |  | A |
| Variante 2 |  |  | P |  |  | P |  |  |  |
| Variante 3 |  |  | A | I | G | Y |  |  |  |
| Variante 4 |  |  |  |  |  |  | - |  |  |

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Tal como se ha indicado, un espacio en blanco en esta tabla indica que un aminoácido en la variante es el mismo que el aminoácido correspondiente en la “Secuencia”. Por lo tanto, los aminoácidos de las variantes de secuencias están enumerados y específicamente definidos.

Puestos que las cuatro variantes de secuencias contienen, cada una, más de cuatro aminoácidos enumerados y específicamente definidos, el párrafo 7.~~a~~b) de la Norma ST.26 exige que cada una de las secuencias se incluya en una lista de secuencias con un identificador de secuencia distinto.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

AVLTYLRGE (SEQ ID NO: 76)

AVLTYLRGA (SEQ ID NO: 77)

AVPTYPRGE (SEQ ID NO: 78)

AVAIGYRGE (SEQ ID NO: 79)

AVLTYLGE (SEQ ID NO: 80)

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.b), 26 y **91**.

**Ejemplo 91-3: Representación de una secuencia de consenso**

Una solicitud de patente incluye la Figura 1, con el siguiente alineamiento múltiple de secuencias:

*Consensus LEG****n****EQFINA****ak****IIRHP****k****Y****nrk****T****ln****NDI****m****LIK*

*Homo sapiens* LEG**N**EQFINA**AK**IIRHP**Q**Y**DRK**T**LN**NDI**M**LIK

*Pongo abelii* LEG**N**EQFINA**AK**IIRHP**Q**Y**DRK**T**VN**NDI**M**LIK

*Papio Anubis* LEG**T**EQFINA**AK**IIRHP**D**Y**DRK**T**LN**NDI**L**LIK

*Rhinopithecus roxellana* LEG**T**EQFINA**AK**IIRHP**N**Y**NRI**T**LD**NDI**L**LIK

*Pan paniscus* LEG**N**EQFINA**AK**IIRHP**K**Y**NRI**T**LN**NDI**M**LIK

*Rhinopithecus bieti* LEG**N**EQFINA**TK**IIRHP**K**Y**NGN**T**LN**NDI**M**LIK

*Rhinopithecus roxellana* LEG**N**EQFINA**TQ**IIRHP**K**Y**NGN**T**LN**NDI**M**LIK

1. La secuencia de consenso incluye letras mayúsculas para representar residuos de aminoácidos conservados, así como las letras minúsculas “n”, “a”, “k”, “r”, “l” y “m” que representan los residuos de aminoácidos predominantes entre las secuencias alineadas.

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

Las letras minúsculas en la secuencia de consenso representan, cada una, un único residuo de aminoácido. En consecuencia, la secuencia de consenso, así como cada una de las siete secuencias restantes de la Figura 1, incluye al menos cuatro aminoácidos específicamente definidos. El párrafo 7.b) de la Norma ST.26 exige la inclusión de las ocho secuencias en la lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

Las letras minúsculas en la secuencia de consenso se utilizan como símbolos de ambigüedad para representar el aminoácido predominante entre las posibles variantes para una posición específica. Por lo tanto, las letras minúsculas “n”, “a”, “k”, “r”, “l” y “m” son símbolos convencionales utilizados de manera no convencional, y la secuencia de consenso debe representarse utilizando un símbolo de ambigüedad en lugar de cada una de las letras minúsculas.

Debería utilizarse el símbolo de ambigüedad más restrictivo. Para la mayoría de las posiciones en la secuencia de consenso, “X” es el símbolo de ambigüedad más restrictivo; sin embargo, el símbolo de ambigüedad más restrictivo para“D” o “N” en las posiciones 20 y 25 es “B”. La secuencia de consenso debería incluirse en la lista de secuencias como:

LEGXEQFINAXXIIRHPXYBXXTXBNDIXLIK (SEQ ID NO: 81)

Con arreglo al párrafo 27, el símbolo “X” se interpretará como cualquiera de los símbolos “A”, “R”, “N”, “D”, “C”, “Q”, “E”, “G”, “H”, “I”, “L”, “K”, “M”, “F”, “P”, “O”, “S”, “U”, “T”, “W”, “Y”, o “V”, excepto cuando se utilice con una descripción más detallada en el cuadro de características. Por lo tanto, cada “X” en la secuencia de consenso deberá describirse en mayor detalle en un cuadro de características utilizando la clave de caracterización “VARIANT” y el calificador “NOTE” para indicar las posibles variantes para cada posición.

