Standard ST.26: Common Mistakes and Best Practices

Mary Till, Senior Legal Advisor, Office of Patent Legal Administration (OPLA)
Kathleen Kalafus, Technical Information Specialist, Sequence Systems Service Center (S3C)
Christopher Babic, Supervisory Patent Examiner, Technology Center 1600
Anish Gupta, PCT Legal Examiner, Office of International Patent Legal Administration (IPLA)

November 15, 2022
To send in questions or comments during the webinar, please email:

PatentQuality@uspto.gov
Agenda

• Compliance with US rules
• Electronic filing
• WIPO Sequence authoring tool
• Compliance with ST.26 requirements
• ST.26 sequence listings in PCT and national phase applications
Compliance with US rules
Compliance with US Rules

Filing date, filing date, filing date. . .

• Application filing date controls what format of sequence listing is required. Applications filed before July 1, 2022, must comply with ST.25 (37 CFR 1.821-1.825). Applications filed on or after July 1, 2022, must comply with ST.26 (37 CFR 1.831-1.835).

• See Federal Register notice 87 Fed. Reg. 30806, 30806: “Applicability date: Patent applications filed on or after July 1, 2022, having disclosures of nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b) must comply with new rules for submission of a “Sequence Listing XML” in accordance with 37 CFR 1.831 through 1.835.”

• See Federal Register notice 87 Fed. Reg. 30806, 30807: “Applications pending prior to July 1, 2022, will not have to comply with WIPO Standard ST.26; rather, such applications will require the submission of a “Sequence Listing,” as defined in 37 CFR 1.821(a), in compliance with 37 CFR 1.821-[sic]1.825.”
Compliance with US Rules

What we have seen:

• Applicant receives pre-examination notice regarding sequence compliance on July 11, 2022 for an application filed June 27, 2022.

• Applicant erroneously submits a sequence listing in xml file format.

• Applicant receives a notice of incomplete response to the original notice, time period continues to run from the original notice.
Compliance with US Rules

Sequence compliance in applications with benefit and/or priority claims to prior filed applications:

• As previously stated, it is all about the actual filing date of the application.

• Whether the present application has a benefit or priority claim under 35 USC 119, 120, 121 or 365 to a prior filed application, if the actual filing date is on or after July 1, 2022, the benefit or priority claim is of no consequence. A sequence listing in XML file format in compliance with ST.26 (37 CFR 1.831-1.835) is REQUIRED for applications filed on or after July 1, 2022 having nucleotide and/or amino acid sequence disclosures that meet the definition of 1.831(b).
Electronic filing
Electronic filing

• **Issue applicants observe:** the USPTO Patent Electronic System (PES) will not permit a sequence listing.TXT file for applications filed on/after July 1, 2022.

• This does **DO NOT** mean you should circumvent the PES and upload a sequence listing in ASCII file format as something else, like a table or chemical drawing, when an xml is **REQUIRED**.
Electronic filing

• **Issue applicants observe:** EFS-Web won’t let me file an XML.

• This does **DO NOT** mean that you should change the file extension from .xml to .TXT to file in EFS-Web.
Electronic filing

• **Issue applicants present:** Using characters such as { or / in a “Sequence Listing XML” file name.

• 37 CFR 1.834(a)(3) specifically states:

  “(3) Be named as *.xml, where “*” is one character or a combination of characters limited to upper- or lowercase letters, numbers, hyphens, and underscores, and the name does not exceed 60 characters in total, excluding the extension. No spaces or other types of characters are permitted in the file name.”
Electronic filing

- **Issue applicants observe**: I don’t like using Patent Center so I will file the majority of my application in EFS-Web and file my “Sequence Listing XML” as a follow-on paper in Patent Center.

- Patent Center is not **JUST** for docx submissions. Make sure that if you are doing a hybrid filing (part in EFS-Web and part in Patent Center) that you allow sufficient time to generate an acknowledgement receipt to have information to file the XML in Patent Center as a follow-on document.
Questions?
WIPO sequence authoring tool
WIPO sequence authoring tool

• Always ensure you are using the most up-to-date version of WIPO Sequence.

