**ST.26 – ANNEXE VII**

Recommandation relative à la conversion d’un listage des séquences de la norme ST.25 à la norme ST.26 : ajout ou suppression éventuels d’éléments

*Projet de proposition soumis au CWS/6 pour examen*

## INTRODUCTION

Les exigences en matière de présentation de séquences de nucléotides et d’acides aminés diffèrent entre la norme ST.26 et la norme ST.25 de l’OMPI. Il convient donc de déterminer si la norme ST.26 impose d’ajouter ou de supprimer des éléments dans un listage des séquences faisant partie intégrante d’une demande internationale déposée conformément à cette norme, étant entendu que ces éléments pourraient ne pas figurer dans une demande dont la priorité est revendiquée.

## PORTEE DU DOCUMENT

Le présent document concerne les exigences énoncées dans la norme ST.26 ainsi que toutes les conséquences qui pourraient en découler. Il ne traite pas individuellement de chaque scénario possible; si les moyens prévus dans la norme ST.26 pour représenter les informations figurant dans un listage des séquences selon la norme ST.25 ne sont pas suffisants, il est toujours possible de faire apparaître ces informations dans la description de la demande afin d’éviter la suppression d’éléments.

## Recommandations relatives à l’ajout ou la suppression éventuels d’éléments

L’examen des questions abordées dans le présent document montre que la conversion d’un listage des séquences de la norme ST.25 à la norme ST.26 ne devrait pas nécessiter par elle-même l’ajout ou la suppression d’éléments, surtout si le listage effectué au titre de la norme ST.25 était entièrement conforme à celle-ci. Néanmoins, dans certains cas le déposant devra être prudent. Des recommandations ont été établies pour éviter que des éléments ne soient ajoutés ou supprimés.

### Scénario 1

La norme ST.25 prévoit l’emploi d’identifiants numériques pour étiqueter différents types de données, par exemple <110> pour le nom du déposant. La norme ST.26 prévoit pour sa part l’emploi de termes anglais pour nommer et décrire les données.

#### Recommandation

Les termes prévus dans la norme ST.26 ne font que décrire le contenu des données; dès lors, l’emploi de ces noms et descripteurs ne constitue pas un ajout d’éléments.

### Scénario 2

La norme ST.26 fait obligation d’indiquer : a) les séquences ramifiées; b) les séquences comportant des acides aminés D; c) les analogues nucléotidiques; et d) les séquences comportant des sites abasiques. L’obligation d’indiquer ces séquences ou leur interdiction n’est pas clairement établie dans la norme ST.25.

#### Recommandation

Les informations divulguées dans la demande devraient être suffisantes pour permettre de représenter ces séquences dans un listage des séquences conforme à la norme ST.26, alors qu’elles n’auraient peut-être pas été indiquées dans un listage conforme à la norme ST.25. Pour certains types d’informations exigés par la norme ST.26, il faut veiller à ne pas ajouter d’éléments au-delà de ce qui a déjà été divulgué (voir par exemple l’analyse du scénario 4 concernant le qualificateur “mol\_type” destiné aux séquences de nucléotides).

### Scénario 3

La norme ST.26 interdit les séquences comportant moins de 10 nucléotides définis de manière spécifique (“n” exclu) et moins de quatre acides aminés définis de manière spécifique (“X” exclu).

#### Recommandation

Les séquences interdites peuvent être ajoutées dans le corps de la demande si elles n’y figurent pas déjà.

### Scénario 4

La norme ST.26 impose d’employer les clés de caractérisation “source” pour toutes les séquences de nucléotides et “SOURCE” pour toutes les séquences d’acides aminés, chaque clé étant assortie de deux qualificateurs obligatoires. La norme ST.25 comporte une clé de caractérisation correspondante pour les séquences de nucléotides (qui est rarement employée) sans les qualificateurs correspondants, et ne comporte aucune clé de caractérisation correspondante pour les séquences d’acides aminés.

Séquences de nucléotides

ST.26 – clé de caractérisation 5.37 “source”; qualificateurs obligatoires 6.44 “organism” et 6.38 “mol\_type” (voir le paragraphe 75 de la norme ST.26)

|  |  |
| --- | --- |
| **Qualificateur** | **Valeur** |
| mol\_type | genomic DNA |
| genomic RNA |
| mRNA |
| tRNA |
| rRNA |
| other DNA (applies to synthetic molecules) |
| other RNA (applies to synthetic molecules) |
| transcribed RNA |
| viral cRNA |
| unassigned DNA (applies where *in vivo* molecule is unknown) |
| unassigned RNA (applies where *in vivo* molecule is unknown) |

Séquences d’acides aminés

ST.26 – clé de caractérisation 7.30 “SOURCE”; qualificateurs obligatoires 8.3 “ORGANISM” et 8.1 “MOL\_TYPE” (voir le paragraphe 75 de la norme ST.26)

|  |  |
| --- | --- |
| **Qualificateur** | **Valeur** |
| MOL\_TYPE | protein |

#### Recommandation

La seule difficulté tient aux valeurs du vocabulaire contrôlé liées au qualificateur “mol\_type” dans les séquences de nucléotides. Certaines des valeurs indiquées plus haut peuvent ne pas être entièrement compatibles avec la divulgation. Néanmoins, on peut éviter d’ajouter des éléments si l’on emploie la valeur la plus générique d’une séquence particulière, par exemple “other DNA” et “other RNA” pour une molécule synthétique, ou encore “unassigned DNA” et “unassigned RNA” pour une molécule *in vivo*.