Las siete secuencias restantes deben incluirse en la lista de secuencias como:

LEGNEQFINAAKIIRHPQYDRKTLNNDIMLIK (SEQ ID NO: 82)

LEGNEQFINAAKIIRHPQYDRKTVNNDIMLIK (SEQ ID NO: 83)

LEGTEQFINAAKIIRHPDYDRKTLNNDILLIK (SEQ ID NO: 84)

LEGTEQFINAAKIIRHPNYNRITLDNDILLIK (SEQ ID NO: 85)

LEGNEQFINAAKIIRHPKYNRITLNNDIMLIK (SEQ ID NO: 86)

LEGNEQFINATKIIRHPKYNGNTLNNDIMLIK (SEQ ID NO: 87)

LEGNEQFINATQIIRHPKYNGNTLNNDIMLIK (SEQ ID NO: 88)

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de PI podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.b), 26, 27, **91** y 95.

### *Párrafo 92 – Variante de secuencia divulgada como secuencia única con residuos alternativos enumerados*

**Ejemplo 92-1: Representación de una secuencia única con aminoácidos alternativos enumerados**

En una solicitud de patente se reivindica un péptido de la secuencia:

(i) Gly-Gly-Gly-[Leu o Ile]-Ala-Thr-[Ser o Thr]

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

La secuencia prevé cuatro aminoácidos específicamente definidos, y el párrafo 7.b) de la Norma ST.26 exige la inclusión de la secuencia en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

En el Cuadro 3 de la Sección 3 del Anexo I se define el símbolo de ambigüedad “J” como isoleucina o leucina. Por lo tanto, la representación preferida de la secuencia es:

GGGJATX (SEQ ID NO: 63)

lo que exige una descripción más detallada en un cuadro de características utilizando la clave de caracterización “VARIANT” y el calificador “NOTE” para indicar que “X” es serina o treonina.

Como alternativa, la secuencia puede representarse, por ejemplo, como:

GGGLATS (SEQ ID NO: 64)

lo que exige una descripción más detallada en un cuadro de características utilizando la clave de caracterización “VARIANT” y el calificador “NOTE” para indicar que L puede ser sustituido por I, y S puede ser sustituido por T.

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de P.I. podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.b), 8, 26, 27, **92** y 95.

### *Párrafo 93.a) – Variante de secuencia divulgada únicamente mediante referencia a una secuencia principal con múltiples variaciones independientes*

**Ejemplo 93.a)-1: Representación de una variante de secuencia mediante anotación de la secuencia principal**

Una solicitud contiene la divulgación siguiente:

“El fragmento de péptido 1 es Gly-Leu-Pro-Xaa-Arg-Ile-Cys, donde Xaa puede ser cualquier aminoácido...

En otra realización, el fragmento de péptido 1 es Gly-Leu-Pro-Xaa-Arg-Ile-Cys, donde Xaa puede ser Val, Thr, o Asp...

En otra realización, el fragmento de péptido 1 es Gly-Leu-Pro-Xaa-Arg-Ile-Cys donde Xaa puede ser Val.”

**Pregunta 1: ¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

El “fragmento de péptido 1” en cada una de las tres realizaciones divulgadas provee al menos seis aminoácidos específicamente definidos; por lo tanto, la secuencia debe incluirse en una lista de secuencias, según lo exige el párrafo 7.b) de la Norma ST.26.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

En este ejemplo, la secuencia enumerada de “fragmento de péptido 1” se divulga tres veces, como tres realizaciones diferentes, cada una con una descripción alternativa de Xaa. En este ejemplo, “X” es el símbolo de ambigüedad más restrictivo para la posición Xaa.

La Norma ST.26 exige que se incluya solo una vez la secuencia enumerada divulgada. En la más abarcadora de las tres realizaciones, Xaa es cualquier aminoácido (véase la Introducción del presente documento). Por lo tanto, la secuencia que debe incluirse en la lista de secuencias es:

GLPXRIC (SEQ ID NO: 65)

Con arreglo al párrafo 27, el simbolo "X" se interpretará como cualquiera de los símbolos "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y" o "V", excepto cuando se utilice con una descripción detallada en el cuadro de características. Por lo tanto, si se pretende que "X" represente "cualquier aminoácido", debería anotarse con la clave de caracterización “VARIANT” y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Cuando sea posible, cada "X" debería anotarse individualmente. Sin embargo, toda región que contiene residuos "X" contiguos, o numerosos residuos "X" dispersos en toda la secuencia, puede describirse en conjunto con la clave de caracterización “VARIANT” utilizando la sintaxis "x..y" como descriptor de localización, donde x e y son las posiciones del primer y último residuo "X", y un calificador “NOTE” con el valor "X puede ser cualquier aminoácido".

Se alienta encarecidamente a incluir cualesquiera secuencias adicionales fundamentales para la divulgación o las reivindicaciones de la invención, según se expone en la Introducción del presente documento.

Para el ejemplo que figura más arriba, se recomienda encarecidamente la inclusión en la lista de secuencias de las siguientes tres secuencias adicionales, cada una con su propio identificador de secuencia:

GLPVRIC (SEQ ID NO: 66)

GLPTRIC (SEQ ID NO: 67)

GLPDRIC (SEQ ID NO: 68)

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de P.I. podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.b), 26, 27 y **93.a)**.

