This document is scheduled to be published in the
Federal Register on 05/20/2022 and available online at
federalregister.gov/d/2022-10343, and on **govinfo.gov**

3510-16-P

DEPARTMENT OF COMMERCE

Patent and Trademark Office

37 CFR Part 1

[Docket No. PTO-P-2021-0006]

RIN 0651-AD53

Standard for Presentation of Nucleotide and Amino Acid Sequence Listings Using eXtensible Markup Language (XML) in Patent Applications to Implement WIPO Standard ST.26; Incorporation by Reference

AGENCY: United States Patent and Trademark Office, Department of Commerce.

ACTION: Final rule.

SUMMARY: The United States Patent and Trademark Office (USPTO or Office) is amending the rules of practice for submitting biological sequence data associated with disclosures of nucleotide and amino acid sequences in patent applications by incorporating by reference certain provisions of World Intellectual Property Office Standard ST.26 (WIPO Standard ST.26) into the USPTO rules of practice. Other conforming changes to accommodate the new rules of practice based on the new standard are also included. In addition to simplifying the process for applicants filing in multiple countries, the requirement to submit a single sequence listing in eXtensible Markup Language (XML) format, or "Sequence Listing XML," will result in better preservation, accessibility, and sorting of the submitted sequence data for the public.

DATES: *Effective date:* This final rule is effective on July 1, 2022. The incorporation by reference of certain publications listed in this rule is approved by the Director of the Federal Register as of July 1, 2022.

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. All other provisions of this final rule apply to all patent applications filed before, on, or after July 1, 2022.

FOR FURTHER INFORMATION CONTACT: Mary C. Till, Senior Legal Advisor, Office of Patent Legal Administration, Office of the Deputy Commissioner for Patents, at Mary.Till@uspto.gov or 571-272-7755; or Ali Salimi, Senior Legal Advisor, Office of Patent Legal Administration, Office of the Deputy Commissioner for Patents, at Ali.Salimi@uspto.gov or 571-272-0909.

SUPPLEMENTARY INFORMATION:

Table of Contents

I.   Background

a.   Summary of Changes

WIPO Standard ST.26 is the new international standard developed and adopted by WIPO and member states for purposes of presenting biotechnology information in patent applications. It will apply to international and national applications filed on or after July 1, 2022. New provisions in 37 CFR 1.831 through 1.835 implement WIPO Standard ST.26. 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 1.8211.825.

Under WIPO Standard ST.26 (as implemented by 37 CFR 1.831 through 1.835), patent applications that contain disclosures of nucleotide and/or amino acid sequences must present the associated biological sequence data in a standardized electronic format (a "Sequence Listing XML") as a separate part of the specification. In particular, WIPO Standard ST.26 permits applicants to submit a single, internationally acceptable sequence listing in a language-neutral format using specified International Nucleotide Sequence Database Collaboration (INSDC) identifiers in international applications filed under the Patent Cooperation Treaty (PCT) and in national and regional applications in the intellectual property offices (IPOs) of WIPO member states. As a result, a single sequence listing in compliance with WIPO Standard ST.26 can be prepared for use in the IPOs of WIPO member states.

For applications filed on or after July 1, 2022, the changes in this final rule include the: (1) creation of new rules (37 CFR 1.831 through 1.839) that incorporate by reference WIPO Standard ST.26; (2) use of INSDC sequence data elements to replace numeric identifiers used in the previous Standard ST.25 for the submission of nucleotide and/or amino acid sequences; (3) modification of rules of practice to include reference to a "Sequence Listing XML"; (4) elimination of the ability to file a paper or Portable Document Format (PDF) copy of nucleotide and/or amino acid sequences; (5) elimination of the option to include within a "Sequence Listing XML," sequences with

fewer than 4 specifically defined amino acids and fewer than 10 specifically defined nucleotides; and (6) clarification and simplification of the rules to aid in understanding the requirements set forth.

b.   Introduction

In an effort to streamline and reduce existing procedural requirements and to implement WIPO Standard ST.26, the USPTO is amending its rules of practice (by adding 37 CFR 1.831 through 1.839) for submitting biological sequence data associated with disclosures of nucleotide and/or amino acid sequences in patent applications filed on or after July 1, 2022. These changes also respond to the needs of our customers to comply with WIPO Standard ST.26.

To decrease the burden on applicants who file patent applications containing nucleotide and/or amino acid sequences internationally, the USPTO has worked with other WIPO member states as part of the Committee on WIPO Standards (CWS) to develop a single, internationally acceptable sequence listing standard for use in patent applications filed in those member states. Beginning in October of 2010, the CWS established a task force to propose a revised standard for the filing of nucleotide and/or amino acid sequence listings in XML file format. To obtain public input on the content of WIPO Standard ST.26, the USPTO issued requests for comments in 2012 and 2016. *See* Request for Comments on the Recommendation for the Disclosure of Sequence Listings Using XML (Proposed ST.26), 77 FR 28541 (May 15, 2012); and Standard ST.26-Request for Comments on the Recommended Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings Using XML (eXtensible Markup Language), 81 FR 74775 (October 27, 2016). The adopted version of WIPO Standard ST.26 takes those comments into account. To achieve the goals WIPO and its member states (including the United States) set out by developing the sequence listing standard for presenting data consistently across all IPOs, all WIPO member states agreed to implement WIPO

Standard ST.26 for international and national applications filed on or after July 1, 2022. Therefore, in view of this final rule, applications filed in the United States on or after July 1, 2022, will need to conform to WIPO Standard ST.26 as implemented in 37 CFR 1.831 through 1.839, which requires submitting sequence listings in XML format.

Under the final rule, applications that claim benefit or priority to an earlier application, where the earlier application contained a sequence listing that complied with the requirements of Standard ST.25 or other earlier requirements, must comply with the new rules that incorporate by reference WIPO Standard ST.26. To facilitate compliance, WIPO, with input from WIPO member states, developed WIPO Sequence, a sequence listing authoring and validating tool that applicants can use to prepare and validate their sequence listings in XML format, as discussed below. The USPTO is adding to the patent rules (37 CFR part 1) by incorporating by reference WIPO Standard ST.26, and providing conforming amendments to the current rules.

To ensure that biological sequence data associated with the disclosures of nucleotide and/or amino acid sequences in patent applications can be widely disseminated and searchable by the public and IPOs, the USPTO works with the National Center for Biotechnology Information (NCBI) on the inclusion of patent sequence data in the GenBank searchable database. For the NCBI to include all sequence data from the USPTO, the data must be provided in INSDC format so it is compatible with GenBank. The Standard ST.25 format sequence listings cannot be readily converted to INSDC format, resulting in only a fraction of patent sequence information appearing in GenBank. This data loss limits the sequence information available to the public and exchanged with other sequence database providers (e.g., the National Institute of Genetics (NIG) in Japan, the DNA Data Bank of Japan (DDBJ), and the European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI). WIPO has been working with the WIPO member states to create, adopt, and implement WIPO Standard ST.26 for sequence listing

submissions in XML file format, which has the INSDC data elements to address the data loss. WIPO Standard ST.26 aims to enhance the accuracy and quality of biological sequence data that is publicly disseminated. With the adoption and implementation of WIPO Standard ST.26, more complete biological sequence data from patents and patent applications will be included in GenBank and thus be accessible by the public. The change from American Standard Code for Information Interchange (ASCII) plain text format to XML format will result in sequence data having computer tags that facilitate sorting and retrieving and will permit ease of access to the data. Additionally, the NCBI plans to stop accepting data in Standard ST.25 format for inclusion in GenBank approximately three to five years after the WIPO Standard ST.26 transition date (July 1, 2022).

c.   *Incorporation by Reference of WIPO Standard ST.26*

The WIPO "Handbook on Industrial Property Information and Documentation" sets forth standards for the presentation of data in many contexts. WIPO Standard ST.26 is titled "RECOMMENDED STANDARD FOR THE PRESENTATION OF NUCLEOTIDE AND AMINO ACID SEQUENCE LISTINGS USING XML (EXTENSIBLE MARKUP LANGUAGE)." The CWS adopted the current version (version 1.5) in November of 2021. In October of 2021, at the Assemblies of Member States of WIPO, the member states agreed on July 1, 2022, as the implementation date of WIPO Standard ST.26. This final rule incorporates by reference WIPO Standard ST.26. The standard is available from WIPO, 34 chemin des Colombettes, 1211 Geneva 20 Switzerland, www.wipo.int. and also as provided for in 37 CFR 1.839.