### Scénario 5

Lorsqu’une séquence comporte le symbole “Xaa”, la norme ST.25 prévoit l’ajout d’autres informations concernant ce résidu dans le champ <223>, qui est annexé aux champs <221> (Nom/clé) et <222> (Emplacement). Cette norme ne définit pas de valeur par défaut pour le symbole “Xaa” (appelé “X” dans la norme ST.26). En revanche, la norme ST.26 établit cette valeur par défaut; c’est pourquoi il n’est pas toujours nécessaire de fournir des informations supplémentaires. Les annotations “tout acide aminé” et “tout acide aminé naturel” renseignent très souvent les variables “Xaa” ou “X” dans les séquences de peptides. On pourrait considérer que ces termes recouvrent des acides aminés différents de ceux qui sont énumérés dans les tableaux d’acides aminés figurant dans les normes ST.25 ou ST.26. Dans la norme ST.26, le symbole “X” sans autre annotation prend par défaut la valeur de n’importe quel acide aminé parmi les 22 qui sont énumérés dans l’annexe I (voir le tableau 3 de la section 3). Cette valeur par défaut peut en elle-même constituer un élément ajouté ou supprimé, et avoir par conséquent un effet néfaste sur la portée de la demande de brevet lors de la conversion de la norme ST.25 à la norme ST.26.

#### Recommandations

1. Si le listage des séquences selon la norme ST.25 comprend des champs <221> (Nom/clé), <222> (Emplacement correspondant au symbole “Xaa”) et <223> (Autres informations sur le symbole “Xaa”), et si le champ <221> (Nom/clé) constitue aussi une clé de caractérisation adéquate selon la norme ST.26, en correspondant par exemple aux clés “SITE”, “VARIANT” ou “UNSURE”, il convient d’utiliser la clé de caractérisation de la norme ST.26. En outre, pour éviter toute suppression éventuelle d’éléments, les informations indiquées dans le champ <223> doivent aussi apparaître dans le qualificateur “NOTE” annexé au listage.
2. Si le listage des séquences selon la norme ST.25 comprend des champs <221> (Nom/clé), <222> (Emplacement correspondant au symbole “Xaa”) et <223> (Autres informations sur le symbole “Xaa”), et si le champ <221> (Nom/clé) ne constitue pas de clé de caractérisation selon la norme ST.26, il convient d’utiliser la clé de caractérisation “SITE” ou “REGION” de la norme ST.26, selon les besoins. En outre, pour éviter toute suppression éventuelle d’éléments, les informations indiquées dans le champ <223> et le nom inadéquat du champ <221> doivent aussi apparaître dans le qualificateur “NOTE” annexé au listage. Ainsi, tout listage selon la norme ST.25 employant un nom qui n’est pas prévu dans la norme ST.25 ou ST.26, par exemple <221> (Variable), et s’accompagnant d’autres informations dans le champ <223> (Xaa) décrit un acide aminé. Dans le présent exemple, la valeur du qualificateur “NOTE” selon la norme ST.26 serait la suivante : “Variable – Xaa est un acide aminé”.
3. Si le listage des séquences selon la norme ST.25 ne contient aucun champ <221>, <222>, ou <223> correspondant au symbole “Xaa”, ou si des champs <221> et <222> correspondant au symbole “Xaa” apparaissent mais qu’aucune information ne figure dans un champ <223> correspondant (aucun de ces deux scénarios n’est conforme à la norme ST.25, mais ils se sont néanmoins produits), toute information indiquée dans le corps de la demande pour renseigner le symbole “Xaa” devrait apparaître dans le qualificateur “NOTE” conforme à la norme ST.26 et s’accompagner d’une clé de caractérisation adéquate, par exemple “SITE”, “REGION” ou “UNSURE”, ainsi que d’une indication de l’emplacement.

### Scénario 6

Dans la norme ST.25, l’uracile est représenté dans la séquence par le symbole “u” et la thymine par le symbole “t”. Dans la norme ST.26, l’uracile et la thymine sont tous deux représentés dans la séquence par le symbole “t” sans autre annotation, le symbole “t” représentant l’uracile dans l’ARN et la thymine dans l’ADN.

#### Recommandations

1. Lorsqu’une séquence d’ADN contient de l’uracile, celui-ci est considéré, dans le contexte de la norme ST.26, comme un nucléotide modifié. Dès lors, la norme ST.26 fait obligation de représenter l’uracile par un “t”, de lui adjoindre une description supplémentaire dans la clé de caractérisation “modified\_base”, et de donner au qualificateur “mod\_base” la valeur “OTHER” et au qualificateur “note” la valeur “uracil”. Cette annotation selon la norme ST.26 n’est pas considérée comme un élément ajouté si la séquence d’ADN décrite selon la norme ST.25 contenait un symbole “u”.
2. Lorsqu’une séquence d’ARN contient de la thymine, celle-ci est considérée, dans le contexte de la norme ST.26, comme un nucléotide modifié. Dès lors, la norme ST.26 fait obligation de représenter la thymine par un “t”, de lui adjoindre une description supplémentaire dans la clé de caractérisation “modified\_base”, et de donner au qualificateur “mod\_base” la valeur “OTHER” et au qualificateur “note” la valeur “thymine”. Cette annotation selon la norme ST.26 n’est pas considérée comme un élément ajouté si la séquence d’ARN décrite selon la norme ST.25 contenait un symbole “t”.

### Scénario 7

Les normes ST.25 et ST.26 prévoient toutes deux que des nucléotides modifiés ou des acides aminés doivent être accompagnés d’une description supplémentaire. Dans la norme ST.26, l’identité d’un nucléotide modifié peut être indiquée, le cas échéant, par l’une des abréviations figurant dans le tableau 2 de la section 2 de l’annexe I; si tel n’est pas le cas, le nom complet non abrégé du nucléotide modifié doit être indiqué. De même, l’identité d’un acide aminé modifié peut être indiquée, le cas échéant, par l’une des abréviations figurant dans le tableau 4 de la section 4 de l’annexe I; si tel n’est pas le cas, le nom complet non abrégé de l’acide aminé modifié doit être indiqué. Inversement, si un résidu modifié ne figure pas dans un tableau de la norme ST.25, il n’est pas obligatoire d’indiquer son nom complet non abrégé; il n’est pas rare qu’une abréviation soit alors employée.