### *Párrafo 93.b) – Variante de secuencia divulgada únicamente mediante referencia a una secuencia principal con múltiples variaciones interdependientes*

**Ejemplo 93.b)-1: Representación de distintas variantes de secuencias con múltiples variaciones interdependientes**

En una solicitud de patente se describe la siguiente secuencia de consenso:

cgaatg**n1**cccactacgaatg**n2**cacgaatg**n3**cccaca

donde n1, n2 y n3 puede ser a, t, g, o c.

Distintas variantes de secuencia se divulgan de la manera siguiente:

si n1 es a, n2 y n3 son t, g, o c;

si n1 es t, n2 y n3 son a, g, o c;

si n1 es g, n2 y n3 son t, a, o c;

si n1 es c, n2 y n3 son t, g, o a.

**Pregunta 1:¿Exige la Norma ST.26 la inclusión de la secuencia o secuencias?**

**SÍ**

La secuencia tiene más de diez nucleótidos enumerados y “específicamente definidos”, y el párrafo 7.a) de la Norma ST.26 exige su inclusión en una lista de secuencias.

**Pregunta 3: ¿Cómo deberían representarse la secuencia o secuencias en la lista de secuencias?**

La secuencia enumerada contiene más de diez nucleótidos específicamente definidos y tres residuos “n”. La norma ST.26 exige la inclusión de la secuencia enumerada divulgada y si corresponde utilizar un símbolo de ambigüedad debería utilizarse el símbolo más restrictivo. En este ejemplo, n1, n2 y n3 pueden ser a, t, g, o c, por lo cual “n” es el símbolo de ambigüedad más restrictivo. Por consiguiente, la secuencia que debe incluirse en la lista de secuencias es:

cgaatgncccactacgaatgncacgaatgncccaca (SEQ ID NO: 69)

La secuencia enumerada contiene variaciones en tres localizaciones distintas y esas variaciones aparecen de forma interdependiente. Se alienta **encarecidamente** a incluir secuencias adicionales que representen realizaciones adicionales que son parte fundamental de la invención, según se expone en la Introducción del presente documento. Por lo tanto, con arreglo al párrafo 93.b) de la Norma ST.26, las realizaciones adicionales deberían incluirse en una lista de secuencias como cuatro secuencias distintas, cada una con su propio identificador de secuencia:

cgaatgacccactacgaatgbcacgaatgbcccaca (SEQ ID NO: 70)

cgaatgtcccactacgaatgvcacgaatgvcccaca (SEQ ID NO: 71)

cgaatggcccactacgaatghcacgaatghcccaca (SEQ ID NO: 72)

cgaatgccccactacgaatgdcacgaatgdcccaca (SEQ ID NO: 73)

(Cabe señalar que b = t, g, o c; v = a, g, o c; h = t, a, o c; y d = t, g, o a; véase el Cuadro 1 de la Sección 1 del Anexo I)

Con arreglo al párrafo 15 de la Norma ST.26, debe utilizarse el símbolo más restrictivo para representar las posiciones variables. En consecuencia, n2 y n3 no deben estar representados por “n” en la secuencia.

**ATENCIÓN:** La representación preferida de la secuencia indicada más arriba se refiere al suministro de una lista de secuencias en la fecha de presentación de una solicitud de patente. La misma representación no puede aplicarse a una lista de secuencias suministrada con posterioridad a la fecha de presentación de una solicitud de patente, puesto que es necesario tener en cuenta si una oficina de PI podría considerar que la información suministrada añade materia a la divulgación original.

**Párrafos pertinentes de la Norma ST.26:** ~~párrafos~~7.a), 15 y **93.b)**.

[Sigue el Apéndice del Anexo VI de la Norma ST.26]

**APÉNDICE**

SECUENCIAS DEL DOCUMENTO DE ORIENTACIÓN EN XML

<?xml version="1.0" encoding="UTF-8"?>

<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.~~0~~2//EN" "~~resources/~~ST26SequenceListing\_V1\_~~0~~2.dtd">

<ST26SequenceListing dtdVersion="1.2" fileName="Guidance\_Document\_Sequences\_XML" softwareName="prototype" softwareVersion="1.0" productionDate="~~2017-01-02~~2018-07-09">

<ApplicantFileReference>ABCD#1234</ApplicantFileReference>

<ApplicantName languageCode="de">Johannes ~~Jäger~~JÃ¤ger</ApplicantName>

<ApplicantNameLatin>Johannes Jaeger</ApplicantNameLatin>

<InventionTitle languageCode="de">Pharmakologische Wirkstoffe ~~für~~fÃ¼r das Nervensystem</InventionTitle>

<SequenceTotalQuantity>~~88~~94</SequenceTotalQuantity>

<SequenceData sequenceIDNumber="1">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>REGION</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>circular peptide</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>D-Alanine</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>2</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>D-Glutamic acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Nle</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>D-Methionine</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>D-Norleucine</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>AEKXGMX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="2">

<INSDSeq>

<INSDSeq\_length>6</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>the N-terminus is acetylated</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>6-amino-7-(1H-indol-3-yl)-5-oxoheptanoic acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>the C-terminus is methylated</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>VAFXGK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="3">