WIPO Standard ST.26 is composed of eight documents, namely, the main body of the standard, a first annex (Annex I) setting forth the controlled vocabulary for use with the main body, a second annex (Annex II) setting forth the Document Type Definition

(DTD) for the Sequence Listing, a third annex (Annex III) containing a sequence listing specimen (XML file), a fourth annex (Annex IV) setting forth the character subset from the Unicode Basic Latin Code Table, a fifth annex (Annex V) setting forth additional data exchange requirements for IPOs, a sixth annex (Annex VI) containing a guidance document with illustrated examples, and a seventh annex (Annex VII) setting forth recommendations for the transformation of a sequence listing from Standard ST.25 format to WIPO Standard ST.26 format, including guidance on how to avoid adding or deleting subject matter.

The main body of WIPO Standard ST.26 defines the disclosures of nucleotide and/or amino acid sequences in patent applications that must be presented in a sequence listing in XML format in the manner specified in the standard. As detailed in paragraph eight of the main body, a sequence listing in XML format must not include any sequences having fewer than 10 specifically defined nucleotides, or fewer than 4 specifically defined amino acids. If such sequences are included in the disclosure, they must not be assigned a sequence identification number. The main body establishes the requirements for the representation of nucleotide and/or amino acid sequences and the requirements for the XML file format for a sequence listing. Annex I contains controlled vocabulary that provides nucleotide base codes, lists of modified nucleotides and their abbreviations, amino acid codes, and a list of modified amino acids and their abbreviations. In addition, Annex I provides defined feature keys and qualifiers used for nucleotide and/or amino acid sequences in the XML file for a sequence listing. Annex I specifically identifies qualifiers with language-dependent "free text" values that may require translation for national and regional procedures. Annex II provides the DTD setting forth the technical specifications to which a submitted Sequence Listing XML must conform. Annex III provides a link to a specimen of a sequence listing that is compliant with WIPO Standard ST.26 and that shows a representation of an entire sequence listing in XML format.

Annex IV provides a table of the character subset from the Unicode Basic Latin Code that will be used in the XML file for the sequence listing. Annex V provides guidance to WIPO member states on how certain sequence elements should be populated when data is exchanged with database providers. Annex VI, containing the guidance document, ensures that all applicants and WIPO member states understand the requirements for inclusion and representation of sequence disclosures. This guidance document was developed, in part, to address concerns raised in response to the USPTO's requests for comments in 2012 and 2016, mentioned above. The guidance document illustrates the requirements of selected paragraphs in the main body of WIPO Standard ST.26 through specific examples of nucleotide and amino acid biological sequence data. Additionally, the document provides guidance on the manner in which biological sequence data is represented in a sequence listing in XML format that is compliant with WIPO Standard ST.26. Annex VII addresses the mandatory requirements of WIPO Standard ST.26, and the potential consequence of these requirements when transforming a compliant Standard ST.25 sequence listing into a WIPO Standard ST.26 sequence listing. Annex VII also provides detailed guidance on how to avoid adding or deleting subject matter due to the additional requirements of WIPO Standard ST.26.

d.   Benefits

Transitioning from rules based on WIPO Standard ST.25 (i.e., the basis for USPTO rules 37 CFR 1.821 through 1.825, regarding "Sequence Listings") to rules based on WIPO Standard ST.26 will be beneficial to both patent applicants filing sequence listings and IPOs receiving applications containing disclosures of nucleotide and/or amino acid sequences requiring sequence listings. WIPO Standard ST.26 provides clear requirements for what must be included in a sequence listing and how sequences must be represented. For example, it standardizes the representation of modified nucleotide sequences and amino acid sequences as well as variants derived from primary sequences.

Since WIPO Standard ST.26 contains a guidance document that illustrates the requirements for the inclusion and representation of biological sequence data, patent applicants will have a better understanding of the requirements for the presentation of biological sequence data in a compliant sequence listing under WIPO Standard ST.26 (as implemented by 37 CFR 1.831 through 1.839). Additionally, since WIPO Standard ST.26 only allows XML format (an electronic computer readable format), this final rule eliminates the potential for differences between a sequence listing filed in paper/PDF format and the required electronic computer readable format (CRF). As a further benefit, the IPOs of WIPO member states will no longer need to expend resources to process paper sequence listings and perform necessary checks on the contents of paper documents.

Unlike rules based on Standard ST.25, rules based on WIPO Standard ST.26 will allow patent applicants to file a single sequence listing with the USPTO (with the exception of changes to comply with national language requirements) that will be acceptable to the IPOs of all WIPO member states. Under Standard ST.25, IPOs have interpreted and enforced rules differently due to the imprecise language in that standard. This has resulted in the frustrating situation in which applicants generate sequence listings that may be accepted in one IPO but not another.

WIPO Standard ST.26 was drafted to precisely define what must and must not be included in a sequence listing and how sequences must be represented in a sequence listing. The "Guidance document with illustrated examples" in Annex VI of WIPO Standard ST.26 demonstrates the application of the rules to real-world sequence disclosure examples, reducing the possibility of misinterpretation by IPOs or applicants.

Due to the improved data structure of XML, transitioning to rules based on WIPO Standard ST.26 will increase the quality of the examination of patent applications

containing biological sequence data since a more comprehensive search will be possible. Sequence listings submitted in accordance with WIPO Standard ST.26 allow for targeted searching of both sequence annotation and newly required sequence types, such as D-amino acids, nucleotide analogues, and linear portions of branched sequences. Finally, sequence listing submissions under rules based on WIPO Standard ST.26 will enhance public database content, as they include the sequence annotations (e.g., feature keys and qualifiers) used by database providers to describe biological sequence data. WIPO Standard ST.26 standardizes sequence variant presentation, the annotation of modified and unusual residues, feature location descriptors, the use of feature keys and qualifiers, organism names, and the presentation of coding regions. Incorporation by reference of WIPO Standard ST.26 into USPTO rules promotes data exchange between the USPTO and the NCBI due to the use of INSDC identifiers required by database providers. The presence of additional data, as well as the enhanced compatibility to facilitate the exchange of data, will increase the value of database searches that relate to nucleotide and amino acid sequences for biotechnology stakeholders.