#### Recommandations

1. Lorsque seul un nom abrégé n’apparaissant ni dans le tableau 2 de la section 2 de l’annexe I ni dans le tableau 4 de la section 4 de l’annexe I a été employé à la fois dans la demande et dans un listage des séquences selon la norme ST.25 pour un nucléotide modifié ou un acide aminé modifié, et que l’homme du métier sait que ce nom abrégé ne désigne qu’un seul nucléotide modifié ou acide aminé modifié bien précis, l’emploi du nom complet non abrégé ne constitue pas en lui-même un élément ajouté.
2. Lorsque seul un nom abrégé n’apparaissant ni dans le tableau 2 de la section 2 de l’annexe I ni dans le tableau 4 de la section 4 de l’annexe I a été employé à la fois dans la demande et dans un listage des séquences selon la norme ST.25 pour un nucléotide modifié ou un acide aminé modifié (et que la demande ne comporte aucune structure chimique), et que l’homme du métier ne sait pas que ce nom abrégé ne désigne qu’un seul nucléotide modifié ou acide aminé modifié bien précis, soit parce que l’abréviation lui est totalement inconnue, soit parce qu’elle pourrait représenter différents nucléotides modifiés ou acides aminés modifiés, il est impossible de se conformer à la norme ST.26 sans ajouter un élément. Bien entendu, dans ce cas particulier la demande établissant une priorité et le listage des séquences sont eux-mêmes vagues. Pour éviter une éventuelle suppression d’éléments, le nom abrégé figurant dans le listage des séquences selon la norme ST.25 devrait être placé dans un qualificateur “note” ou “NOTE” conforme à la norme ST.26 en plus du nom complet non abrégé du nucléotide modifié ou de l’acide aminé modifié. Ce nom complet non abrégé, qui est prévu dans les listages des séquences définis par la norme ST.26, n’aura pas la priorité par rapport à la demande antérieure. Il faut veiller à faire apparaître le nom non abrégé dans le listage des séquences original (selon la norme ST.25) et dans la divulgation de la demande afin d’éviter des problèmes par la suite.

### Scénario 8

La norme ST.25 comporte un certain nombre de clés de caractérisation qui n’existent pas dans la norme ST.26. Les déposants doivent donc veiller à reprendre les informations figurant dans ces clés de caractérisation d’une manière conforme à la norme ST.26 sans avoir à ajouter ou à supprimer d’éléments.

#### Recommandations

Le tableau suivant fournit des orientations quant à la manière de reprendre les informations figurant dans une ancienne clé de caractérisation de la norme ST.25 en respectant la norme ST.26 et sans avoir à ajouter ou à supprimer d’éléments. Les numéros 1 à 23 correspondent à des clés de caractérisation de séquences de nucléotides, et les numéros 24 à 43 correspondent à des clés de séquences d’acides aminés.