<INSDSeq>

<INSDSeq\_length>4</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>the N-terminus is acetylated</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>C-terminus linked via a glutaraldehyde bridge to dipeptide GK</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>VAFW</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="4">

<INSDSeq>

<INSDSeq\_length>5</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>REGION</INSDFeature\_key>

<INSDFeature\_location>&gt;5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>The entire sequence of amino acids 1-5 can be repeated one or more times</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GGGGX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="5">

<INSDSeq>

<INSDSeq\_length>12</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>cytosine attached to a C3 spacer, which is joined to another nucleic acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>atgcatgcatgc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="6">

<INSDSeq>

<INSDSeq\_length>12</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>cytosine attached to a C3 spacer, which is joined to another nucleic acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cggcatgcatgc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="7">

<INSDSeq>

<INSDSeq\_length>25</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..25</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Mus musculus</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>atgcatgcatgcvcggcatgcatgc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="8">

<INSDSeq>

<INSDSeq\_length>12</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>cytosine is linked to a C3 spacer, which is linked to 5&amp;apos;-end of another nucleic acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>atgcatgcatgc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="9">

<INSDSeq>

<INSDSeq\_length>12</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>cytosine is linked to a C3 spacer, which is linked to 3&amp;apos;-end of another nucleic acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cggcatgcatgc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="10">

<INSDSeq>

<INSDSeq\_length>18</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..18</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>modified\_base</INSDFeature\_key>

<INSDFeature\_location>11</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>mod\_base</INSDQualifier\_name>

<INSDQualifier\_value>OTHER</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>abasic site</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>gagcattgacntaaggct</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="11">

<INSDSeq>

<INSDSeq\_length>30</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..30</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>GNA sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>modified\_base</INSDFeature\_key>

<INSDFeature\_location>1..30</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>mod\_base</INSDQualifier\_name>

<INSDQualifier\_value>OTHER</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value> 2,3-dihydroxypropyl nucleosides (glycol nucleic acids)</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>tagttcattgactaaggctccccattgact</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="12">

<INSDSeq>

<INSDSeq\_length>10</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..10</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>aatgccggag</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="13">

<INSDSeq>

<INSDSeq\_length>10</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..10</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>aatgccggak</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="14">

<INSDSeq>

<INSDSeq\_length>10</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..10</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>aatgccggac</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="15">

<INSDSeq>

<INSDSeq\_length>10</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..10</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>aatgttggac</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="16">

<INSDSeq>

<INSDSeq\_length>15</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..15</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>nngkngkngkagvcr</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="17">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>XYEKGJL</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="18">

<INSDSeq>

<INSDSeq\_length>30</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..30</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>This sequence is one of four branches of a branched polynucleotide</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cacacaaaaaaaaaaaaaaaaaaaaaaaaa</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="19">

<INSDSeq>

<INSDSeq\_length>28</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..28</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>This sequence is one of four branches of a branched polynucleotide.</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cacataggcatctcctagtgcaggaaga</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="20">

<INSDSeq>

<INSDSeq\_length>75</INSDSeq\_length>

<INSDSeq\_moltype>RNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..75</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>tRNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>modified\_base</INSDFeature\_key>

<INSDFeature\_location>39</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>mod\_base</INSDQualifier\_name>

<INSDQualifier\_value>p</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>modified\_base</INSDFeature\_key>

<INSDFeature\_location>54</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>mod\_base</INSDQualifier\_name>

<INSDQualifier\_value>p</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>gcggatttagctcagctgggagagcgccagactgaatanctggagtcctgtgtncgatccacagaattcgcacca</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="21">

<INSDSeq>

<INSDSeq\_length>44</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..44</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>gatcattttttatattttttatattttttatattttttatgtac</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="22">

<INSDSeq>

<INSDSeq\_length>44</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..44</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>repeat\_region</INSDFeature\_key>

<INSDFeature\_location>5..40</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>rpt\_type</INSDQualifier\_name>

<INSDQualifier\_value>tandem</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>rpt\_unit\_seq</INSDQualifier\_name>

<INSDQualifier\_value>atacgcact</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>attcatacgcactatacgcactatacgcactatacgcactgtac</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="23">

<INSDSeq>

<INSDSeq\_length>44</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..44</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>repeat\_region</INSDFeature\_key>

<INSDFeature\_location>5..40</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>rpt\_type</INSDQualifier\_name>

<INSDQualifier\_value>tandem</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>rpt\_unit\_seq</INSDQualifier\_name>

<INSDQualifier\_value>atacgcact</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>gatcatacgcactatacgcactatacgcactatacgcactgtac</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="24">

<INSDSeq>

<INSDSeq\_length>12</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_difference</INSDFeature\_key>

<INSDFeature\_location>5..8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>A nucleotide in any of positions 5-8 may be replaced with inosine or pseudouridine.</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>modified\_base</INSDFeature\_key>

<INSDFeature\_location>5..8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>mod\_base</INSDQualifier\_name>