• Out of date versions of WIPO Sequence may generate a non-compliant sequence listing and/or have bugs.

• The version number is displayed in the upper left hand corner of the interface.
WIPO sequence authoring tool

- You can get the latest version of WIPO Sequence by allowing the software to auto-update if you have an internet connection.
WIPO sequence authoring tool

- Generate the sequence listing XML file using the “Generate Sequence Listing” button.
WIPO sequence authoring tool

- Do not file a “human readable” print out of your project.
WIPO sequence authoring tool

- This is a “human readable” print out of project data; it is NOT a sequence listing!
- This text file should not be filed at a patent office; it is for applicant use only.
WIPO sequence authoring tool

- **DO NOT** manually edit your XML file after generation.
- Make any changes within the project, and then use the “Generate Sequence Listing” button to create a new XML file.
- If you don’t have the original project used to create the XML file, you can use the “Import Sequence Listing” function to create a new project by importing an XML.
WIPO sequence authoring tool

- Always validate your XML file after generation before filing.
WIPO sequence authoring tool

- **Helpful hint** – import of ST.25 sequence listings with a large number of nucleotide sequences.
- Use the “bulk edit” function to add “mol_type” qualifier values to multiple sequences at once.
WIPO sequence authoring tool

1. Under “Type of bulk edit*” select “Qualifier molecule type”;
2. Select the “Molecule Type*”;
3. Enter the SEQ ID numbers of the sequence you want to bulk edit;
4. Select the desired “mol_type” qualifier value.

Type of bulk edit*

Qualifier molecule type

Select the type of bulk edit to continue

Molecule Type*

DNA

Qualifier Molecule Type*

genomic DNA

Select Range of Sequence IDs*

1,3,5, 8

Use commas to separate individual sequences (for example: “2,5,8”); use a dash to indicate a range of sequences (for example: “2-8”).

Only DNA/RNA sequences allow editing of the qualifier ‘mol_type’ because for amino acid sequences this qualifier value is automatically set to ‘protein’.

An error will be generated during project validation for sequences where the organism is specified as ‘synthetic construct’ and the qualifier value for ‘mol_type’ is not ‘other DNA’ or ‘other RNA’.

The qualifier ‘mol_type’ of hybrid DNA/RNA sequences, where the ‘combinedIndicator’ = ‘Yes’, cannot be bulk edited.

Enter one or more of the following DNA sequences to edit the qualifier ‘mol_type’.
Questions?
Compliance with ST.26 requirements
Compliance with ST.26 requirements

• The “organism” qualifier value must be one of the following options:

1. A full Latin “Genus species” name or “Genus sp.” (see WIPO ST.26 paragraphs 78-80), wherein the Genus name is capitalized and the species name is lower-case;

2. A virus name (see WIPO ST.26 paragraph 82);

3. “unidentified”, for naturally occurring sequences where the Latin Genus and/or species is unknown. Any other taxonomic information should be contained in a “note” qualifier (see WIPO ST.26 paragraph 81); or

4. “synthetic construct” for sequences that are not naturally occurring (see WIPO ST.26 paragraph 83).
 Compliance with ST.26 requirements

• If you use “Genus species” or “Genus sp.” for the “organism” qualifier value.

• **DO NOT** abbreviate the Genus name.
  
  o Staphylococcus aureus ✔
  
  o S. aureus ❌
Compliance with ST.26 requirements

• If you use “Genus species” or “Genus sp.” for the “organism” qualifier value.

• **DO NOT** leave the species name out, or spell out “species.”
  
  o Staphylococcus sp. ✓
  
  o Staphylococcus ✗
  
  o Staphylococcus species ✗
Compliance with ST.26 requirements

• If you use “Genus species” or “Genus sp.” for the “organism” qualifier value.

• **DO NOT** include the Genus name in lowercase or the species name in uppercase.
  
  - Homo sapiens ✔
  - homo sapiens ✗
  - Homo Sapiens ✗
Compliance with ST.26 requirements

• If you use “Genus species” or “Genus sp.” for the “organism” qualifier value.

