

Comité des normes de l'OMPI (CWS)

Dixième session
Genève, 21 – 25 novembre 2022

PROPOSITION DE REVISION DE LA NORME ST.26 DE L'OMPI

Document établi par le Bureau international

RESUME

1. L'Équipe d'experts chargée du listage des séquences propose des révisions à la version actuelle de la norme ST.26 de l'OMPI, afin de corriger les erreurs rédactionnelles relevées par l'Équipe d'experts et d'inclure trois nouveaux exemples dans l'annexe VI. À la neuvième session du Comité des normes de l'OMPI (CWS), le Bureau international a pris note des préoccupations des offices concernant les nouvelles révisions de la norme ST.26 de l'OMPI et a fourni des conseils sur la manière de gérer ces modifications.

CONTEXTE

2. À la reprise de sa quatrième session, qui s'est tenue en mars 2016, le CWS a adopté la norme ST.26 de l'OMPI (voir le paragraphe 52 du document CWS/4BIS/16), intitulée "Recommandation de norme relative à la présentation des listages de séquences de nucléotides et d'acides aminés en langage XML (eXtensible Markup Language)". L'Équipe d'experts est responsable des modifications de la norme ST.26 de l'OMPI, dans le cadre de la tâche n° 44 dont la description est la suivante :

"Fournir un appui au Bureau international en communiquant les besoins et le retour d'information des utilisateurs sur l'outil d'édition et de validation de la norme ST.26; fournir un appui au Bureau international pour les révisions à apporter en conséquence aux Instructions administratives du PCT; et préparer les révisions à apporter à la norme ST.26 de l'OMPI."

3. Le CWS a approuvé les révisions de la norme ST.26 de l'OMPI à ses cinquième, sixième, septième et huitième sessions, donnant lieu aux versions 1.1, 1.2, 1.3 et 1.4 respectivement (voir le paragraphe 41 du document CWS/5/22, le paragraphe 112 du document CWS/6/34, le

paragraphe 133 du document CWS/7/29 et le paragraphe 49 du document CWS/8/24). Plus récemment, à sa neuvième session, tenue en novembre 2021, plusieurs modifications de forme et de fond ont été approuvées par le CWS et publiées en tant que dernière version de la norme ST.26 de l'OMPI, version 1.5 (voir le paragraphe 63 du document CWS/9/25). Il s'agissait notamment de modifications visant à respecter les exigences de l'International Nucleotide Sequence Database Collection (INSDC).

4. Au cours de la neuvième session, des délégations ont exprimé des préoccupations concernant les incidences potentielles des révisions de la norme ST.26 de l'OMPI après que tous les États membres ont mis en œuvre cette norme, par exemple au cas où leurs systèmes informatiques devraient être réorganisés (voir le paragraphe 62 du document CWS/9/25). Le Bureau international a indiqué qu'il fournirait des orientations à ce sujet à la dixième session.

5. Plusieurs modifications rédactionnelles ont été apportées à l'annexe VI et à l'appendice de l'annexe VII après qu'il a été constaté qu'il y avait plusieurs erreurs de référencement pour les numéros d'identification de séquence et des emplacements de caractéristiques incorrects fournis après la révision effectuée à la neuvième session. Ces modifications ont été effectuées et publiées en mai 2022.

PROPOSITION DE REVISION DE LA NORME ST.26 DE L'OMPI

6. L'Équipe d'experts chargée du listage des séquences a préparé une proposition de révision de la norme ST.26 de l'OMPI pour examen et, le cas échéant, approbation par le CWS. Lorsque cela est indiqué, le texte barré indique une modification et le texte en gras indique un ajout.