Requiring compliance with WIPO Standard ST.26 for an application filed on or after July 1, 2022, will reduce the complexity and cost of the long-term maintenance of information technology (IT) systems for accepting sequence listings in multiple formats, provide a clear implementation date, and facilitate the transition to the format requirements of database providers. In addition, a requirement to submit a single sequence listing in XML format will result in better preservation, accessibility, and sorting of the submitted sequence data for the public.

e.   *WIPO Authoring and Validation Tool (WIPO Sequence)*

To comply with rules based on WIPO Standard ST.26, patent applicants will be able to generate a sequence listing compliant with WIPO Standard ST.26 using WIPO

Sequence, a desktop application developed by WIPO and adopted by WIPO member states. WIPO Sequence has two functions: an authoring function and a validation function. Patent applicants will be able to author and validate their sequence listing using WIPO Sequence to comply with the requirements of WIPO Standard ST.26. Such a sequence listing will be accepted by all the IPOs of the WIPO member states. Thus, the burden of generating a sequence listing that is acceptable across all WIPO member states will be significantly decreased for patent applicants under WIPO Standard ST.26. This tool is downloadable, free of charge, from the WIPO website. The current version of WIPO Sequence is accessible at www.wipo.int/standards/en/sequence/index.html. This version, subject to updates, will allow the public to become familiar with the tool and its dual functionalities.

WIPO Sequence will allow a user to create and save (author) patent application data and biological sequence data in a project, validate the project to ensure all required information is present, and generate a sequence listing in WIPO Standard ST.26 XML format. Information can be entered into a project manually, or data can be imported from a source file in one of a number of file types. WIPO Sequence can import data from other WIPO Standard ST.26 projects, WIPO Standard ST.26 XML sequence listings, Standard ST.25 sequence listing text files, raw files, multi-sequence format files, and FASTA (FAST-All-a DNA and protein sequence alignment software package) files. Feature keys, qualifiers, and organism names are available to select from drop-down lists, simplifying the creation of sequence listings. Applicant and inventor names, as well as custom organism names, can be stored in WIPO Sequence for easy access. To facilitate the review of data entered into a project, WIPO Sequence can generate a "human-readable" version (a text version of the sequence data) of the sequence listing in addition to the XML sequence listing.

WIPO Sequence includes an integrated validation function that will alert users to most errors in a project or sequence listing data. The validation function generates a report that clearly lists every detected error, the location of the error, and the detected value of the error, along with a link to the sequence in question, thereby ensuring users can correct errors before generating a final sequence listing. While the validation function will alert a user to most errors in a project or sequence listing, there are a small number of errors that can be detected only by human review (for example, an inappropriate organism name). In those cases, the integrated validation function will list a "warning" in the validation report, reminding users of the applicable/relevant rule and urging them to check their input values before generating a final sequence listing.

A sequence listing in Standard ST.25 format cannot automatically be converted into WIPO Standard ST.26 format because certain data elements required for a sequence listing compliant with WIPO Standard ST.26 are not present in Standard ST.25. Therefore, conversion of a sequence listing in Standard ST.25 format to Standard ST.26 format necessarily requires additional input from the applicant. WIPO Sequence, supplemented by significant guidance from WIPO and the USPTO (in Annex VI and Annex VII of WIPO Standard ST.26), will help applicants accomplish this task. Users can import a Standard ST.25 sequence listing into a project, and WIPO Sequence automatically performs many of the necessary conversions. An Import Report is generated that alerts the user to all data conversions and lists all sequence entries that require additional input. In response to concerns raised regarding the USPTO's requests for comments in 2012 and 2016, the USPTO, in conjunction with WIPO, developed Annex VII to provide detailed guidance to help applicants avoid added or deleted subject matter when converting a sequence listing from Standard ST.25 format into Standard ST.26 format.

To ensure that IPOs can validate and accept sequence listing projects from applicants generated with WIPO Sequence, WIPO is developing a Standard ST.26 sequence listing validation tool, WIPO Sequence Validator. WIPO Sequence Validator will be for use by IPOs. WIPO Sequence Validator will be synchronized with the validation function in the WIPO Sequence tool. The USPTO is integrating WIPO Sequence Validator into its internal IT systems. The WIPO Sequence Validator will apply the same validation rules as WIPO Sequence. Therefore, filers will have a greater level of confidence that a sequence listing authored and validated by WIPO Sequence will comply with the USPTO rules for a "Sequence Listing XML" (37 CFR 1.831 through 1.835) and be accepted, given that the WIPO Sequence Validator that the USPTO will use is based on WIPO Standard ST.26.

f.   *Applicability*

In accordance with this final rule, an application that has a filing date on or after July 1, 2022, will be required to provide a "Sequence Listing XML" in accordance with 37 CFR 1.831 through 1.835 for disclosures of any nucleotide and/or amino acid sequences that meet the definitions of 37 CFR 1.831(a) and (b). This includes applications having an international filing date on or after July 1, 2022, that claim benefit or priority to applications with filing dates before July 1, 2022. Such applications include, but are not limited to, applications having one or more benefit or priority claims under 35 U.S.C. 119(e) (claiming the benefit of a provisional), 35 U.S.C. 120 (claiming the benefit as a continuation and/or continuation-in-part), 35 U.S.C. 121 (claiming the benefit as a divisional), 35 U.S.C. 365(c) (claiming the benefit as a continuing application to a PCT application), or 35 U.S.C. 119(a)-(d) or 35 U.S.C. 365(a) (claiming the priority to a foreign filed application or a prior filed PCT). If a prior application to which benefit or priority is claimed contains a "Sequence Listing" in Standard ST.25 format (in compliance with 37 CFR 1.821 through 1.825), the applicant will be required to convert

that "Sequence Listing" to WIPO Standard ST.26 format (a "Sequence Listing XML" in compliance with 37 CFR 1.831 through 1.835) for inclusion in the new application filed on or after July 1, 2022.

As provided in 35 U.S.C. 363, the filing date of an international stage application is also the filing date for the national stage application filed under 35 U.S.C. 371. Accordingly, for applications submitted under 35 U.S.C. 371, WIPO Standard ST.26 will apply to such applications based on the international filing date of the corresponding international application, rather than the date of submission of the national stage application in the USPTO.

Compliance with 37 CFR 1.831 through 1.835 (rules based on WIPO Standard ST.26) is also applicable to any reissue application filed on or after July 1, 2022, where the disclosure or claims contain nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(a) or (b). The filing date of the originally granted patent for which reissue is sought is not relevant in determining the applicability date of this final rule.

Relying on the actual filing date of an application to determine whether sequence information must conform to 37 CFR 1.821 through 1.825 (rules based on Standard ST.25) or 37 CFR 1.831 through 1.835 (rules based on WIPO Standard ST.26) will simplify the application of the sequence rules, both for the USPTO and the applicant. Though 37 CFR 1.821 through 1.825 are not revised by this final rule, note that 37 CFR 1.821 through 1.825 will not be applicable to applications filed on or after July 1, 2022, as a result of this final rule.

For applications filed on or after July 1, 2022, the USPTO patent electronic filing system will prohibit an applicant from submitting both a "Sequence Listing XML" (a sequence listing that conforms to WIPO Standard ST.26 as implemented in 37 CFR 1.831 through 1.835) and a "Sequence Listing" (a sequence listing that conforms to ST.25 as

implemented in 37 CFR 1.821 through 1.825) in the same submission. Filing a "Sequence Listing" in an application filed on or after July 1, 2022, will result in a notice informing applicant that the submission fails to comply with 37 CFR 1.831 through 1.834 and will require submission of a "Sequence Listing XML."

While implementing regulations and procedures for ST.26, the USPTO recognized that an applicant might erroneously provide a "Sequence Listing" (one in ASCII plain text file format) even though a "Sequence Listing XML" is required. Therefore, in the rare circumstance in which a "Sequence Listing" is submitted in an application filed on or after July 1, 2022, the "Sequence Listing" present in the Office file wrapper of the application at issue may be used to provide support for the submission of a compliant "Sequence Listing XML." The applicant's reliance on the "Sequence Listing" to support the compliant "Sequence Listing XML" would be by way of the safeguard under 37 CFR 1.57(b), if an earlier filed application contains a proper "Sequence Listing" in .txt file format, or via a grantable petition under 37 CFR 1.182, only if the application does not have a proper benefit or priority claim present on the filing date to an earlier filed application.