• **DO NOT** use a taxonomic name that is broader than a Genus.
  
  o Staphylococcaceae sp. ✗

  o Staphylococcaceae is a Family of bacteria, **NOT** a Genus!

• If the only taxonomic information about the source organism is broader than a genus, use “unidentified” as the “organism” qualifier value and include any known taxonomic information in a “note” qualifier.
Compliance with ST.26 requirements

• If you use “Genus species” or “Genus sp.” for the “organism” qualifier value.

• **DO NOT** include strain, substrain, isolate, or subspecies information.
  
  - Staphylococcus aureus ✓
  
  - Staphylococcus aureus strain MRSA252 ❌

• Strain, substrain, isolate, subspecies, cultivar, clone, serovar, cell line, and other source organism information can be included in an appropriate “source” feature qualifier (see WIPO Sequence Annex I, Section 5.37).
Compliance with ST.26 requirements

• Correctly describing modified residues and variant sequences:
  
  o Modified nucleotides must be described using the feature key "modified_base" (ST.26, para. 17);

  o Modified amino acids must be described using the feature key "MOD_RES", "SITE", "CARBOHYD" or "LIPID", as appropriate (ST.26, pgh. 30); and

  o Variants must be described using the feature keys "variation" or "misc_difference" for nucleotide sequences, and "VAR_SEQ" or "VARIANT" for amino acid sequences (ST.26, para. 96).
Compliance with ST.26 requirements

- **Common mistake:** use of “modified_base” or “misc_feature” to describe a nucleotide sequence variant.

  Example disclosure:

  \[5'-aggtcgatgttacgn-3' \text{ where } n=g, \ t, \ c, \text{ or } i\]

- The correct feature key to describe the “n” is “variation” or “misc_difference” along with a “note” qualifier with the value “n=g, t, c, or i.”

- One of the alternatives is “i”, or inosine, which is a modified nucleotide. You can also include a “modified_base” feature key to note that the n may be “i”, but it is not necessary.
Compliance with ST.26 requirements

• **Common mistake:** use of “SITE” or “MOD_RES” to describe an amino acid sequence variant.

  Example disclosure:

  MDCWIKSLSX  where X= P, L, I, or V

• The correct feature key to describe the “X” is “VAR_SEQ” or “VARIANT” along with a “note” qualifier with the value “X= P, L, I, or V.”
Compliance with ST.26 requirements

• **Common mistake:** use of “misc_feature” or “variation” or “misc_difference” to describe a modified residue

  Example disclosure:

  5’-aggtcgatgttacgn-3’ where n = i (inosine)

• The correct feature key to describe the “n” is “modified_base” along with a “mod_base” qualifier with the value “i.”

• One of the alternatives is “i”, or inosine, which is a modified nucleotide. You can also include a “modified_base” feature key to note that the n may be “i”, but it is not necessary.
Compliance with ST.26 requirements

• **Helpful hint** – describing modified residues at multiple locations within a nucleotide sequence using a single feature key.

Example disclosure:

```
atggtc*gatgttac*c*c*ggatgtggtc*gatgttaatgttattac*
```

1 5 10 15 20 25 30 35 40

where c* = 5-methylcytidine

- How can feature keys be used to describe these modified residues?
Compliance with ST.26 requirements

• Method 1 – use a separate feature key to describe each modified cytosine or contiguous span of modified cytosines.

atggtc*gatgttaac*c*c*ggatgtggtc*gatgttaatgttattac*
1 5 10 15 20 25 30 35 40

Feature key 1:
“modified_base”
location = 6
qualifier “mod_base”
qualifier value = m5c

Feature key 2:
“modified_base”
location = 14..16
qualifier “mod_base”
qualifier value = m5c

Feature key 3:
“modified_base”
location = 26
qualifier “mod_base”
qualifier value = m5c

Feature key 4:
“modified_base”
location = 43
qualifier “mod_base”
qualifier value = m5c
Compliance with ST.26 requirements

- Method 2 – use a single feature key with an “order” location operator (ST.26 pgh. 67).