7. Les modifications qu'il est proposé d'apporter au corps du texte de la norme ST.26 de l'OMPI peuvent être résumées comme suit :

a) mettre à jour le paragraphe 44 afin d'améliorer la valeur du nom de fichier fournie de manière à ce qu'elle comprenne "-" en lieu et place de "_" :
filename="US11-405455-SEQL.xml", étant donné que certains offices de la propriété intellectuelle n'autorisent pas le trait de soulignement "_" dans le nom de fichier des fichiers électroniques soumis;

b) mettre à jour le paragraphe 45 pour améliorer la définition de ApplicationNumberText comme suit : *"Le numéro de la demande tel que fourni par l'office de dépôt (par exemple, PCT/IB2013/099999)."*;

c) mettre à jour le paragraphe 73 pour garantir la cohérence avec la description de la section 6 en incluant la phrase suivante dans la deuxième moitié de la première phrase : *"La section 8 contient la liste exclusive des qualificateurs **et la définition de leurs formats de valeur, le cas échéant**, pour chaque clé de caractérisation de chaque séquence d'acide aminé."*;

d) mettre à jour le paragraphe 76 pour apporter une nouvelle définition de l'attribut "id", qui est prévu pour les qualificateurs de texte libre dépendant de la langue et se trouve dans la nouvelle quatrième ligne du tableau inclus. La description se lit comme suit : *"Un qualificateur de texte libre dépendant de la langue peut être identifié de manière unique en utilisant l'attribut facultatif XML "id" dans l'élément INSDQualifier. La valeur de l'attribut "id" doit commencer par la lettre "q" et être suivi par tout nombre entier positif. Les valeurs "id" ne doivent pas être numérotées consécutivement, mais un attribut "id" doit être propre à chaque élément INSDQualifier, c'est-à-dire que la valeur de l'attribut ne doit être utilisée qu'une seule fois dans un fichier de listage des séquences."*;

- e) mettre à jour le paragraphe 77 pour inclure un espace et le symbole ” entre “organisme” et “pour” à la première ligne : “ ... et “organisme” pour les séquences d’acides aminés...”;
- f) mettre à jour du paragraphe 79, exemple 2 : supprimer l’espace avant “mol_type”;
- g) mettre à jour le paragraphe 86 comme suit : “L’utilisation du texte libre doit être limitée à quelques termes courts indispensables à la compréhension d’une caractéristique de la séquence. Pour chaque qualificateur **autre que le qualificateur “traduction”**, le texte libre ne doit pas dépasser 1000 caractères”;
- h) mettre à jour le début du paragraphe 87.d) comme suit : “Pour les qualificateurs dont la valeur en texte libre dépend de la langue, ...”; et
- i) mettre à jour toutes les occurrences de “propriété industrielle” en “propriété intellectuelle”, après l’approbation de cette mise à jour à la neuvième session du CWS visant à élargir le champ d’application du Manuel de l’OMPI.
8. Les modifications proposées aux annexes de la norme ST.26 de l’OMPI peuvent être résumées comme suit :
- a) annexe I, section 6 : mettre à jour le tableau 5 pour inclure la fréquence comme qualificateur avec une valeur de texte libre dépendant de la langue;
- b) annexe I, section 6.21 : mettre à jour le format de la valeur obligatoire pour indiquer que le qualificateur de fréquence dépend de la langue;
- c) annexe I, section 6 : mettre à jour le titre du tableau 5 en “Liste des valeurs des qualificateurs pour les séquences de nucléotides ayant des valeurs de texte libre dépendant de la langue”;
- d) annexe I, sections 6 et 8 : mettre à jour la note, qui ne fait actuellement référence qu’aux “procédures nationales ou régionales”, pour inclure “international”;
- e) annexe I, section 6.77 : supprimer les espaces dans l’exemple après la valeur de l’anticodon, mais avant le symbole inférieur à;
- f) annexe I, section 8 : mettre à jour la note pour qu’elle soit conforme à la note qui figure à la section 6. La note mise à jour commencera par : “Toute valeur de qualificateur fournie pour un qualificateur dont le format de valeur “texte libre” dépend de la langue peut nécessiter une traduction...”;
- g) annexe II : inclure le commentaire suivant dans la DTD afin de fournir des précisions sur la valeur qui doit être fournie dans l’attribut `fileName` : “Par défaut, il s’agit de la valeur fournie pour le nom d’un projet dans WIPO Séquence. Si la valeur est identique au nom de fichier XML ST.26 réel, il convient de noter que les offices peuvent appliquer leurs exigences concernant le nom de fichier utilisé, ce qui peut restreindre les caractères autorisés pour les fichiers électroniques soumis. Veuillez vous référer à la base de connaissances de WIPO Séquence et de la norme ST.26 pour plus de détails sur les conventions de dénomination des offices pour les fichiers électroniques”;
- h) annexe VI, exemple 3.c)-1, question 3 : corriger le terme “acylation” par “acétylation”;

