|  |  |  |
| --- | --- | --- |
|  | WIPO-E | **E** |
| CWS/9/12 REV. |
| ORIGINAL: english   |
|  DATE: October 27, 2021   |

# Committee on WIPO Standards (CWS)

**Ninth Session**

**Geneva, November 1 to 5, 2021**

## Proposal for the revision of WIPO Standard ST.26

*Document prepared by the International Bureau*

## Introduction

 At its reconvened fourth session, held in March 2016, the Committee on WIPO Standards (CWS) adopted WIPO Standard ST.26 (see paragraph 52 of document CWS/4BIS/16), entitled “Recommended standard for the presentation of nucleotide and amino acid sequence listings using XML (eXtensible Markup Language)”.

 Further revisions to WIPO Standard ST.26, versions 1.1 and 1.2, were approved by the CWS at its fifth and sixth sessions, respectively (see paragraph 41 of document CWS/5/22, and paragraph 112 of document CWS/6/34). At its seventh session, held in July 2019, the CWS approved a further revision to the Standard, version 1.3 (see paragraph 133 of document CWS/7/29). Most recently at its eighth session, held in December 2020, several editorial and substantial revisions were approved by the CWS and published as the most recent version of WIPO ST.26, version 1.4 (see paragraph 49 of document CWS/8/24).

 The Sequence Listings Task Force has been working under the framework of Task No.44 whose description reads:

 *"Support the International Bureau by providing users’ requirements and feedback on the ST.26 authoring and validation software tool; support the International Bureau in the consequential revision of the PCT Administrative Instructions; and prepare necessary revisions of WIPO Standard ST.26."*

## Proposed Revision of WIPO Standard ST.26

 The Sequence Listings Task Force has prepared a proposal for revision to WIPO ST.26 for consideration and where appropriate approval by the CWS. The proposal is identified in the Annexes I to III of the present document, with yellow highlighting indicating an addition of new text and purple highlighting indicating a removal.

 The proposed changes to the Main Body of WIPO Standard ST.26 can be summarized as follows:

1. all instances of the “SOURCE” feature for amino acid sequences were changed to “source”, in order to comply with recent UniProt changes;
2. all instances of the qualifiers “ORGANISM”, “MOL\_TYPE”, and “NOTE” were changed to “organism”, “mol\_type”, and “note”, in order to comply with recent UniProt changes;
3. all instances of 'patent Office' were changed to 'IPO';
4. changed the position of the acronym 'DTD' from the end of the sentence next to 'Document Type Definition';
5. an edit to paragraph 3(g)(i)(2) to replace "nucleotide analogues" with "backbone moieties";
6. an edit to the code segment in paragraph 44 to correct the quotation marks used for softwareVersion, which should be: softwareVersion="1.0";
7. an edit to paragraph 46, Examples 1-2 to change the nonEnglishFreeTextLanguageCode from "jp" to "ja" and also correcting the indentation of <FilingDate> to be aligned with <ApplicationNumberText>; and
8. an edit to paragraph 97 to replace the phrase “…a subgroup of…” with “…a value other than…”.

 The proposed changes to the Annexes of WIPO Standard ST.26 are highlighted for consideration by the CWS. For a full list of all of the proposed changes, refer to Annexes I-III of the present document:

1. all instances of the “SOURCE” feature for amino acid sequences were changed to “source”, including Annex I, Section 7, in order to comply with recent UniProt changes;
2. all instances of the qualifiers “ORGANISM”, “MOL\_TYPE”, and “NOTE” were changed to “organism”, “mol\_type”, and “note”, including Annex I, Section 8, in order to comply with recent UniProt changes;
3. the column header of the second columns of Tables 1 and 3 and Tables A and B to all read 'Definition' for consistency. Also the addition of a 3-letter code column to Table B;
4. the addition of new qualifiers to Annex I, to align with the latest INSDC updates;
5. the addition of new Example 29-2 in Annex VI to provide clarification on the inclusion of the corresponding unmodified amino acid;
6. the edit of the Introduction to Annex VI to provide justification for the responses provided for “when should a modified residue be represented by the corresponding unmodified residue or the variable ‘n’ or ‘X’?” and “what modifications result in a ‘modified residue’ and MUST be annotated?” and additional language added to address how to represent disclosures where sequence length can vary under "Most encompassing sequence";
7. a revision of Annex VI in general to ensure consistency with the use of “must”, “should”, "preferably" and “may”, to ensure there is no confusion, as "preferably" is not defined within paragraph 4, WIPO ST.26;
8. an edit to Annex VI, Example 3(c)-2 to correct "should" to "must" so that the following phrase now reads:

'Therefore, if “X” is intended to represent “any amino acid”, then it ~~should~~ must be annotated with the feature key VARIANT and a NOTE qualifier with the value, “X can be any amino acid”.';

1. the inclusion of new SEQ ID NO. 51 which forms part of Example 29-2 to Appendix to Annex VI and the renumbering of subsequent sequences to ensure the numbering within this sequence listing example remains consecutive;
2. an edit to Annex VI, Example 30-1 to correct the last paragraph in Question 3 as follows:

'In addition, there is a disulfide bond between the two Cys residues. Therefore the feature key “DISULFID” is used to describe an intrachain crosslink. The feature location element is the residue position numbers of the linked Cys residues in “x..y” format, i.e., “4..15”. The mandatory qualifier note~~NOTE~~ should describe the intrachain disulfide bond. ~~The qualifier NOTE is not mandatory.~~';

 and

1. an edit to the Appendix to Annex VI, existing SEQ ID NO.51 (new SEQ ID NO:52) to add a note qualifier to the DISULFID feature key, as a result of the change proposed under paragraph 6 (j) above:

<INSDFeature>

 <INSDFeature\_key>DISULFID</INSDFeature\_key>

 <INSDFeature\_location>4..15</INSDFeature\_location>

 <INSDFeature\_quals>

 <INSDQualifier>

 <INSDQualifier\_name>note</INSDQualifier\_name>

 <INSDQualifier\_value>intrachain disulfide bond</INSDQualifier\_value>

 </INSDQualifier>

 </INSDFeature\_quals>

</INSDFeature>

 The proposed revision of WIPO Standard ST.26 is provided as Annexes I-III to this document for consideration by the CWS as follows:

* Annex I contains the proposed amendment to WIPO Standard ST.26, where the revisions are highlighted, except its Annex IV where no amendment is proposed (as well as the two components below);
* Annex II contains WIPO ST.26, Annex III, which is published as a separate file; and
* Annex III contains WIPO ST.26, Annex VI, Appendix, which is published as a separate file.

 *The CWS is invited to:*

1. *note the content of this document and Annexes to this document; and*
2. *consider and approve the proposed revisions to WIPO Standard ST. 26 as referred to in paragraphs 5 and 6 above and reproduced in the Annexes to this document.*

[Annexes (the proposed revision of WIPO Standard ST.26) follow]