<INSDQualifier\_value>OTHER</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>i or p</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>gatcaaaagtac</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="25">

<INSDSeq>

<INSDSeq\_length>74</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..74</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>1..8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>10..19</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>21..48</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>50..68</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>70..74</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>XXXXXXXXDXXXXXXXXXXFXXXXXXXXXXXXXXXXXXXXXXXXXXXXAXXXXXXXXXXXXXXXXXXXGXXXXX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="26">

<INSDSeq>

<INSDSeq\_length>4</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1..4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>This sequence is one part of a branched amino acid sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Leu is linked to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>RISL</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="27">

<INSDSeq>

<INSDSeq\_length>4</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1..4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>This sequence is one part of a branched ~~peptide~~amino acid sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>3</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>The side chain of Lysine is linked via an amide linkage to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>The side chain of Lysine is linked via an amide linkage to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LLKK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="28">

<INSDSeq>

<INSDSeq\_length>6</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>This sequence is one part of a branched ~~peptide~~amino acid sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Ala is linked to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>IPACTA</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="29">

<INSDSeq>

<INSDSeq\_length>6</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>This sequence is one part of a branched ~~peptide~~amino acid sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>The side chain of Lysine is linked via an amide linkage to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>FRAGGK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="30">

<INSDSeq>

<INSDSeq\_length>5</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1..5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>This sequence is one part of a branched ~~peptide~~amino acid sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Ala is linked to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>HQYFA</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="31">

<INSDSeq>

<INSDSeq\_length>8</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1..8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>This sequence is one part of a branched ~~peptide~~amino acid sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>The side chain of Lysine is linked via an amide linkage to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>The side chain of Lysine is linked via an amide linkage to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>The side chain of Lysine is linked via an amide linkage to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>ATFGKKKA</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="32">

<INSDSeq>

<INSDSeq\_length>~~7~~8</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..~~7~~8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>~~Bound~~The side chain of Lysine is linked via an amide linkage ~~to Cly at position 5 of AASHG~~another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>~~DCSAKKK~~DGSAKKKK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="33">

<INSDSeq>

<INSDSeq\_length>5</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>~~Bound~~Gly is linked to ~~Lys at position 5 of DCSAKKK~~another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>AASHG</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="34">

<INSDSeq>

<INSDSeq\_length>13</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..13</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>atcgcaattggcc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="35">

<INSDSeq>

<INSDSeq\_length>13</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..13</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>ggccaattgcgat</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="36">

<INSDSeq>

<INSDSeq\_length>58</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..58</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>tagttcattgactaaggctccccattgactaaggcgactagcattgactaaggcaagc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="37">

<INSDSeq>

<INSDSeq\_length>16</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..16</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>modified\_base</INSDFeature\_key>

<INSDFeature\_location>~~6~~1..16</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>mod\_base</INSDQualifier\_name>

<INSDQualifier\_value>OTHER</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>N-(2-aminoethyl) glycine nucleosides (PNA)</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>modified\_base</INSDFeature\_key>

<INSDFeature\_location>6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>mod\_base</INSDQualifier\_name>

<INSDQualifier\_value>OTHER</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>N-(2-aminoethyl) glycine 5-nitroindole or N-(2-aminoethyl) glycine 3-nitroindole</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

~~<INSDFeature>~~

~~<INSDFeature\_key>modified\_base</INSDFeature\_key>~~

~~<INSDFeature\_location>1..16</INSDFeature\_location>~~

~~<INSDFeature\_quals>~~

~~<INSDQualifier> <INSDQualifier\_name>mod\_base</INSDQualifier\_name> <INSDQualifier\_value>OTHER</INSDQualifier\_value>~~

~~</INSDQualifier>~~

~~<INSDQualifier> <INSDQualifier\_name>note</INSDQualifier\_name>~~

~~<INSDQualifier\_value>N-(2-aminoethyl) glycine nucleosides (PNA)</INSDQualifier\_value>~~

~~</INSDQualifier>~~

~~</INSDFeature\_quals>~~

~~</INSDFeature>~~

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cgcctnagtcaatggg</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="38">

<INSDSeq>

<INSDSeq\_length>25</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..25</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>atcgggatcgcatattcgattggcc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="39">

<INSDSeq>

<INSDSeq\_length>25</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..25</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>ggccaatatggcttgcgatcccgat</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="40">

<INSDSeq>

<INSDSeq\_length>20</INSDSeq\_length>

<INSDSeq\_moltype>RNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..20</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>~~ccugucgt is attached at its 3&amp;apos;-end to a linker which is attached to the~~ The 5~~&amp;~~apos; oxygen of the thymidine~~. The~~ is attached through the linker ~~is~~(4-(3-hydroxybenzamido)butyl) phosphinic acid to ccugucgt</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>modified\_base</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>mod\_base</INSDQualifier\_name>

<INSDQualifier\_value>OTHER</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>thymine</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>tcctgtccggagatgttgat</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="41">

<INSDSeq>

<INSDSeq\_length>8</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GGGXGGGX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="42">