- i) annexe VI, exemple 3.k)-2, question 3 : corriger la référence à SEQ ID NO. 12 dans le troisième paragraphe comme suit : *“la clé de caractérisation “misc_difference” ayant l’emplacement de la caractéristique “10” devrait être utilisée avec deux qualificatifs “replace” où la valeur de l’un serait “gk” et la seconde serait “c”*;
- j) annexe VI, exemple 7.b)-4 et 7b)-5 : remplacer les numéros d’identification de séquence SEQ ID NO. 89, SEQ ID NO.90 et NO. 91 en SEQ ID NO. 90, SEQ ID NO. 91 et SEQ ID NO. 92;
- k) annexe VI et appendice à l’annexe VI : incluent une série de trois nouveaux exemples, l’exemple 94-2, l’exemple 30-2 et l’exemple 12-1. Les nouveaux exemples sont fournis dans leur intégralité à l’annexe I du présent document;
- l) annexe VII, scénario 4 : remplacer la première phrase par le texte suivant : *“Tant pour les séquences de nucléotides que pour les séquences d’acides aminés, la norme ST.26 a la clé de caractéristique obligatoire “source avec deux qualificatifs obligatoires.”*;
- m) annexe VII, scénario 4 : remplacer la légende du tableau des séquences de nucléotides par : *“ST.26 – clé de caractérisation 5.37 source; qualificateur obligatoire 6.39 mol_type (voir ST.26 paragraphe 75)”*;
- n) annexe VII, scénario 4 : remplacer la légende du tableau des séquences d’acides aminés par *“ST.26 – feature key 7.30 source; qualificatif obligatoire 8.1 mol_type (voir ST.26, paragraphe 75)”*;
- o) annexe VII, scénario 8 : mettre à jour la ligne 18 du tableau inclus pour ajouter dans la quatrième colonne le texte suivant : **“TATA_signal et (si <223> présent) .”**, et supprimer également de la note le texte suivant : *“(si <223> présent)”*; et
- p) appendice à l’annexe VI : toute modification correspondante requise en raison des mises à jour de l’annexe VI. L’appendice à l’annexe VI est fourni en tant qu’annexe II du présent document.

PRINCIPES DIRECTEURS POUR LA REVISION DE LA NORME ST.26

9. Comme indiqué ci-dessus, le Bureau international a pris note, à la neuvième session du CWS, des préoccupations concernant les mises à jour régulières de la norme ST.26 de l’OMPI et de leurs incidences potentielles (voir le paragraphe 4 du présent document).

10. Il convient de noter que la norme ST.26 de l’OMPI ne sera mise à jour qu’une fois par an au maximum, selon les besoins et avec l’approbation du CWS. Dans la proposition de révision présentée au CWS/10, il y a à la fois des modifications de forme et de fond. Les modifications de forme apportées au corps du texte ne devraient pas avoir de conséquences majeures sur les utilisateurs, notamment les offices de propriété intellectuelle. Les modifications de fond apportées aux annexes VI et VII ne devraient pas non plus avoir de conséquences sur les offices de propriété intellectuelle et les utilisateurs.