An applicant may rely on the provisions in 37 CFR 1.57(b), as described in the Manual of Patent Examining Procedure at section 217, to support the required "Sequence Listing XML" as an "inadvertently omitted portion of the specification or drawing(s)." To rely on 37 CFR 1.57(b), a compliant "Sequence Listing" must have been submitted in an earlier filed application to which the present application makes a proper benefit or priority claim, and the "Sequence Listing" was present on the filing date of the earlier filed application (i.e., the earlier filed application contains a compliant "Sequence Listing" submitted under 37 CFR 1.821(c)(1) as an ASCII plain text file (with a proper incorporation by reference statement in the specification), 37 CFR 1.821(c)(2) as a PDF copy, or 37 CF 1.821(c)(3) on physical sheets of paper). An applicant would be required

to submit: (1) a compliant "Sequence Listing XML" under 37 CFR 1.835(a)(1); (2) a statement identifying where the inadvertently omitted portion of the specification can be found (e.g., identifying the nucleotide and/or amino acid sequence information in the compliant "Sequence Listing" from the earlier filed application that forms the basis for the "Sequence Listing XML"), *see* 37 CFR 1.835(a)(3); (3) a statement identifying the nucleotide and/or amino acid sequences of the "Sequence Listing," submitted (in the earlier filed application) under 37 CFR 1.821(c)(1) as an ASCII plain text file (with a proper incorporation by reference statement in the specification), 37 CFR 1.821(c)(2) as a PDF copy, or 37 CFR 1.821(c)(3) as physical sheets of paper, which forms the basis for the compliant "Sequence Listing XML"; (4) a statement that the "Sequence Listing XML" does not introduce new matter into the application, *see* 37 CFR 1.835(a)(4); and (5) a statement that all or a portion of the specification or drawings, as found in the "Sequence Listing XML," were inadvertently omitted from the application. The availability of relief under 37 CFR 1.57(b) precludes the filing of a grantable petition under 37 CFR 1.182 seeking the same relief.

A petition under 37 CFR 1.182 would require: (1) a compliant "Sequence Listing XML" under 37 CFR 1.835(a)(1); (2) a statement identifying the nucleotide and/or amino acid sequence information of the "Sequence Listing" submitted as an ASCII plain text file that forms the basis for the "Sequence Listing XML" (i.e., identifying the nucleotide and/or amino acid sequence information found in the "Sequence Listing" from the earlier submitted ASCII "Sequence Listing") that is relied on for submission of a compliant "Sequence Listing XML," *see* 37 CFR 1.835(a)(3); and (3) a statement that the "Sequence Listing XML" does not introduce new matter into the application, as required by 37 CFR 1.835(a)(4). In such circumstances, for record retention purposes, any "Sequence Listing" submitted as an ASCII plain text file will be retained in the official record for the application.

II.   Discussion of Specific Rules

Section 1.52: Section 1.52 (e)(1)(ii) is amended to include reference to a "Sequence Listing XML" submitted under § 1.831(a) in compliance with §§ 1.832 through 1.834.

Section 1.52(e)(3)(iii) is amended to more explicitly indicate that the contents of each read-only optical disc must be in ASCII plain text and if compressed, must be compressed in accordance with § 1.58 for "Large Tables," § 1.96 for a "Computer Program Listing Appendix," or § 1.824 for a "Sequence Listing" or CRF of the "Sequence Listing," as applicable.

Section 1.52(e)(3)(iv) is added to require that the contents of each read-only optical disc for a "Sequence Listing XML" must be in XML file format and, if compressed, must be compressed in accordance with § 1.834.

Section 1.52(e)(7) is amended to add that any amendment to the information on a read-only optical disc previously submitted in relation to a "Sequence Listing XML" must be made by way of a replacement read-only optical disc in accordance with § 1.835(b).

Section 1.52(f)(1) is amended to add that any XML file submitted on a read-only optical disc is excluded from the application size fee determination if the read-only optical disc contains a "Sequence Listing XML" in compliance with § 1.831(a). The provision at 35 U.S.C. 41(a)(1)(G) provides the basis for excluding "any sequence listing," when filed in electronic medium, from the application size fee determination. A "Sequence Listing XML" is considered as "any sequence listing."

Section 1.52(f)(1)(i) is amended to reference any "Sequence Listing XML" in compliance with § 1.831(a).

Section 1.52(f)(2) is amended to indicate that any XML file, submitted via the USPTO patent electronic filing system for a "Sequence Listing XML" in compliance with § 1.831(a) is excluded from the application size fee determination. The provision at 35 U.S.C. 41(a)(1)(G) provides the basis for excluding "any sequence listing," when filed in an electronic medium, from the application size fee determination. A "Sequence Listing XML" is considered as "any sequence listing."

Section 1.52(f)(2)(i) is amended to add a reference to any "Sequence Listing XML" in compliance with § 1.831(a).

Section 1.52(f)(3) is amended to add that any "Sequence Listing XML" of 300 MB-800 MB is subject to the surcharge set forth in § 1.21(o)(1) and also add that any "Sequence Listing XML" over 800 MB is subject to the surcharge set forth in § 1.21(o)(2).

Section 1.53: Section 1.53(c)(4) is revised to indicate that a separate sequence listing in a provisional application disclosing nucleotide and/or amino acid sequences is not required, but any biological sequence data submitted in a provisional application filed on or after July 1, 2022, must be a "Sequence Listing XML" in compliance with §§ 1.831 through 1.834. This change does not apply to provisional applications filed before July 1, 2022.

Section 1.77: Section 1.77(b)(5) is amended to reorganize the provisions to § 1.77(b)(5)(i) for an incorporation by reference statement for ASCII plain text files submitted for a "Computer Program Listing Appendix" (§ 1.77(b)(5)(i)(A)), a "Sequence Listing" (§ 1.77(b)(5)(i)(B)), and "Large Tables" (§ 1.77(b)(5)(i)(C)). Section 1.77(b)(5)(ii) is added to provide for the provisions for an incorporation by reference statement for a "Sequence Listing XML" submitted via the USPTO patent electronic filing system or on one or more read-only optical discs.

Section 1.121: Section 1.121(b) is amended to revise the reference for a "Sequence Listing" and eliminate the reference to a CRF of a "Sequence Listing," since a separate CRF (under § 1.821(e)(1) or (2)) is not part of the specification. The amendment also adds an exception to amendment practice for a "Sequence Listing XML" (§ 1.831(a)).

Section 1.121(b)(6) is amended to require that changes to a "Sequence Listing XML" be made in accordance with § 1.835.

Section 1.173: The heading of § 1.173(b)(1) is amended to include "'Sequence Listing XML' (§ 1.831(a))."

Section 1.173(b)(1)(i) is amended to add an exception to reissue amendment practice for a "'Sequence Listing XML' (§ 1.831(a))."

Section 1.173(b)(1)(ii) is amended to provide that changes to a "Sequence Listing XML" must be made in accordance with § 1.835.

Section 1.173(d) is amended to add a "Sequence Listing XML" (§ 1.831(a)) among the items that are excluded from the manner of making amendments in a reissue application. Reference to specific CFR provisions for "Large Tables" (§ 1.58(c)), a "Computer Program Listing Appendix" (§ 1.96(c)), and a "Sequence Listing" (§ 1.821(c)) were added.