<INSDSeq>

<INSDSeq\_length>400</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..400</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>~~9~~4..400</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>9..400</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Up to 98 copies of GGGX can be deleted</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGXGGGX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="43">

<INSDSeq>

<INSDSeq\_length/>

<INSDSeq\_moltype/>

<INSDSeq\_division/>

<INSDSeq\_sequence>000</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="44">

<INSDSeq>

<INSDSeq\_length>8</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GGGXGGGX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="45">

<INSDSeq>

<INSDSeq\_length>12</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GGGXGGGXGGGX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="46">

<INSDSeq>

<INSDSeq\_length>306</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..306</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>CDS</INSDFeature\_key>

<INSDFeature\_location>19..114</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>translation</INSDQualifier\_name>

<INSDQualifier\_value>MAPNTQTASPRALADSLMQLARQVSRLESGQ</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>protein\_id</INSDQualifier\_name>

<INSDQualifier\_value>47</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>CDS</INSDFeature\_key>

<INSDFeature\_location>115..222</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>translation</INSDQualifier\_name>

<INSDQualifier\_value>MTMITDSLAVVLQRRDWENPGVTQLNRLAAHWCQK</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>protein\_id</INSDQualifier\_name>

<INSDQualifier\_value>48</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>CDS</INSDFeature\_key>

<INSDFeature\_location>251..283</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>translation</INSDQualifier\_name>

<INSDQualifier\_value>MLRRQVNEVA</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>protein\_id</INSDQualifier\_name>

<INSDQualifier\_value>49</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>caattcagggtggtgaatatggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatggaaagcgggcagtgaatgaccatgattacggattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacattggtgtcaaaaataataataaccggatgtactatttatccctgatgctgcgtcgtcaggtgaatgaagtcgcttaagcaatcaatgtcggatgcggcgcgacgcttatccgaccaacatatcataa</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="47">

<INSDSeq>

<INSDSeq\_length>31</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..31</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>MAPNTQTASPRALADSLMQLARQVSRLESGQ</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="48">

<INSDSeq>

<INSDSeq\_length>35</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..35</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>MTMITDSLAVVLQRRDWENPGVTQLNRLAAHWCQK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="49">

<INSDSeq>

<INSDSeq\_length>10</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..10</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>MLRRQVNEVA</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="50">

<INSDSeq>

<INSDSeq\_length>10</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..10</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>2</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Homoserine</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>3..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>A or G</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>AXXXXXYLGS</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="51">

<INSDSeq>

<INSDSeq\_length>17</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..17</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>DISULFID</INSDFeature\_key>

<INSDFeature\_location>join(4,15)</INSDFeature\_location>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>CARBOHYD</INSDFeature\_key>

<INSDFeature\_location>9</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Asn side-chain N-linked to asialyloligosaccharide</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEYCLKRWNETISHCAW</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="52">

<INSDSeq>

<INSDSeq\_length>106</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..106</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LLXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXKYMR</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="53">

<INSDSeq>

<INSDSeq\_length>53</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..53</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>~~30~~2..3</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>5..24</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>26..28</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>30</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>deleted</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>30..49</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>51..52</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>KXXKXXXXXXXXXXXXXXXXXXXXKXXXKXXXXXXXXXXXXXXXXXXXXKXXK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="54">

<INSDSeq>

<INSDSeq\_length>52</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..52</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>~~29^30~~2..3</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>5..24</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>26..28</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>29^30</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X is inserted</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>30..48</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>50..51</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>KXXKXXXXXXXXXXXXXXXXXXXXKXXXKXXXXXXXXXXXXXXXXXXXKXXK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="55">

<INSDSeq>

<INSDSeq\_length>63</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..63</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

~~<INSDFeature\_location>5..29</INSDFeature\_location>~~

<INSDFeature\_location>2..3</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>~~Up to 10 residues may~~ X can be ~~deleted~~ any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>~~31..59~~5..29</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>~~Up~~X can be any amino acid; up to 10 residues may be deleted</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>31..33</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>35..59</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid; up to 10 residues may be deleted</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>61..62</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>KXXKXXXXXXXXXXXXXXXXXXXXXXXXXKXXXKXXXXXXXXXXXXXXXXXXXXXXXXXKXXK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="56">

<INSDSeq>

<INSDSeq\_length>43</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..43</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

~~<INSDFeature\_location>5..~~ <INSDFeature\_location>2..3</INSDFeature\_location>

~~19</INSDFeature\_location>~~

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>~~Up to 10~~X ~~residues may~~ can be ~~inserted~~any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>~~21..39~~5..19</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>~~Up~~X can be any amino acid; up to 10 X residues may be inserted</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>21..23</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>25..39</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid; up to 10 X residues may be inserted</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>41..42</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>KXXKXXXXXXXXXXXXXXXKXXXKXXXXXXXXXXXXXXXKXXK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="57">

<INSDSeq>

<INSDSeq\_length>6</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>1</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>This residue is linked N-terminally to a peptide having an N-terminal Gly-Gly and a gap of undefined length</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>5..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GGAGXX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="58">