11. Le Bureau international fournira, dans les futures versions de la norme ST.26 de l’OMPI, un numéro de version approprié, X.Y, qui indiquera aux utilisateurs de la norme si cette révision est considérée comme rétrocompatible avec la version précédente. Les mises à jour mineures de la norme ST.26 de l’OMPI sont indiquées par un incrément au deuxième chiffre seulement (“Y”) et celles-ci sont considérées comme rétrocompatibles. Les mises à jour majeures, qui sont indiquées par un incrément au premier chiffre (“X”), auront probablement un impact sur vos systèmes informatiques et peuvent nécessiter une réorganisation. Le numéro de version de la DTD peut être différent de celui du corps du texte et des autres annexes.

12. Étant donné qu'aucune modification n'est proposée à la DTD de la norme ST.26 de l'OMPI, à l'exception de l'ajout d'un commentaire dans cette nouvelle version, et que les modifications proposées au corps du texte et aux annexes portent principalement sur la forme ou sont envisagées à des fins de clarification, le numéro de version de la DTD dans ce cas restera la version 1.3 et le numéro de version de la norme sera mis à jour à la version 1.6.

13. Si la révision proposée obtient l'approbation du CWS et l'appui de l'Équipe d'experts chargée du listage des séquences, le Bureau international intégrera les modifications adoptées à la norme ST.26 de l'OMPI et la publiera en 2022.

14. *Le CWS est invité à*

a) *prendre note du contenu du présent document et de ses annexes et*

b) *examiner et approuver les propositions de révision de la norme ST.26 telles que visées aux paragraphes 7 et 8 ci-dessus et reproduites dans les annexes I et II du présent document.*

[L'annexe I (les nouveaux exemples proposés pour l'annexe VI) suit]

PROPOSITION DE NOUVEAUX EXEMPLES POUR L'ANNEXE VI DE LA NORME ST.26

Paragraphe 94 – Toute séquence variante divulguée comme une séquence unique avec des résidus alternatifs énumérés

Exemple 94-2 – Représentation d'une séquence unique avec des acides aminés alternatifs énumérés qui peuvent être des acides aminés modifiés

Une demande de brevet décrit le polypeptide suivant :

Leu-Glu-Tyr-Cys-Leu-Lys-Arg-Trp-Xaa-Glu-Thr-Ile-Ser-His-Cys-Ala-Trp

où Xaa peut être Ile, Ala, Phe, Tyr, alle, Melle ou Nle.

Question 1 : La norme ST.26 prescrit-elle l'intégration de la ou des séquences?

OUI

Le peptide énuméré fournit 16 acides aminés définis avec précision. Par conséquent, la séquence doit être intégrée dans un listage des séquences comme l'exige le paragraphe 7.b) de la norme ST.26.

Question 3 : Comment la ou les séquences devraient-elles être représentées dans le listage des séquences?

Le symbole ambigu le plus restrictif pouvant englober "Ile, Ala, Phe, Tyr, alle, Melle, ou Nle" est "X", la séquence doit être intégrée dans un listage de séquences comme suit :

LEYCLKRWXETISHCAW (SEQ ID NO : xx)

Le paragraphe 30 de la norme ST.26 requiert qu'"[u]n acide aminé modifié doit être accompagné d'une description supplémentaire dans le tableau des caractéristiques". Cependant, il n'exige pas l'utilisation d'une clé de caractérisation spécifique pour décrire les acides aminés modifiés. Même si le paragraphe 30 décrit l'utilisation des clés de caractérisation "CARBOHYD", "LIPID", "MOD_RES", et "SITE", ces clés sont plus appropriées pour les scénarios où l'acide aminé modifié ne figure pas dans la liste d'alternatives d'un emplacement donné. Dans cet exemple, la clé de caractérisation "VARIANT" satisfait à l'exigence du paragraphe 30, puisqu'elle permet de représenter toutes les alternatives à l'emplacement de la variante. Ainsi, la clé "VARIANT" accompagnée du qualificateur "note" et d'une valeur de qualificateur "Ile, Ala, Phe, Tyr, alle, Melle, ou Nle" doit être utilisée pour décrire l'emplacement de la variante en position 9. L'utilisation d'une deuxième clé de caractérisation telle que "SITE" et du qualificateur "note" peut être employée pour mieux définir les acides aminés modifiés trouvés en position 9.