Section 1.211: Section 1.211(c) is amended to add a "Sequence Listing" in compliance with §§ 1.821 through 1.825 (if applicable) for an application filed before July 1, 2022, and a "Sequence Listing XML" in compliance with §§ 1.831 through 1.835 (if applicable) for an application filed on or after July 1, 2022, to the currently listed items that may delay application publication if not present.

Section 1.495: Section 1.495(c)(5) is amended to delineate between translations needed for a sequence listing in international applications entering the national stage in the United States and having an international filing date before July 1, 2022, and a sequence listing in XML format for international applications entering the national stage in the United States and having an international filing date on or after July 1, 2022. Specifically, the amendment indicates that a sequence listing need not be translated for national stage entry if it complies with PCT Rule 12.1(d) and the description complies with PCT Rule 5.2(b) for applications having an international filing date before July 1, 2022. However, the amendment indicates that a sequence listing in XML format must be translated for national stage entry if it was submitted in an international application having an international filing date on or after July 1, 2022, with non-English language values for any language-dependent free text qualifiers. Note that an invention title is not considered a "language-dependent free text qualifier" for purposes of this rule, and translation of the invention title is not required.

Section 1.495(c)(5), as well as §§ 1.833(b)(3) and 1.835(d)(2) as discussed below, were proposed to require that the "Sequence Listing XML" contain at least one invention title in English. This proposal has not been adopted in this final rule. The proposed requirement for a translation of the title into English was not adopted in the final rule because applicants in the international phase need only provide a title in the language of filing, which can be in a language other than English.

Section 1.530: The heading of § 1.530(d)(1) is amended to include "'Sequence Listing XML' (§ 1.831(a))."

Section 1.530(d)(1)(i) is amended to add an exception to reexamination amendment practice for a "'Sequence Listing XML' (§ 1.831(a))."

Section 1.530(d)(1)(ii) is amended to provide that changes to a "Sequence Listing XML" must be made in accordance with § 1.835.

Section 1.704: Section 1.704(f) is amended to add a "Sequence Listing XML" in compliance with §§ 1.831 through 1.835 (if applicable) to the list of items required for an application filed under 35 U.S.C. 111(a) to be in condition for examination for purposes of calculating a reduction in patent term adjustment. The amendment also adds a "Sequence Listing XML" in compliance with §§ 1.831 through 1.835 (if applicable) to the list of items that must be submitted in an international application for such an application to be in condition for examination when the application has entered the national stage as defined in § 1.491(b). Lastly, the rule is also amended to add a "Sequence Listing XML" in compliance with §§ 1.831 through 1.835 (if applicable) to the current list of items required for an application to be considered compliant, for purposes of determining a patent term adjustment reduction, on the filing date of the latest reply (if any) correcting the papers, drawings, or "Sequence Listing" that is prior to the date of the mailing of either an action under 35 U.S.C. 132 or a notice of allowance under 35 U.S.C. 151, whichever occurs first. Lastly, the term "Sequence Listing" replaces "sequence listing," since §§ 1.821 through 1.825 specifically define a "Sequence Listing."

Section 1.831: Section 1.831 is added to provide the heading of "requirements for patent applications filed on or after July 1, 2022, having disclosures of nucleotide and/or amino acid sequences."

Section 1.831(a) is added to specify that patent applications disclosing nucleotide and/or amino acid sequences by enumeration of their residues, as defined in paragraph (b) of the section, must contain, as a separate part of the disclosure, a "Sequence Listing XML". Disclosed nucleotide and/or amino acid sequences that do not meet the definition

in paragraph (b) of the section must not be included in the "Sequence Listing XML." The "Sequence Listing XML" contains information of the nucleotide and/or amino acid sequences disclosed in the patent application using the symbols and format in accordance with the requirements of §§ 1.832 through 1.834.

Section 1.831(b)(1) and (2) are added to define the nucleotide and amino acid sequences for which a "Sequence Listing XML" is required. Specifically, nucleotide and/or amino acid sequences, as used in these rules, encompass: an unbranched sequence or linear region of a branched sequence containing 4 or more specifically defined amino acids, wherein the amino acids form a single peptide backbone or an unbranched sequence or linear region of a branched sequence of 10 or more specifically defined nucleotides, wherein adjacent nucleotides are joined by a 3' to 5' (or 5' to 3') phosphodiester linkage or, for nucleotide analogs, any chemical bond that results in an arrangement of adjacent nucleobases that mimics the arrangement of nucleobases in naturally occurring nucleic acids.

Section 1.831(c) is added to state that, where the description or claims of a patent application discuss a nucleotide and/or amino acid sequence that is set forth in the "Sequence Listing XML" in accordance with paragraph (a) of the section, reference must be made to the sequence by use of the sequence identifier, preceded by "SEQ ID NO:" or the like, in the text of the description or claims, even if the sequence is also embedded in the text of the description or claims of the patent application. Where a sequence is presented in a drawing, reference must be made to the sequence by use of the sequence identifier (§ 1.832(a)), either in the drawing or in the Brief Description of the Drawings, where the correlation between multiple sequences in the drawing and their sequence identifiers (§ 1.832(a)) in the Brief Description is clear. The use of SEQ ID NO: is preferred, but including "or the like" is intended to ensure that a formalities notice is not sent when an application uses, for example, "SEQ NO." or "Seq. Id. No." or any similar

identification of an amino acid or nucleotide sequence in the description or claims where it is clear that a sequence from the "Sequence Listing XML" is shown in the description, claims, or drawings. When identifying the sequence in the description, claims, or drawings, the numeric sequence identifier from the "Sequence Listing XML" must identify the same sequence.

Section 1.831(d) is added to define the expression "enumeration of its residues," consistent with the definition in paragraph 3(c)(i) or (ii) of WIPO Standard ST.26 (incorporated by reference, *see* 37 CFR 1.839).

Section 1.831(e) is added to define the expression "specifically defined," consistent with the definition in paragraph 3(k) of WIPO Standard ST.26.

Section 1.831(f) is added to define the expression "amino acid," consistent with the definition in paragraph 3(a) of WIPO Standard ST.26.

Section 1.831(g) is added to define the expression "modified amino acid," consistent with the definition in paragraph 3(e) of WIPO Standard ST.26.

Section 1.831(h) is added to define the expression "nucleotide," consistent with paragraphs 3(f) and 3(g) of WIPO Standard ST.26.

Section 1.831(i) is added to define the expression "modified nucleotide," consistent with paragraph 3(f) of WIPO Standard ST.26.

Section 1.831(j) is added to indicate that a "Sequence Listing XML" must not include any sequences having fewer than 10 specifically defined nucleotides, or fewer than 4 specifically defined amino acids. Even though § 1.831(a) states that "[d]isclosed nucleotide or amino acid sequences that do not meet the definition in paragraph (b) of this section must not be included in the "Sequence Listing XML,"" adding § 1.831(j)

makes explicit the prohibition of including such sequences in the "Sequence Listing XML."

Section 1.832: Section 1.832 is added to provide the manner in which a nucleotide and/or amino acid sequence is represented in the "Sequence Listing XML" part of a patent application having a filing date on or after July 1, 2022.

Section 1.832(a) is added to define the requirements for the representation of sequences in the "Sequence Listing XML" part of the application. Specifically, each nucleotide and/or amino acid sequence represented in the "Sequence Listing XML" must be assigned a separate sequence identifier, and sequence identifiers must begin with the number 1 and increase sequentially by integers, as defined in paragraph 10 of WIPO Standard ST.26 (incorporated by reference, *see* 37 CFR 1.839).