| **N°** | **Clé de caractérisation du champ <221> dans la norme ST.25** | **Équivalent dans la norme ST.26** | | |
| --- | --- | --- | --- | --- |
| **Clé de caractérisation** | **Qualificateur** | **Valeur du qualificateur** |
| 1 | allele | misc\_feature | allele | valeur du champ <223> |
| 2 | attenuator | regulatory[[1]](#footnote-2) | regulatory\_class1 | “attenuator” |
| note *(if <223> present)* | valeur du champ <223> |
| 3 | CAAT\_signal | regulatory1 | regulatory\_class1 | “CAAT\_signal” |
| note *(if <223> present)* | valeur du champ <223> |
| 4 | conflict | misc\_feature | note | “conflict” et valeur du champ <223> |
| 5 | enhancer | regulatory1 | regulatory\_class1 | “enhancer” |
| note *(if <223> present)* | valeur du champ <223> |
| 6 | GC\_signal | regulatory1 | regulatory\_class1 | “GC\_signal” |
| note *(if <223> present)* | valeur du champ <223> |
| 7 | LTR | mobile\_element1 | rpt\_type1 | “long\_terminal\_repeat” |
| note *(if <223> present)* | valeur du champ <223> |
| 8 | misc\_signal | regulatory1 | regulatory\_class1 | “other” |
| note *(if <223> present)* | valeur du champ <223> |
| 9 | mutation | variation | note | “mutation” et valeur du champ <223> |
| 10 | old\_sequence | misc\_feature | note | “old\_sequence” et valeur du champ <223> |
| 11 | polyA\_signal | regulatory1 | regulatory\_class1 | “polyA\_signal\_sequence” |
| note *(if <223> present)* | valeur du champ <223> |
| 12 | promoter | regulatory1 | regulatory\_class1 | “promoter” |
| note *(if <223> present)* | valeur du champ <223> |
| 13 | RBS | regulatory1 | regulatory\_class1 | “ribosome\_binding\_site” |
| note *(if <223> present)* | valeur du champ <223> |
| 14 | repeat\_unit a) when repeat\_region not used | misc\_feature | note | “repeat\_unit” et valeur du champ <223> |
| repeat\_unit b) when repeat\_region used | repeat\_region | rpt\_unit\_range | 1st residue..last residue |
| note *(if <223> present)* | valeur du champ <223> |
| 15 | satellite | repeat\_region | satellite | “satellite” (ou “microsatellite” ou  “minisatellite” – si pris en charge) |
| note *(if <223> present)* | valeur du champ <223> |
| 16 | scRNA | ncRNA1 | ncRNA\_class1 | “scRNA” |
| note *(if <223> present)* | valeur du champ <223> |
| 17 | snRNA | ncRNA1 | ncRNA\_class1 | “snRNA” |
| note *(if <223> present)* | valeur du champ <223> |
| 18 | TATA\_signal | regulatory1 | regulatory\_class1 | “TATA\_box” |
| note *(if <223> present)* | valeur du champ <223> |
| 19 | terminator | regulatory1 | regulatory\_class1 | “terminator” |
| note *(if <223> present)* | valeur du champ <223> |
| 20 | 3’clip | misc\_feature | note | “3’clip” et valeur du champ <223> |
| 21 | 5’clip | misc\_feature | note | “5’clip” et valeur du champ <223> |
| 22 | -10\_signal | regulatory1 | regulatory\_class1 | “minus\_10\_signal” |
| note *(if <223> present)* | valeur du champ <223> |
| 23 | -35\_signal | regulatory1 | regulatory\_class1 | “minus\_35\_signal” |
| note *(if <223> present)* | valeur du champ <223> |
| 24 | NON\_CONS | Cette caractéristique décrit une brèche composée d’un nombre inconnu de résidus dans une même séquence, ce qui est interdit aussi bien par la norme ST.25  (paragraphe 22) que par la norme ST.26 (paragraphe 37). Dès lors, il convient d’intégrer chaque région composée de résidus définis de manière spécifique et visée par le paragraphe 7 de la norme ST.26 dans le listage des séquences à titre de séquence distincte, et de lui attribuer son propre numéro d’identification de séquence. Pour éviter d’ajouter ou de supprimer un élément, chacune de ces séquences doit être annotée pour indiquer qu’elle fait partie d’une séquence plus longue comportant une brèche indéfinie. | | |
| SITE | NOTE | Description |
| Description de l’emplacement et de la cible de la liaison de la séquence, par exemple “Ce résidu est relié par une liaison N terminale à un peptide comportant un Gly Gly N terminal et une brèche de longueur indéfinie”. | | |
| 25 | SIMILAR | REGION | NOTE | “SIMILAR” et  valeur du champ <223> si présent |
| 26 | THIOETH | CROSSLNK | NOTE | “THIOETH” et  valeur du champ <223> si présent |
| Pour plus de détails sur les informations d’emplacement, voir le commentaire sur la clé de caractérisation CROSSLNK dans l’annexe I de la norme ST.26. | | |
| 27 | THIOLEST | CROSSLNK | NOTE | “THIOLEST” et  valeur du champ <223> si présent |
| Pour plus de détails sur les informations d’emplacement, voir le commentaire sur la clé de caractérisation CROSSLNK dans l’annexe I de la norme ST.26. | | |
| 28 | VARSPLIC | Analysé dans le scénario 13 | | |
| 29 | ACETYLATION | MOD\_RES | NOTE | “ACETYLATION” et  valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 30 | AMIDATION | MOD\_RES | NOTE | “AMIDATION” et  valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 31 | BLOCKED | MOD\_RES | NOTE | “BLOCKED” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 32 | FORMYLATION | MOD\_RES | NOTE | “FORMYLATION” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 33 | GAMMA-  CARBOXYGLUTAMIC  ACID  HYDROXYLATION | MOD\_RES | NOTE | “GAMMA-CARBOXYLGLUTAMIC ACID HYDROXYLATION” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 34 | METHYLATION | MOD\_RES | NOTE | “METHYLATION” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 35 | PHOSPHORYLATION | MOD\_RES | NOTE | “PHOSPHORYLATION” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 36 | PYRROLIDONE  CARBOXYLIC ACID | MOD\_RES | NOTE | “PYRROLIDONE CARBOXYLIC ACID” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 37 | SULFATATION | MOD\_RES | NOTE | “SULFATATION” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “MOD\_RES” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 38 | MYRISTATE | LIPID | NOTE | “MYRISTATE” et  valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “LIPID” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 39 | PALMITATE | LIPID | NOTE | “PALMITATE” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “LIPID” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 40 | FARNESYL | LIPID | NOTE | “FARNESYL” et  valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “LIPID” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 41 | GERANYL-GERANYL | LIPID | NOTE | “GERANYL-GERANYL” et  valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “LIPID” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 42 | GPI-ANCHOR | LIPID | NOTE | “GPI-ANCHOR” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “LIPID” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |
| 43 | N-ACYL  DIGLYCERIDE | LIPID | NOTE | “N-ACYL DIGLYCERIDE” et valeur du champ <223> si présent |
| NOTE | Informations requises si possible en vertu du commentaire sur la clé de caractérisation “LIPID” figurant dans l’annexe I de la norme ST.26 (sans ajouter d’élément) |

### Scénario 9

Certaines clés de caractérisation présentes à la fois dans les normes ST.25 et ST.26, tant pour les séquences de nucléotides que pour les séquences d’acides aminés, ont des qualificateurs obligatoires dans la norme ST.26, comme indiqué plus loin. La norme ST.25 ne prévoit pas de qualificateur, mais elle comporte un champ de texte libre portant le numéro <223>. Si les informations figurant dans le champ <223> de la norme ST.25 sont compatibles avec la valeur du qualificateur obligatoire de la norme ST.26, elles doivent être reprises telles quelles. Si le champ <223> de la norme ST.25 est absent ou contient des informations qui ne sont pas compatibles avec la valeur du qualificateur obligatoire de la norme ST.26, le déposant doit veiller à reprendre les informations figurant dans la clé de caractérisation ou dans le champ <223> de la norme ST.25 d’une manière qui soit conforme à la norme ST.26 sans ajouter ni supprimer d’élément.

Séquences de nucléotides[[2]](#footnote-3)

|  |  |
| --- | --- |
| **Clé de caractérisation** | **Qualificateur obligatoire** |
| 5.12 – misc\_binding | 6.3 – bound\_moiety |
| 5.30 – protein\_bind | 6.3 – bound\_moiety |

#### Recommandations

1. Si le champ <223> de la norme ST.25 est absent ou incompatible et que la description de la demande révèle le nom de la molécule ou du complexe susceptible d’assurer la liaison avec l’emplacement de la caractéristique de l’acide nucléique, ce nom doit être repris dans le qualificateur “bound\_moiety”.
2. Toute information figurant dans le champ <223> de la norme ST.25 qui est incompatible avec le qualificateur “bound\_moiety” doit être reprise dans un qualificateur facultatif adéquat de la clé de caractérisation, par exemple “note”.
3. Si le champ <223> de la norme ST.25 est absent ou incompatible et que la description de la demande révèle le nom de la molécule ou du complexe susceptible d’assurer la liaison avec l’emplacement de la caractéristique de l’acide nucléique, il convient d’employer la clé de caractérisation “misc\_feature” de la norme ST.26 au lieu de la clé “misc\_binding” ou “protein\_bind”, ainsi que le qualificateur “note”.
4. Si le champ <223> de la norme ST.25 est absent, la valeur du qualificateur “note” doit être le nom de la clé de caractérisation de cette norme;
5. Si le champ <223> de la norme ST.25 contient des informations incompatibles, la valeur du qualificateur “note” doit être le nom de la clé de caractérisation de cette norme auquel s’ajoutent les informations du champ <223>.