Feature key 1:
“modified_base”
location = order(6,14..16,26,43)
qualifier “mod_base”
qualifier value = m5c

Note that the “order” location operator can only be used for locations in nucleotide sequences.
Questions?
ST.26 sequence listings in international (PCT) applications
**ST.26 sequence listings in PCT apps**

- A sequence listing (SL) is required as part of the description where the PCT applicant contains nucleotide and/or amino acid sequences.
  - PCT Rule 5.2: “Where the international application contains disclosure of nucleotide and/or amino acid sequences... the description shall include a sequence listing part of the description...”
- The SL must comply with WIPO Standard ST.26 and the PCT Administrative Instructions.
- A SL CANNOT be text (.TXT) format and CANNOT be filed on sheets of paper or as a separate PDF file.
ST.26 sequence listings in PCT apps

- A SL included with IA on the international filing date (IFD) and indicated as part of the application within Box No. IX of the Request (PCT/RO/101) checklist will form “part of the international application.”

- Filing ST.26 SL using the USPTO electronic filing system, ST.26 SL must be filed via Patent Center.
  - EFS-Web does not accept ST.26 XML sequence listings. If SL is greater than 100MB, on physical data carrier (disc) labeled “sequence listing” filed on the same day as IA.

- An SL filed with RO/US must be in English only, including any language dependent text.
ST.26 sequence listings in PCT apps

- The request (PCT/RO/101) must properly identify ST.26 XML file within the checklist (box IX(f)) filed using Patent Center.

- FYI: new request form for IA filed on/after July 1, 2022 has updated Box No. IX check list (see left).
ST.26 sequence listings in PCT apps

- PDF request form has option when filing a sequence listing in an IA with RO/US “forming Part of the international application” submitted on a physical carrier (disc) on the same day as IA.

- PDF request form when IA is filed on paper with SL on physical carrier (disc) (see left).
ST.26 sequence listings in PCT apps

- Request form using ePCT to indicate SL is “forming part of the international application” (see left).

- **Note:** As of July 2022, WIPO is no longer providing support for PCT-SAFE and applicants should no longer generate a request form using PCT-SAFE.

  o **FYI:** ePCT must be used to generate the new request form.

<table>
<thead>
<tr>
<th>IX</th>
<th>Check list</th>
<th>Number of sheets</th>
<th>Electronic file(s) attached</th>
</tr>
</thead>
<tbody>
<tr>
<td>IX-1</td>
<td>Request (including declaration sheets)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX-2</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX-3</td>
<td>Claims</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX-4</td>
<td>Abstract</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX-5</td>
<td>Drawings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX-6a</td>
<td>Sequence listing part of the description (also to be used for the purposes of international search)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>IX-7</td>
<td>TOTAL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX-8</td>
<td>Accompanying items</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fee calculation sheet</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

0-4 Form PCT/RO/101 Request
0-4-1 Prepared Using
  ePCT-Filing for data package download
  Version
0-5 Petition

The undersigned requests that the present international application be processed pursuant to the Patent Cooperation Treaty.
ST.26 sequence listings in PCT apps

- A SL may “form part of the International application,” filed after the IFD where the SL was:
  - included in the application under PCT Rule 20.5(b) or (c), or PCT Rule 20.5bis(b) or (c); or
  - considered to have been contained in application under PCT Rule 20.6(b); or
  - corrected under PCT rule 26; or
  - rectified under PCT Rule 91; or
  - added as an amendment under Article 34.

*Amendment under Article 34 requires the filing of a Demand to the International Preliminary Examination Authority (IPEA).
Where sequences are disclosed, but no SL in XML format was provided, the International Searching Authority (ISA) may issue an invitation to furnish Sequence Listing (Form PCT/ISA/225) under PCT rule 13ter.1.