Section 1.832(b)(1) through (4) are added to define the requirements for the representation of nucleotide sequence data in the "Sequence Listing XML." Specifically, a nucleotide sequence must be represented in the manner described in paragraphs 11-12 of WIPO Standard ST.26. All nucleotides, including nucleotide analogs, modified nucleotides, and "unknown" nucleotides, within a nucleotide sequence must be represented and described using symbols in the manner described in paragraphs 13-19 and 21 of WIPO Standard ST.26. For a region containing a known number of contiguous "a," "c," "g," "t," or "n" residues for which the same description applies, the entire region may be jointly described as provided in paragraph 22 of WIPO Standard ST.26.

Section 1.832(c)(1) through (4) are added to define the requirements for the representation of amino acid sequence data in the "Sequence Listing XML." Specifically, an amino acid sequence must be represented in the manner described in paragraphs 24 and 25 of WIPO Standard ST.26. All amino acids, including modified amino acids and "unknown" amino acids, within an amino acid sequence must be represented and

described using symbols in the manner described in paragraphs 26-30 and 32 of WIPO Standard ST.26. For a region containing a known number of contiguous "X" residues for which the same description applies, the entire region may be jointly described as provided in paragraph 34 of WIPO Standard ST.26.

Section 1.832(d) is added to define the manner in which a single continuous sequence, derived from one or more non-contiguous segments of a larger sequence, or of segments from different sequences, must be represented in the "Sequence Listing XML," as described in paragraph 35 of WIPO Standard ST.26.

Section 1.832(e) is added to define the manner in which a nucleotide and/or amino acid sequence that contains regions of specifically defined residues separated by one or more regions of contiguous "n" or "X" residues of specified length must be represented in the "Sequence Listing XML," as described in paragraph 36 of WIPO Standard ST.26.

Section 1.832(f) is added to define the manner in which a nucleotide and/or amino acid sequence that contains regions of specifically defined residues separated by one or more gaps of an unknown or undisclosed number of residues must be represented in the "Sequence Listing XML," as described in paragraph 37 of WIPO Standard ST.26.

Section 1.833: Section 1.833 is added to describe the requirements for a "Sequence Listing XML," which is required by § 1.831(a) for disclosures of nucleotides and/or amino acid sequences in patent applications with a filing date on or after July 1, 2022, to comply with WIPO Standard ST.26 (incorporated by reference, *see* 37 CFR 1.839).

Section 1.833(a) is added to require that the "Sequence Listing XML" must be presented as a single XML 1.0 file and encoded using Unicode UTF-8. Section 1.833(a)

also incorporates by reference paragraphs 40 and 41, and Annex IV of WIPO Standard ST.26 for character sets.

Section 1.833(b)(1) is added to require that the "Sequence Listing XML" presented in accordance with § 1.833(a) must further be valid according to the DTD as presented in Annex II of WIPO Standard ST.26.

Section 1.833(b)(2) is added to recite that a "Sequence Listing XML" must comply with the requirements of WIPO Standard ST.26, to include the items enumerated in § 1.833(b)(2)(i) through (v) as discussed in the following paragraphs.

Section 1.833(b)(2)(i) is added to require that the "Sequence Listing XML" contain an XML declaration as defined in paragraph 39(a) of WIPO Standard ST.26.

Section 1.833(b)(2)(ii) is added to require that the "Sequence Listing XML" contain a document type declaration as defined in paragraph 39(b) of WIPO Standard ST.26.

Section 1.833(b)(2)(iii) is added to require that the "Sequence Listing XML" contain a root element as defined in paragraph 43 of WIPO Standard ST.26.

Section 1.833(b)(2)(iv) is added to require that the "Sequence Listing XML" contain a general information part that complies with paragraphs 45, 47, and 48 of WIPO Standard ST.26, as applicable.

Section 1.833(b)(2)(v) is added to require that the "Sequence Listing XML" contain a sequence data part that complies with paragraphs 50-55, 57, 58, 60-69, 71-78, 80-87, 89-98, and 100 of WIPO Standard ST.26, as applicable.

Section 1.833(b)(3) is added to require that an INSDQualifier_value element includes a value for that element in English for each language-dependent free text

qualifier in the "Sequence Listing XML," as required by § 1.52(b)(1)(ii), and where an INSDQualifier_value element is defined in paragraphs 76 and 85-87 of WIPO Standard ST.26. The proposed requirement for a translation of the title into English was not adopted in the final rule because applicants in the international phase need only provide a title in the language of filing, which can be in a language other than English.

Section 1.834: Section 1.834 is added to provide details on the form and format for nucleotide and/or amino acid sequence submissions as the "Sequence Listing XML" in patent applications filed on or after July 1, 2022.

Section 1.834(a) is added to indicate that a "Sequence Listing XML" in Unicode UTF-8 created by any means (e.g., text editors, nucleotide/amino acid sequence editors, or other custom computer programs) in accordance with §§ 1.831 through 1.833 must: (1) be compatible with a PC or Mac® and with MS-DOS®, MS-Windows®, Mac OS®, or Unix®/Linux® operating systems; (2) be in XML format, where all permitted printable characters (including the space character) and non-printable (control) characters are defined in paragraph 40 of WIPO Standard ST.26 (incorporated by reference, *see* 37 CFR 1.839); and (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.

Section 1.834(b) is added to require that the "Sequence Listing XML" must be in a single file containing the sequence information and be submitted either: (1) electronically via the USPTO patent electronic filing system, where the file size must not exceed 100 MB and file compression is not permitted; or (2) on read-only optical disc(s) in compliance with § 1.52(e), where (i) a file that is not compressed must be contained on a single read-only optical disc, (ii) the file may be compressed using

WinZip®, 7-Zip, or Unix®/Linux® Zip, (iii) a compressed file must not be self-extracting, and (iv) a compressed XML file that does not fit on a single read-only optical disc may be split into multiple file parts in accordance with the target read-only optical disc size and labeled in compliance with § 1.52(e)(5)(vi).

Section 1.834(c)(1) is added to require that when a "Sequence Listing XML" required by § 1.831(a) is submitted in XML file format via the USPTO patent electronic filing system or on a read-only optical disc (in compliance with § 1.52(e)), the specification must contain a statement in a separate paragraph (*see* § 1.77(b)(5)) that incorporates by reference the material in the XML file identifying: (1) the name of the file, (2) the date of creation, and (3) the size of the file in bytes, so long as § 1.834(c)(2) does not apply. This provision was added in the final rule to expressly require an incorporation by reference statement in the specification to the "Sequence Listing XML," which was only implicitly required by § 1.835(c).

Section 1.834(c)(2) is added to indicate that if the "Sequence Listing XML" required by § 1.831(a) is submitted in XML file format via the USPTO patent electronic filing system or on a read-only optical disc (in compliance with § 1.52(e)) for an international application during the international stage, then an incorporation by reference statement of the material in the XML file is not required. This provision was added in the final rule to specifically exempt the requirement for an incorporation by reference statement in the specification to the "Sequence Listing XML" (as in § 1.834(c)(1)) for a national stage application when the "Sequence Listing XML" constituted part of the international application during the international stage.

Section 1.835: Section 1.835 is added to provide the requirements for submission of an amendment to add or replace a "Sequence Listing XML" for applications filed on or after July 1, 2022.