<INSDSeq>

<INSDSeq\_length>45</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..45</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>CDS</INSDFeature\_key>

<INSDFeature\_location>1..45</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>translation</INSDQualifier\_name>

<INSDQualifier\_value>MTGNKPETDVYEIL</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>protein\_id</INSDQualifier\_name>

<INSDQualifier\_value>59</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>atgaccggaaataaacctgaaaccgatgtttacgaaattttatga</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="59">

<INSDSeq>

<INSDSeq\_length>14</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..14</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>MTGNKPETDVYEIL</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="60">

<INSDSeq>

<INSDSeq\_length>22</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..22</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Drosophila melanogaster</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>acattgaatctcataccacttt</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="61">

<INSDSeq>

<INSDSeq\_length>19</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..19</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Drosophila virilis</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>acatggatcccacgacttt</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="62">

<INSDSeq>

<INSDSeq\_length>22</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..22</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Drosophila simulans</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>gtatggcgtcgtatsgtagttt</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="63">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X is S or T</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GGGJATX</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="64">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>replace with I</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>replace with T</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GGGLATS</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="65">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>X can be any amino acid</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GLPXRIC</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="66">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GLPVRIC</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="67">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GLPTRIC</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="68">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>GLPDRIC</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="69">

<INSDSeq>

<INSDSeq\_length>36</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..36</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cgaatgncccactacgaatgncacgaatgncccaca</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="70">

<INSDSeq>

<INSDSeq\_length>36</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..36</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cgaatgacccactacgaatgbcacgaatgbcccaca</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="71">

<INSDSeq>

<INSDSeq\_length>36</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..36</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cgaatgtcccactacgaatgvcacgaatgvcccaca</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="72">

<INSDSeq>

<INSDSeq\_length>36</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..36</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cgaatggcccactacgaatghcacgaatghcccaca</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="73">

<INSDSeq>

<INSDSeq\_length>36</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..36</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>cgaatgccccactacgaatgdcacgaatgdcccaca</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="74">

<INSDSeq>

<INSDSeq\_length>400</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..400</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>CDS</INSDFeature\_key>

<INSDFeature\_location>join(1..79,142..212,272..400)</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>protein\_id</INSDQualifier\_name>

<INSDQualifier\_value>75</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>translation</INSDQualifier\_name>

<INSDQualifier\_value>MKTFAALLSAVTLALSVRAQAAVWSQCGGTPGWTGETTCVAGSVCTSLSSSYSQCVPGSATSSAPAAPSATTSGPAPTDGTCSASGAWPPLT</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>atgaagactttcgcagccttgctttccgctgtcactctcgcgctctcggtgcgcgcccaggcggctgtctggagtcaatgtaagtgccgctgcttttcattgatacgagactctacgccgagctgacgtgctaccgtataggtggcggtacaccgggttggacgggcgagaccacttgcgttgctggttcggtttgtacctccttgagctcagtgagcgactttcaatccgtcgtcattgctcctcatgtattgacgattggccttcatagtcatactctcaatgcgttccgggctccgcaacgtccagcgctccggcggccccctcagcgacaacttcaggccccgcacctacggacggaacgtgctcggccagcggggcatggccgccattgacctga</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="75">

<INSDSeq>

<INSDSeq\_length>92</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..92</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>MKTFAALLSAVTLALSVRAQAAVWSQCGGTPGWTGETTCVAGSVCTSLSSSYSQCVPGSATSSAPAAPSATTSGPAPTDGTCSASGAWPPLT</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="76">

<INSDSeq>

<INSDSeq\_length>9</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..9</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>AVLTYLRGE</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="77">

<INSDSeq>

<INSDSeq\_length>9</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..9</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>AVLTYLRGA</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="78">

<INSDSeq>

<INSDSeq\_length>9</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..9</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>AVPTYPRGE</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="79">

<INSDSeq>

<INSDSeq\_length>9</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..9</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>AVAIGYRGE</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="80">

<INSDSeq>

<INSDSeq\_length>8</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..8</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>AVLTYLGE</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="81">

<INSDSeq>

<INSDSeq\_length>32</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..32</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>N or T</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>11</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>A or T</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>12</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>K or Q</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>18</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>K or Q or N or D</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>21</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>R or G</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>22</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>K or I or N</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>24</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>L or V</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>VARIANT</INSDFeature\_key>

<INSDFeature\_location>29</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>M or L</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEGXEQFINAXXIIRHPXYBXXTXBNDIXLIK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="82">

<INSDSeq>

<INSDSeq\_length>32</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..32</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEGNEQFINAAKIIRHPQYDRKTLNNDIMLIK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="83">

<INSDSeq>

<INSDSeq\_length>32</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..32</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Pongo abelii</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEGNEQFINAAKIIRHPQYDRKTVNNDIMLIK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="84">

<INSDSeq>

<INSDSeq\_length>32</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..32</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Papio anubis</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEGTEQFINAAKIIRHPDYDRKTLNNDILLIK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="85">