Séquences d’acides aminés2

|  |  |
| --- | --- |
| **Clé de caractérisation** | **Qualificateur obligatoire** |
| 7.2 – BINDING | 8.2 – NOTE |
| 7.4 – CARBOHYD | 8.2 – NOTE |
| 7.10 – DISULFID | 8.2 – NOTE |
| 7.11 – DNA\_BIND | 8.2 – NOTE |
| 7.12 – DOMAIN | 8.2 – NOTE |
| 7.16 – LIPID | 8.2 – NOTE |
| 7.17 – METAL | 8.2 – NOTE |
| 7.18 – MOD\_RES | 8.2 – NOTE |
| 7.23 – NP\_BIND | 8.2 – NOTE |
| 7.29 – SITE | 8.2 – NOTE |
| 7.39 – ZN\_FING | 8.2 – NOTE |

#### Recommandations

1. Si le champ <223> de la norme ST.25 est absent ou incompatible et que la description de la demande révèle les informations spécifiques devant figurer dans le qualificateur obligatoire, ces informations doivent être reprises dans le qualificateur obligatoire “NOTE”.
2. Toute information figurant dans le champ <223> de la norme ST.25 qui est incompatible avec le qualificateur obligatoire “NOTE” (voir la définition et les commentaires de la clé de caractérisation) doit être reprise dans un second qualificateur “NOTE”.
3. Si le champ <223> de la norme ST.25 est absent ou incompatible et que la description de la demande ne révèle pas les informations spécifiques devant figurer dans le qualificateur obligatoire, il convient d’employer plutôt la clé de caractérisation “SITE” de la norme ST.26 (pour un seul acide aminé) ou “REGION” (pour une série d’acides aminés) ainsi que le qualificateur “NOTE”.
4. Si le champ <223> de la norme ST.25 est absent, la valeur du qualificateur “NOTE” doit être le nom de la clé de caractérisation de cette norme;
5. Si le champ <223> de la norme ST.25 contient des informations incompatibles, la valeur du qualificateur “NOTE” doit être le nom de la clé de caractérisation de cette norme auquel s’ajoutent les informations du champ <223>.

### Scénario 10

Chaque clé de caractérisation spécifique de la norme ST.25 comporte un champ <222> permettant d’indiquer un emplacement; cependant, d’une part cette norme ne fait pas obligation d’indiquer l’emplacement pour la plupart des caractéristiques, et d’autre part le format de ces informations d’emplacement n’est pas normalisé. Au demeurant, la norme ST.25 ne comporte pas d’opérateur d’emplacement tel que “join”. La norme ST.26 comporte quant à elle des descripteurs et des opérateurs d’emplacement, chacune de ses caractéristiques devant comporter au moins un descripteur d’emplacement. (Les caractéristiques CDS constituent un cas particulier examiné dans le scénario 11).

#### Recommandations

1. Si le listage des séquences selon la norme ST.25 comporte un champ <222>, l’importation directe et la conversion au format ST.26 ne devraient pas nécessiter l’ajout d’éléments;
2. Si le listage des séquences selon la norme ST.25 ne comporte pas de champ <222>, mais que les informations sur l’emplacement figurent dans la description de la demande, l’importation directe et la conversion au format ST.26 ne devraient pas nécessiter l’ajout d’éléments;
3. Si ni le listage des séquences selon la norme ST.25 ni la description de la demande ne contiennent d’informations sur l’emplacement, la caractéristique s’applique par hypothèse à l’ensemble de la séquence. (Le fait d’indiquer un emplacement moins long que l’ensemble de la séquence sans le justifier dans la description de la demande constituerait probablement un élément ajouté ou supprimé.) Il faut veiller à faire apparaître, dans la mesure du possible, les informations d’emplacement dans le listage des séquences original (selon la norme ST.25) et dans la divulgation de la demande afin d’éviter des problèmes par la suite.

### Scénario 11

Selon la norme ST.25, toute séquence codant un polypeptide unique et continu mais ayant été interrompue par une ou plusieurs séquences non codantes (par exemple des introns) apparaît sous la forme de plusieurs caractéristiques CDS, de la manière suivante :

<220>

<221> CDS

<222> (1)..(571)

<220>

<221> CDS

<222> (639)..(859)

La norme ST.26, pour sa part, dispose d’un opérateur d’emplacement “join” qui permet de préciser que les polypeptides codés par les emplacements indiqués sont joints et forment un polypeptide unique et continu. (Note : les normes ST.25 et ST.26 imposent toutes deux que le codon d’arrêt figure dans l’emplacement de la caractéristique CDS.)

#### Recommandations

1. Si le listage des séquences selon la norme ST.25 ou la description de la demande indique clairement que les séquences de polypeptides codées par plusieurs caractéristiques CDS distinctes constituent un polypeptide unique et continu, toute séquence de codage interrompue par un intron dans une caractéristique CDS particulière doit être représentée par l’opérateur d’emplacement “join”, de la manière décrite ci-après, afin de n’ajouter ou de ne supprimer aucun élément :

<INSDFeature\_key>CDS</INSDFeature\_key>

<INSDFeature\_location>join(1..571,639..859)</INSDFeature\_location>

1. Si le listage des séquences selon la norme ST.25 ou la description de la demande n’indique pas que les séquences de polypeptides codées par les deux caractéristiques CDS distinctes constituent un polypeptide unique et continu, l’emploi de l’opérateur d’emplacement “join” constituera probablement un ajout d’élément.