Paragraphe pertinents de la norme ST.26 : 3.a), 7.b), 27, 30, **94**, 96, et annexe I, section 4 et tableau 4

Paragraphe 30 – Annotation d'un acide aminé modifié

Exemple 30-2 – Acides aminés modifiés après traduction

Une demande de brevet décrit le polypeptide suivant :

Leu-Glu-Tyr-Cys-Leu-Lys-Arg-Trp-Glu-Thr-Ile-Ser-His

dans lequel l'Arg en position 7 peut être modifié en citrulline après traduction.

Question 1 : La norme ST.26 exige-t-elle l'intégration de la ou des séquences?

OUI

Le peptide énuméré fournit 13 acides aminés spécialement définis. Par conséquent, la séquence doit être intégrée dans un listage des séquences comme l'exige le paragraphe 7.b) de la norme ST.26.

Question 3 : Comment la ou les séquences devraient-elles être représentées dans le listage des séquences?

Selon le paragraphe 29 de la norme ST.26, un acide aminé modifié doit être représenté dans la séquence comme l'acide aminé non modifié correspondant chaque fois que cela est possible.

Par conséquent, la séquence doit être intégrée dans un listage de séquences comme suit :

LEYCLKRWETISH (numéro d'identification : xx)

où le symbole "R" est utilisé pour représenter l'arginine en position 7.

Une description supplémentaire indiquant que l'arginine en position 7 peut être modifiée en citrulline est requise. La modification de l'arginine en citrulline est une modification après traduction. Par conséquent, la clé de caractérisation "MOD_RES" doit être utilisée avec le qualificateur obligatoire "note" pour indiquer que l'arginine peut être modifiée pour devenir de la citrulline. Dans l'emplacement d'une caractéristique, le descripteur d'emplacement est le numéro de position du résidu de l'arginine modifiée.

Paragraphe(s) pertinent(s) de la norme ST.26 : 3.a), 7.b), 30, et annexe I, section 7, clé de caractérisation 7.18

Paragraphe 12 – Séquence nucléotidique circulaire

Exemple 12-1 : séquence de nucléotides circulaire

Une demande de brevet contient la figure suivante, divulguant la séquence d'ADN du plasmide pCIRC1 :



Question 1 : La norme ST.26 exige-t-elle l'intégration de la ou des séquences?

OUI

La séquence nucléotidique énumérée comporte plus de 10 nucléotides spécialement définis. Par conséquent, la séquence doit être intégrée dans un listage de séquences comme l'exige le paragraphe 7.a) de la norme ST.26.

Question 3 : Comment la ou les séquences devraient-elles être représentées dans le listage des séquences?

Selon le paragraphe 12 de la norme ST.26, lorsque les séquences de nucléotides sont de configuration circulaire, le demandeur doit choisir le nucléotide en position de résidu numéro 1. Pour les besoins de cet exemple, le résidu "a" identifié par la flèche dans la figure sera utilisé comme position 1. Cependant, n'importe quel résidu peut être choisi comme position 1. Avec le résidu indiqué par la flèche comme position 1, la séquence doit être intégrée dans un listage de séquences comme suit :

atggataatgaagaagttaacgaagaatgtatgagatttttcaagaacgctcgtgcgcatctggataaacatctaacaatcaaggttgacatgcatgaaaatgaaaatgcatatcacgttcagatgcttccctggatggaatacactcgaaatctactaggtttctcgaagagctacttttgaacaagaaaatgtaccaccaatg (SEQ ID NO : xx)

La séquence doit être décrite plus en détail à l'aide de la clé de caractérisation "misc_feature" avec un emplacement de "212^1", qui indique que le dernier résidu de la séquence, la position 212, est lié au résidu 1. Le qualificateur "note" doit être inclus avec une valeur indiquant que la molécule est circulaire.

Paragraphe pertinent de la norme ST.26 : 7.a), 12, et annexe I, section 5, clé de caractérisation 5.15

[L'annexe II (appendice de l'annexe VI) suit]