<INSDSeq>

<INSDSeq\_length>32</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..32</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Rhinopithecus roxellana</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEGTEQFINAAKIIRHPNYNRITLDNDILLIK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="86">

<INSDSeq>

<INSDSeq\_length>32</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..32</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Pan paniscus</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEGNEQFINAAKIIRHPKYNRITLNNDIMLIK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="87">

<INSDSeq>

<INSDSeq\_length>32</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..32</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Rhinopithecus bieti</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEGNEQFINATKIIRHPKYNGNTLNNDIMLIK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="88">

<INSDSeq>

<INSDSeq\_length>32</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..32</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Rhinopithecus roxellana</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LEGNEQFINATQIIRHPKYNGNTLNNDIMLIK</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="89">

<INSDSeq>

<INSDSeq\_length>7</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>join(1,7)</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Cysteine Leucine thioester (Cys-Leu)</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Lys side chain links the tripeptide ALE</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>CALRDKL</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="90">

<INSDSeq>

<INSDSeq\_length>5</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>5</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Ser is linked to another sequence through an amide linkage between Ser and a side chain of a Lys in the other sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>LRDQS</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="91">

<INSDSeq>

<INSDSeq\_length>6</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>REGION</INSDFeature\_key>

<INSDFeature\_location>1..6</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>Circular peptide</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>SITE</INSDFeature\_key>

<INSDFeature\_location>4</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>NOTE</INSDQualifier\_name>

<INSDQualifier\_value>The side chain of Lys is linked via an amide linkage to another sequence</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>ALFKNG</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="92">

<INSDSeq>

<INSDSeq\_length>43</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..43</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>synthetic construct</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>other DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>1..7</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>DNA fragment</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>8..42</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>RNA fragment</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>misc\_feature</INSDFeature\_key>

<INSDFeature\_location>43</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>note</INSDQualifier\_name>

<INSDQualifier\_value>DNA fragment</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>agaccttcggagtctcctgttgaacagatagtcaaagtagatc</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="93">

<INSDSeq>

<INSDSeq\_length>216</INSDSeq\_length>

<INSDSeq\_moltype>DNA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>source</INSDFeature\_key>

<INSDFeature\_location>1..216</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>organism</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>mol\_type</INSDQualifier\_name>

<INSDQualifier\_value>genomic DNA</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

<INSDFeature>

<INSDFeature\_key>CDS</INSDFeature\_key>

<INSDFeature\_location>&lt;1..&gt;216</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>protein\_id</INSDQualifier\_name>

<INSDQualifier\_value>94</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>translation</INSDQualifier\_name>

<INSDQualifier\_value>HHAAECGFCPQQWQVLRGSLCICEGPAEGWFISRCWLWCGPQVQGFIFGEGKEGGGDRRAEASPQEFWECTW</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>catcacgcagcagaatgtggattttgtcctcaacaatggcaagttctacgtgggagtctgtgcatttgtgagggtccagctgaaggatggttcatatcaagatgttggttatggtgtgggcctcaagtccaaggctttatctttggagaaggcaaggaaggaggcggtgacagacgggctgaagcgagccctcaggagttttgggaatgcacttgg</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

<SequenceData sequenceIDNumber="94">

<INSDSeq>

<INSDSeq\_length>72</INSDSeq\_length>

<INSDSeq\_moltype>AA</INSDSeq\_moltype>

<INSDSeq\_division>PAT</INSDSeq\_division>

<INSDSeq\_feature-table>

<INSDFeature>

<INSDFeature\_key>SOURCE</INSDFeature\_key>

<INSDFeature\_location>1..72</INSDFeature\_location>

<INSDFeature\_quals>

<INSDQualifier>

<INSDQualifier\_name>ORGANISM</INSDQualifier\_name>

<INSDQualifier\_value>Homo sapiens</INSDQualifier\_value>

</INSDQualifier>

<INSDQualifier>

<INSDQualifier\_name>MOL\_TYPE</INSDQualifier\_name>

<INSDQualifier\_value>protein</INSDQualifier\_value>

</INSDQualifier>

</INSDFeature\_quals>

</INSDFeature>

</INSDSeq\_feature-table>

<INSDSeq\_sequence>HHAAECGFCPQQWQVLRGSLCICEGPAEGWFISRCWLWCGPQVQGFIFGEGKEGGGDRRAEASPQEFWECTW</INSDSeq\_sequence>

</INSDSeq>

</SequenceData>

</ST26SequenceListing>

[Sigue el Anexo VII (Anexo VII de la Norma ST.26)]

1. NOTA: Mientras que en la divulgación de una solicitud pueden representarse los nucleótidos o los aminoácidos mediante símbolos en letra tanto minúscula como mayúscula, para una secuencia incluida en una lista de secuencias deberá utilizarse únicamente la letra minúscula para la representación de una secuencia de nucleótidos (véase el párrafo 13 de la Norma ST.26) y únicamente la letra mayúscula para la representación de una secuencia de aminoácidos (véase el párrafo 26 de la Norma ST.26). [↑](#footnote-ref-1)