### Scénario 12

La norme ST.25 dispose que le nom des caractéristiques doit être repris des tableaux 5 ou 6. Toutefois, selon la réglementation des États-Unis d’Amérique, il s’agit d’une recommandation et non d’une obligation. Dès lors, toute séquence d’un listage établi selon la norme ST.25 (et conforme à la réglementation des États-Unis d’Amérique) peut comporter une clé de caractérisation dont le nom est “Custom” et qui n’a pas de correspondance dans la norme ST.26. Il est également possible qu’aucun nom de caractéristique n’ait été indiqué dans le champ <221>, ou que ce champ soit absent. Ces autres scénarios peuvent être gérés de la même manière.

#### Recommandation

Le nom de la clé de caractérisation “Custom” de la norme ST.25 peut apparaître dans un listage des séquences selon la norme ST.26 sans ajouter d’élément, de la manière suivante :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type | Clé de caractérisation  du champ <221> dans la norme ST.25 | Équivalent possible dans la norme ST.26 | | |
| Clé de caractérisation | Qualificateur | Valeur du qualificateur |
| NA | Clé de caractérisation “Custom” | misc\_feature | note | Nom de la clé de caractérisation “Custom” et valeur du champ <223> si présent |
| AA | Clé de caractérisation “Custom” | SITE ou REGION | NOTE | Nom de la clé de caractérisation “Custom” et valeur du champ <223>  si présent |

### Scénario 13

La norme ST.25 comporte la clé de caractérisation “VARSPLIC” qui est définie comme une “description des variants de la séquence produits par un épissage alternatif”. Dans la norme ST.26, la clé “VARSPLIC” a été remplacée par une clé de caractérisation plus générale appelée “VAR\_SEQ” et définie comme une “description des variants de la séquence produits par un épissage alternatif, l’usage de promoteurs alternatifs, une initiation alternative et un déphasage ribosomique”. Il ne faut donc pas utiliser la clé “VAR\_SEQ” pour remplacer la clé “VARSPLIC” sans ajouter d’explication supplémentaire dans un listage des séquences conforme à la norme ST.26.

#### Recommandation

Selon la norme ST.26, la caractéristique “VAR\_SEQ” doit être employée conjointement avec le qualificateur “NOTE”, dont la valeur doit indiquer que la portée était plus restreinte dans la norme ST.25, par exemple avec la mention “Variant de la séquence produit par un épissage alternatif”. Toute information supplémentaire figurant dans un champ <223> de la norme ST.25 devrait aussi être reprise dans le qualificateur “NOTE”.

### Scénario 14

Si la source de la séquence est artificielle, il convient de porter la mention “Artificial Sequence” dans le champ <213> (Organisme) de la norme ST.25. Selon la norme ST.26, la clé de caractérisation “source” ou “SOURCE” doit comporter le qualificateur “organism” ou “ORGANISM”, dont la valeur doit être “synthetic construct” et non “Artificial Sequence”.

#### Recommandation

La valeur du qualificateur “organism” ou “ORGANISM” de la norme ST.26 doit être “synthetic construct”. Pour éviter une éventuelle suppression d’élément, toute explication figurant dans le champ obligatoire <213> de la norme ST.25 doit apparaître dans un qualificateur “note” ou “NOTE” (de la clé de caractérisation “source” ou “SOURCE”).

### Scénario 15

Si le nom scientifique de l’organisme source d’une séquence est inconnu, il faut porter la mention “Unknown” dans le champ <213> (Organisme) de la norme ST.25. Selon la norme ST.26, la clé de caractérisation “source” ou “SOURCE” doit comporter le qualificateur “organism” ou “ORGANISM”, dont la valeur doit être “unidentified” et non “Unknown”.

#### Recommandation

La valeur du qualificateur “organism” ou “ORGANISM” selon la norme ST.26 doit être “unidentified”. Pour éviter une éventuelle suppression d’élément, toute explication figurant dans le champ obligatoire <223> de la norme ST.25 doit apparaître dans un qualificateur “note” ou “NOTE” (de la clé de caractérisation “source” ou “SOURCE”).

### Scénario 16

Pour les acides aminés précédant la protéine mature, par exemple les préséquences, les proséquences et les pré-proséquences ainsi que les séquences signal, la norme ST.25 permet d’ajouter à la liste des acides aminés, à titre facultatif, des numéros négatifs en comptant à rebours et en commençant par l’acide aminé voisin de l’acide portant le numéro 1. La norme ST.26 ne permet pas d’utiliser des numéros négatifs dans l’emplacement de la caractéristique.

#### Recommandations

1. Si le listage des séquences selon la norme ST.25 comporte une ou plusieurs caractéristiques dans un champ <221> et un champ <222> annexe, et que ces caractéristiques s’accompagnent d’une numérotation négative ou positive, par exemple “PROPEP” ou “CHAIN”, il convient d’employer la clé de caractérisation adéquate dans le listage des séquences selon la norme ST.26, par exemple “PROPEP” ou “CHAIN”. On peut employer un qualificateur “NOTE” pour reprendre les informations d’un champ <223>, le cas échéant;
2. Si le listage des séquences selon la norme ST.25 ne comporte pas de caractéristique dans un champ <221> et un champ <222> annexe, mais que la description de la demande contient des informations relatives à une numérotation négative ou positive, il convient d’employer la clé de caractérisation adéquate dans le listage des séquences selon la norme ST.26, par exemple “PROPEP” ou “CHAIN”. Une autre possibilité consiste à utiliser la clé de caractérisation “REGION”. On peut employer un qualificateur “NOTE” pour reprendre les informations figurant dans la description de la demande, le cas échéant;
3. Si ni le listage des séquences selon la norme ST.25 ni la description de la demande ne contiennent d’informations sur une numérotation négative ou positive, pour éviter une éventuelle suppression d’élément dans le listage des séquences selon la norme ST.26, il convient d’utiliser la clé de caractérisation “REGION” lorsque l’emplacement de la caractéristique recouvre la région portant un numéro négatif dans la séquence conforme à la norme ST.25. Il faut aussi utiliser un qualificateur “NOTE” pour indiquer que la séquence d’acides aminés portait un numéro négatif dans le listage des séquences établi selon la norme ST.25 aux fins de la demande dont la priorité est revendiquée.