- SL filed in response to PCT/ISA/225 does “not forming part of the international application” since it is furnished for the purposes of the international search or international preliminary examination.
- SL filed in response to PCT/ISA/225 must be accompanied by a late furnishing fee under 37 C.F.R. 1.445(a)(5) when filed with ISA/US.
- SL in response to PCT/ISA/225 must be accompanied by the “a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.”
- Time period for reply to the PCT/ISA/225 is one month and is NOT extendable.
- Failure to timely response to the PCT/ISA/225 may result in search and examination by the ISA to the extent that a meaningful search can be carried out without the sequence listing under PCT (Rule 13ter.1(d)).
- SL in response to PCT/ISA/225 must be filed with the chosen ISA.
- IF ISA is US, SL can be filed using Patent Center or if SL is greater than 100MB, on physical data carrier (disc).
ST.26 sequence listings in U.S. national stage applications (35 U.S.C §371)
ST.26 sequence listings in §371 apps

• A SL in a U.S. national stage application under 35 U.S.C. §371 must be in compliance with 37 C.F.R. §1.831-1.835

• Where the ST.26 sequence listing XML forms part of the international application and the international publication has occurred, the USPTO will obtain the ST.26 sequence listing XML as part of the Article 20 documents from the International Bureau for the national stage application.
  
  o Where sequence listing formed part of the IA and the USPTO obtain the ST.26 sequence listing XML as part of the Article 20 documents, an incorporation by reference, under 37 C.F.R. 1.834(c), is NOT required in the 35 U.S.C. §371 application.

  o Where the ST.26 SL XML formed part of the international application, applicant should not provide a SL in the U.S. National Stage upon filing the U.S. National Stage application.

  o A SL provided under PCT Rule 13ter.1 in the IA is does “not forming part of the international application” and will not be retrieved by the USPTO.
Only four circumstances under which an ST.26 sequence listing XML should be submitted at time of entry into national phase under 35 U.S.C. 371.

1. The international phase application (PCT) discloses sequences but does not contain a sequence listing as “forming part of the international application” (Annex C of the Administrative Instructions under the PCT), as required under PCT Rule 5.2(a); or

2. The international phase application (PCT) discloses sequences and contains a sequence listing as “forming part of the international application”, as required under PCT Rule 5.2(a), but the PCT application was neither filed with the USPTO as receiving Office nor published by the International Bureau at the time of entry; or

3. A translation of the sequence listing as “forming part of the international application” is necessary because the values of one or more qualifiers that contain language-dependent free text were not provided in English in the ST.26 sequence listing XML of the PCT application; or

4. Applicant desires to amend the sequence listing as “forming part of the international application”.

ST.26 sequence listings in §371 apps
ST.26 sequence listings in §371 apps

• Where sequences are disclosed, but no SL in XML format was provided in the U.S. national stage application or the SL contained in the U.S. national stage application is not in compliance with 37 C.F.R. §1.831-1.834, the Office will issue “Notification to Comply With Requirements for Patent Applications Containing Nucleotide And/Or Amino acid Sequence Disclosure,” specifying that 37 CFR 1.831-1.835 apply to the present application and compliance is required.
A ST.26 sequence listing XML submitted in response to a “Notification to Comply With Requirements for Patent Applications Containing Nucleotide And/Or Amino acid Sequence Disclosure,” or as an addition to the application, or to replace an existing SL in the application is an amendment.

The submission must be compliance with 37 C.F.R. 1.835(a) or (b).

- Must include an Incorporation by reference statement in a substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 consisting of:
  - A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);
  - A copy of the amended specification without markings (clean version); and
  - A statement that the substitute specification contains no new matter.

A translated sequence listing XML is not considered and added or amended submission (37 CFR 1.835(d)(2)).
Questions?
ST.26 resources

**USPTO Resources**
- Sequence Listing Help Desk via SequenceHelpDesk@USPTO.GOV or 571-272-2510
- PCT Help Desk via 571-272-4300
- Patent Legal Administration Help Desk via 571-272-7701

**WIPO Resources**
- WIPO Standard ST.26 Knowledge Base: [https://www3.wipo.int/confluence/display/WSSKB/WIPO+Sequence+and+ST.26+Knowledge+Base+Home](https://www3.wipo.int/confluence/display/WSSKB/WIPO+Sequence+and+ST.26+Knowledge+Base+Home)