Section 1.835(a) is added to require that any amendment to a patent application adding an initial submission of a "Sequence Listing XML" as required by § 1.831(a) after the application filing date must include: (1) a "Sequence Listing XML" file submitted either (i) via the USPTO patent electronic filing system, or (ii) on a read-only optical disc in compliance with § 1.52(e); (2) a request to amend the specification to include an incorporation by reference statement of the material in the "Sequence Listing XML" file, identifying the name of the file, the date of creation, and the size of the file in bytes (*see* § 1.77(b)(5)(ii)), except when submitted to the United States International Preliminary Examining Authority for an international application; (3) a statement that indicates the basis for the amendment, with specific references to particular parts of the application as originally filed (specification, claims, drawings) for all sequence data in the "Sequence Listing XML"; and (4) a statement that the "Sequence Listing XML" includes no new matter.

Section 1.835(b) is added to require that any amendment adding to, deleting from, or replacing sequence information in a "Sequence Listing XML" submitted as required by § 1.831(a) must include: (1) a replacement "Sequence Listing XML" containing the entire "Sequence Listing XML," including any additions, deletions, or replacements of sequence information, and shall be submitted either (i) via the USPTO patent electronic filing system, or (ii) on a read-only optical disc, in compliance with § 1.52(e) labeled as "REPLACEMENT MM/DD/YYYY" (with the month, day, and year of creation indicated); (2) an instruction to amend the specification to include an incorporation by reference statement of the material in the replacement "Sequence Listing XML" file that identifies the name of the file, the date of creation, and the size of the file in bytes (*see* § 1.77(b)(5)(ii)), except when the replacement "Sequence Listing XML" is submitted to the United States International Preliminary Examining Authority for an international application; (3) a statement that identifies the location of all additions, deletions, or

replacements of sequence information relative to the replaced "Sequence Listing XML";

(4) a statement that indicates the support for the additions, deletions, or replacements of

the sequence information, with specific references to particular parts of the application as

originally filed (specification, claims, drawings) for all amended sequence data in the

replacement "Sequence Listing XML"; and (5) a statement that the replacement

"Sequence Listing XML" includes no new matter.

Section 1.835(c) is added to require that the specification of a complete

application with a "Sequence Listing XML" as required under § 1.831(a), present on the

application filing date but without an incorporation by reference of the material contained

in the "Sequence Listing XML" file, must be amended to contain a separate paragraph

incorporating by reference the material contained in the "Sequence Listing XML" file, in

accordance with § 1.77(b)(5)(ii), except for international applications.

Section 1.835(d)(1) is added to provide that, when any of the requirements of

§§ 1.831 through 1.834 are not satisfied in an application under 35 U.S.C. 111(a) or in a

national stage application under 35 U.S.C. 371, the applicant will be notified and given a

period of time in which to comply with such requirements to prevent the abandonment of

the application. This final rule indicates that, subject to § 1.835(d)(2), any amendment to

add or replace a "Sequence Listing XML" in response to a requirement under this

paragraph must be submitted in accordance with the requirements of § 1.835(a) through

(c).

Section 1.835(d)(2) is added to explicitly provide that compliance with § 1.835(a)

through (c) is not required for the submission of a "Sequence Listing XML" that is solely

an English translation of a previously submitted "Sequence Listing XML" that contains

non-English values for any language-dependent free text elements (as per § 1.833(b)(3)).

The required submission will be a translated "Sequence Listing XML" in compliance

with §§ 1.831 through 1.834. Updated values for attributes in the root element (§ 1.833(b)(2)(iii)) or elements of the general information part (§ 1.833(b)(2)(iv)) are not considered amendments for purposes of complying with § 1.835(a) through (c). Even though §§ 1.52(b)(1)(ii) and 1.495(c)(1)(i) require a translation for applications filed under 35 U.S.C. 111(a) and for those entering the national stage, respectively, this rule makes explicit that when a translated "Sequence Listing XML" is provided as a reply to a notice that the "Sequence Listing XML" contains non-English values for any language-dependent free text elements, and the translation does not include the deletion, addition, or replacement of sequence information, the translated "Sequence Listing XML" need not comply with the requirements for an amended "Sequence Listing XML" as set forth in § 1.835(a) through (c). The proposed requirement for a translation of the title into English was not adopted in the final rule because applicants in the international phase need only provide a title in the language of filing, which can be in a language other than English.

Section 1.835(e) is added to provide that, when any of the requirements of §§ 1.831 through 1.834 are not satisfied at the time of filing an international application under the PCT, where the application is to be searched by the United States International Searching Authority or examined by the United States International Preliminary Examining Authority, the applicant may be sent a notice calling for compliance with the requirements within a prescribed time period. Under PCT Rule 13*ter*, the applicant may provide, in response to such a requirement or otherwise, a sequence listing that is a "Sequence Listing XML" in accordance with § 1.831(a). The "Sequence Listing XML" must be accompanied by a statement that the information recorded does not go beyond the disclosure in the international application as filed. In response to such a requirement, the late furnishing fee set forth in § 1.445(a)(5) is also required. If the applicant fails to timely provide the required "Sequence Listing XML," the United States International Searching Authority shall search only to the extent that a meaningful search can be

performed without the "Sequence Listing XML," and the United States International Preliminary Examining Authority shall examine only to the extent that a meaningful examination can be performed without the "Sequence Listing XML."

Section 1.835(f) is added to provide that any appropriate amendments to the "Sequence Listing XML" in a patent (e.g., by reason of reissue, reexamination, or certificate of correction) must comply with the requirements of paragraph (b) of this section.

Section 1.839: Section 1.839 is added to provide the location of WIPO Standard ST.26 that is being incorporated by reference.

III. Comments and Responses and Changes from Proposed Rule

The USPTO published a proposed rule on July 6, 2021, at 86 FR 35432, soliciting public comments on the proposed amendments to 37 CFR part 1 being adopted in this final rule. The USPTO received no comments from the public on the proposed rule. Even though no comments were received, the proposed changes to §§ 1.495(c)(5), 1.833(b)(3) and 1.835(d)(2) to require a title in English in the "Sequence Listing XML" were not adopted in the final rule. The proposed requirement for a translation of the title into English was not adopted since applicants in the international phase need only provide a title in the language of filing, which can be in a language other than English. Additionally, even though § 1.831(a) states that "[d]isclosed nucleotide or amino acid sequences that do not meet the definition in paragraph (b) of this section must not be included in the "Sequence Listing XML,"" § 1.831(j) was added to make explicit the prohibition of including such sequences in the "Sequence Listing XML." Section 1.834(c)(1) was added to expressly require an incorporation by reference statement in the specification to the "Sequence Listing XML," which was only implicitly required by § 1.835(c). Lastly, § 1.834(c)(2) was added to specifically exempt the requirement for an

incorporation by reference statement in the specification to the "Sequence Listing XML" (as in § 1.834(c)(1)) for a national stage application when the "Sequence Listing XML" constituted part of the international application during the international stage.

IV.   Rulemaking Considerations

A. Administrative Procedure Act: The changes in this rulemaking involve rules of agency practice and procedure, and/or interpretive rules. See *Bachow Commc'ns Inc.* v. *FCC*, 237 F.3d 683, 690 (D.C. Cir. 2001) (changes to procedural rules are not subject to notice and comment review under the Administrative Procedure Act (APA)); *Inova Alexandria Hosp.* v. *Shalala*, 244 F.3d 342, 349 (4th Cir. 2001) (rules for handling appeals are procedural where they do not change the substantive standard for reviewing claims); *Nat'l Org. of Veterans' Advocates* v. *Sec'y of Veterans Affairs*, 260 F.3d 1365, 1375 (Fed. Cir. 2001) (Substantive rules "effect a change in existing law or policy or which affect individual rights and obligations," whereas interpretative rules "clarify or explain existing law or regulation and are exempt from notice and comment" review under the APA.).