### Scénario 17

La norme ST.25 permet de fournir des informations concernant la publication dans les champs <300> à <313>. La norme ST.26 ne permet pas de communiquer ces informations.

#### Recommandation

Les informations figurant dans les champs <300> à <313> de la norme ST.25 doivent être reprises dans le corps de la demande annexe si elles n’y figurent pas déjà.

### Scénario 18

La norme ST.25 ne prévoit pas de méthode normalisée pour indiquer qu’une région CDS d’une séquence de nucléotides doit être traduite au moyen d’un tableau de codes génétiques différent du tableau classique. Inversement, la norme ST.26 dispose d’un qualificateur “transl\_table” qui peut être employé dans la clé de caractérisation “CDS” pour indiquer que la région doit être traduite au moyen d’un tableau de codes génétiques différent. Si le qualificateur “transl\_table” n’est pas employé, on présume qu’il faut utiliser le tableau de codes génétiques classique.

#### Recommandations

1. Si le listage des séquences selon la norme ST.25 ou la description de la demande indique clairement qu’il faut traduire une région CDS au moyen d’un tableau de codes génétiques différent, il faut employer le qualificateur “transl\_table” en précisant le numéro du tableau dans sa valeur. Ne pas employer ce qualificateur reviendrait à ajouter un élément car on présumerait qu’il faut utiliser le tableau classique (“Standard Code”). En outre, si l’on ne reprend pas les informations sur l’emploi d’un tableau différent dans le listage des séquences conforme à la norme ST.26, ces informations qui figuraient dans le listage des séquences selon la norme ST.25 ou dans la description de la demande vont probablement constituer des éléments supprimés.
2. Si le listage des séquences selon la norme ST.25 ou la description de la demande n’indique pas qu’il faut traduire une région CDS au moyen d’un tableau de codes génétiques différent, le qualificateur “transl\_table” ne doit pas être utilisé; il peut toutefois l’être si la valeur du qualificateur est “1”, ce qui désigne le tableau de codes classique. Toute indication d’une autre valeur dans ce qualificateur constituerait probablement un ajout ou une suppression d’élément.

### Scénario 19

La norme ST.25 ne prévoit pas de méthode normalisée pour indiquer l’emplacement d’une caractéristique, notamment si celle-ci se trouve sur un site ou dans une région qui s’étend au-delà du résidu ou de la série de résidus définis, par exemple une région CDS d’une séquence de nucléotides qui s’étend au-delà de l’une ou des deux extrémités d’une séquence divulguée. En revanche, la norme ST.26 dispose d’un descripteur d’emplacement de la caractéristique qui offre une méthode normalisée pour indiquer l’emplacement d’un tel site ou d’une telle région en s’appuyant sur les symboles “<” ou “>”. Ainsi, l’emplacement de la caractéristique “CDS” doit comporter le codon d’arrêt, même si celui-ci ne figure pas dans la séquence divulguée elle-même; à cette fin, on peut par exemple indiquer l’emplacement de la manière suivante : 1..>321.

#### Recommandations

1. Lorsque le listage des séquences selon la norme ST.25 n’indique pas explicitement que l’emplacement d’une caractéristique s’étend au-delà de la séquence, mais que cet emplacement ressort clairement de la divulgation ou de la séquence elle-même, par exemple si le codon d’arrêt d’une caractéristique CDS ne figure pas dans la séquence, on peut employer les symboles “<” ou “>” dans le listage des séquences selon la norme ST.26 sans ajouter d’élément.
2. Lorsque le listage des séquences selon la norme ST.25 n’indique pas explicitement que l’emplacement d’une caractéristique s’étend au-delà de la séquence, et que cet emplacement ne ressort pas clairement de la divulgation ou de la séquence elle-même, il peut être impossible de se conformer à la norme ST.26 sans ajouter d’élément. Dans ce cas, on peut faire valoir que la demande établissant une priorité et le listage des séquences sont eux-mêmes incomplets. La description de la caractéristique dans le listage des séquences selon la norme ST.26 n’obtiendra alors pas la priorité par rapport à la demande antérieure. Il faut veiller à fournir des informations complètes sur les caractéristiques lorsqu’on établit le listage des séquences et la divulgation de la demande originaux (selon la norme ST.25).

### Scénario 20

L’appendice 1 de la norme ST.25 prévoit que si une séquence de nucléotides contient à la fois des fragments d’ADN et d’ARN, le champ <212> doit prendre la valeur “DNA” et la molécule combinée d’ADN/ARN doit être décrite plus en détail dans les champs <220> à <223>, qui sont consacrés aux caractéristiques. Cependant, la nature exacte de la description détaillée n’étant pas claire, cette obligation n’est généralement pas respectée. Aux termes du paragraphe 55 de la norme ST.26, chaque fragment d’ADN et d’ARN de la molécule combinée d’ADN/ARN doit en outre être décrit par la clé de caractérisation “misc\_feature” et par le qualificateur “note”, ce dernier indiquant s’il s’agit d’un fragment d’ADN ou d’ARN.

#### Recommandations

1. Si le listage des séquences selon la norme ST.25 décrit des fragments d’ADN et d’ARN au moyen d’une ou plusieurs caractéristiques en utilisant les clés de caractérisation <221> (misc\_feature), <222> (Emplacement) et <223> (Autres informations) pour préciser si chaque fragment est de l’ADN ou de l’ARN, puis que ces informations sont converties au format ST.26 en utilisant une clé “misc\_feature” pour chaque fragment d’ADN ou d’ARN, cette méthode ne devrait pas entraîner d’ajout d’élément.
2. Si le listage des séquences selon la norme ST.25 décrit des fragments d’ADN et d’ARN au moyen d’une ou plusieurs caractéristiques en utilisant une clé de caractérisation <221> autre que “misc\_feature”, ainsi que les clés <222> (Emplacement) et <223> (Autres informations) pour préciser si chaque fragment est de l’ADN ou de l’ARN, puis que ces informations sont converties au format ST.26 en utilisant une clé “misc\_feature” pour chaque fragment d’ADN ou d’ARN, cette méthode ne devrait pas entraîner d’ajout ni de suppression d’élément.
3. Si le listage des séquences selon la norme ST.25 indique l’identité (ADN ou ARN) et l’emplacement de chaque segment dans un champ <223> qui n’est pas lié à des champs <221> et <222>, par exemple pour préciser qu’une séquence est artificielle, la conversion de ces informations au format ST.26 utilisant une clé “misc\_feature” pour chaque fragment d’ADN ou d’ARN ne devrait pas entraîner d’ajout d’élément.
4. Si le listage des séquences selon la norme ST.25 décrit la molécule dans une caractéristique au moyen d’une clé <221> (misc\_feature) et d’une clé <223>, en précisant que la molécule combine de l’ADN et de l’ARN, mais qu’il ne fournit pas d’informations sur l’emplacement de chaque fragment, et
   * 1. Si la description indique l’emplacement de chaque fragment d’ADN ou d’ARN, puis ces informations sont converties au format ST.26 en utilisant une clé “misc\_feature” pour chaque fragment d’ADN ou d’ARN, cette méthode ne devrait pas entraîner d’ajout d’élément.
     2. Si la description n’indique pas l’emplacement de chaque fragment d’ADN ou d’ARN, il peut être impossible de se conformer à la norme ST.26 sans ajouter d’élément. Dans ce cas, on peut faire valoir que la demande établissant une priorité et le listage des séquences sont eux-mêmes incomplets. Les descriptions d’emplacement des caractéristiques dans le listage des séquences selon la norme ST.26 n’obtiendront alors pas la priorité par rapport à la demande antérieure. Il faut veiller à fournir des informations complètes sur les caractéristiques lorsqu’on établit le listage des séquences et la divulgation de la demande originaux (selon la norme ST.25).
5. Si le listage des séquences selon la norme ST.25 décrit la molécule dans une caractéristique au moyen d’une clé <221> autre que “misc\_feature” et d’une clé <223>, en précisant que la molécule combine de l’ADN et de l’ARN, mais qu’il ne fournit pas d’informations sur l’emplacement de chaque fragment, et
6. Si la description indique l’emplacement de chaque fragment d’ADN ou d’ARN, puis ces informations sont converties au format ST.26 en utilisant une clé “misc\_feature” pour chaque fragment d’ADN ou d’ARN et un qualificateur “note” supplémentaire prenant la valeur de la clé de caractérisation <221> originale, cette méthode ne devrait pas entraîner d’ajout d’élément.
7. Si la description n’indique pas l’emplacement de chaque fragment d’ADN ou d’ARN, il peut être impossible de se conformer à la norme ST.26 sans ajouter d’élément. Dans ce cas, on peut faire valoir que la demande établissant une priorité et le listage des séquences sont eux-mêmes incomplets. Les descriptions d’emplacement des caractéristiques dans le listage des séquences selon la norme ST.26 n’obtiendront alors pas la priorité par rapport à la demande antérieure. Il faut veiller à fournir des informations complètes sur les caractéristiques lorsqu’on établit le listage des séquences et la divulgation de la demande originaux (selon la norme ST.25).
8. Si le listage des séquences selon la norme ST.25 indique dans un champ <223> (par exemple pour signaler une séquence artificielle) que la molécule est une combinaison d’ADN et d’ARN, mais qu’il ne fournit pas de clé de caractérisation ou d’informations sur l’emplacement de chaque fragment, et
9. Si la description indique l’emplacement de chaque fragment d’ADN ou d’ARN, puis ces informations sont converties au format ST.26 en utilisant une clé “misc\_feature” pour chaque fragment d’ADN ou d’ARN, cette méthode ne devrait pas entraîner d’ajout d’élément.
10. Si la description n’indique pas l’emplacement de chaque fragment d’ADN ou d’ARN, il peut être impossible de se conformer à la norme ST.26 sans ajouter d’élément. Dans ce cas, on peut faire valoir que la demande établissant une priorité et le listage des séquences sont eux-mêmes incomplets. Les descriptions d’emplacement des caractéristiques dans le listage des séquences selon la norme ST.26 n’obtiendront alors pas la priorité par rapport à la demande antérieure. Il faut veiller à fournir des informations complètes sur les caractéristiques lorsqu’on établit le listage des séquences et la divulgation de la demande originaux (selon la norme ST.25).

[Fin de l’annexe VII et du document]

1. La norme ST.26 peut prévoir de remplacer une clé de caractérisation de la norme ST.25, par exemple TATA\_signal, par une clé de caractérisation, un qualificateur ou une valeur plus généraux, par exemple regulatory/regulatory\_class/ TATA\_box. Dans ce cas, la clé plus restreinte de la norme ST.25 aura priorité par rapport à la demande antérieure. En revanche, la clé de caractérisation ou le qualificateur plus généraux de la norme ST.26 (par exemple regulatory/regulatory\_class) n’aura pas priorité dans sa portée complète par rapport à la demande antérieure. [↑](#footnote-ref-2)
2. Les numéros de référence indiqués dans le tableau ci-après renvoient aux numéros de clé de caractérisation et de qualificateur figurant dans le vocabulaire contrôlé de l’annexe I de la norme ST.26. [↑](#footnote-ref-3)