Accordingly, prior notice and opportunity for public comment for the changes in this rulemaking were not required pursuant to 5 U.S.C. 553(b) or (c), or any other law. See *Cooper Techs. Co.* v. *Dudas*, 536 F.3d 1330, 1336-37 (Fed. Cir. 2008) (stating that 5 U.S.C. 553, and thus 35 U.S.C. 2(b)(2)(B), do not require notice and comment rulemaking for "interpretative rules, general statements of policy, or rules of agency organization, procedure, or practice" (quoting 5 U.S.C. 553(b)(A))). However, the USPTO chose to seek public comment before implementing the rule to benefit from the public's input.

B. Regulatory Flexibility Act: For the reasons set forth in this notice, the Senior Counsel for Regulatory and Legislative Affairs of the USPTO has certified to the Chief Counsel

for Advocacy of the Small Business Administration that this rule will not have a significant economic impact on a substantial number of small entities. See 5 U.S.C. 605(b).

The USPTO amends the rules of practice to require the submission of biological sequence data in XML where the rules of practice incorporate by reference WIPO Standard ST.26, "Recommended Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings Using XML (eXtensible Markup Language)," including Annexes I-VII, version 1.5, approved November 5, 2021, as disclosed in the WIPO Handbook on Industrial Property Information and Documentation.

This rulemaking makes more technical data associated with biotechnology inventions available to the public because the new rules of practice based on WIPO Standard ST.26 provide for enhanced biological sequence data related to disclosures of nucleotide and/or amino acid sequences in patent applications. WIPO Standard ST.26 provides clear rules as to what must be included in a sequence listing and how sequences must be represented (e.g., standardization of the representation of modified nucleic acids and amino acids as well as variants derived from primary sequences). WIPO Standard ST.26 contains a guidance document that demonstrates the requirement for inclusion and representation of biological sequence data. As a result, patent applicants will have a clearer understanding as to the requirements and presentation of biological sequence data in a compliant sequence listing under WIPO Standard ST.26. Additionally, since WIPO Standard ST.26 only allows XML format, applicants will not be burdened with or confused by the requirements of filing a sequence listing in paper or PDF format, and IPOs will not be burdened with processing paper sequence listings and performing necessary checks on the contents of the paper documents. The changes in this rulemaking are largely procedural in nature, and do not impose any additional requirements or fees

on applicants. For the foregoing reasons, the changes in this rule will not have a significant economic impact on a substantial number of small entities.

C. Executive Order 12866 (Regulatory Planning and Review): This rulemaking has been determined to be not significant for purposes of Executive Order 12866 (Sept. 30, 1993).

D. Executive Order 13563 (Improving Regulation and Regulatory Review): The USPTO has complied with Executive Order 13563 (Jan. 18, 2011). Specifically, to the extent feasible and applicable, the USPTO has: (1) reasonably determined that the benefits of the rule justify its costs; (2) tailored the rule to impose the least burden on society consistent with obtaining the agency's regulatory objectives; (3) selected a regulatory approach that maximizes net benefits; (4) specified performance objectives; (5) identified and assessed available alternatives; (6) involved the public in an open exchange of information and perspectives among experts in relevant disciplines, affected stakeholders in the private sector, and the public as a whole, and provided online access to the rulemaking docket; (7) attempted to promote coordination, simplification, and harmonization across government agencies and identified goals designed to promote innovation; (8) considered approaches that reduce burdens while maintaining flexibility and freedom of choice for the public; and (9) ensured the objectivity of scientific and technological information and processes.

E. Executive Order 13132 (Federalism): This rulemaking does not contain policies with federalism implications sufficient to warrant preparation of a Federalism Assessment under Executive Order 13132 (Aug. 4, 1999).

F. Executive Order 13175 (Tribal Consultation): This rulemaking will not: (1) have substantial direct effects on one or more Indian tribes; (2) impose substantial direct compliance costs on Indian tribal governments; or (3) preempt tribal law. Therefore, a

tribal summary impact statement is not required under Executive Order 13175 (Nov. 6, 2000).

G. Executive Order 13211 (Energy Effects): This rulemaking is not a significant energy action under Executive Order 13211 because this rulemaking is not likely to have a significant adverse effect on the supply, distribution, or use of energy. Therefore, a Statement of Energy Effects is not required under Executive Order 13211 (May 18, 2001).

H. Executive Order 12988 (Civil Justice Reform): This rulemaking meets applicable standards to minimize litigation, eliminate ambiguity, and reduce burden as set forth in sections 3(a) and 3(b)(2) of Executive Order 12988 (Feb. 5, 1996).

I. Executive Order 13045 (Protection of Children): This rulemaking does not concern an environmental risk to health or safety that may disproportionately affect children under Executive Order 13045 (Apr. 21, 1997).

J. Executive Order 12630 (Taking of Private Property): This rulemaking will not affect a taking of private property or otherwise have taking implications under Executive Order 12630 (Mar. 15, 1988).

K. Congressional Review Act: Under the Congressional Review Act provisions of the Small Business Regulatory Enforcement Fairness Act of 1996 (5 U.S.C. 801 et seq.), the USPTO will submit a report containing the final rule and other required information to the United States Senate, the United States House of Representatives, and the Comptroller General of the Government Accountability Office. The changes in this rulemaking are not expected to result in an annual effect on the economy of $100 million or more, a major increase in costs or prices, or significant adverse effects on competition, employment, investment, productivity, innovation, or the ability of United States-based enterprises to compete with foreign-based enterprises in domestic and export markets.

Therefore, this rulemaking is not expected to result in a "major rule" as defined in 5 U.S.C. 804(2).

L. Unfunded Mandates Reform Act of 1995: The changes set forth in this rulemaking do not involve a Federal intergovernmental mandate that will result in the expenditure by State, local, and tribal governments, in the aggregate, of $100 million (as adjusted) or more in any one year, or a Federal private sector mandate that will result in the expenditure by the private sector of $100 million (as adjusted) or more in any one year, and will not significantly or uniquely affect small governments. Therefore, no actions are necessary under the provisions of the Unfunded Mandates Reform Act of 1995. See 2 U.S.C. 1501 et seq.

M. National Environmental Policy Act of 1969: This rulemaking will not have any effect on the quality of the environment and is thus categorically excluded from review under the National Environmental Policy Act of 1969. See 42 U.S.C. 4321 et seq.

N. National Technology Transfer and Advancement Act of 1995: The requirements of section 12(d) of the National Technology Transfer and Advancement Act of 1995 (15 U.S.C. 272 note) are not applicable because this rulemaking does not contain provisions that involve the use of technical standards.

O. Paperwork Reduction Act of 1995: The Paperwork Reduction Act of 1995 (44 U.S.C. 3501-3549) requires that the USPTO consider the impact of paperwork and other information collection burdens imposed on the public. In accordance with section 3507(d) of the Paperwork Reduction Act of 1995, the majority of the paperwork and other information collection burdens discussed in this rule have already been approved under the following Office of Management and Budget (OMB) Control Numbers: 0651-0024 (Sequence Listing), 0651-0031 (Patent Processing), 0651-0032 (Initial Patent Applications), and 0651-0064 (Patent Reexaminations and Supplemental Examinations).

Modifications to 0651-0024 because of this rulemaking will be submitted to OMB for approval. Modifications include the removal of the Sequence Listing in Application (paper), which will result in an estimated reduction in the burden associated with this information collection by 5,000 responses and 30,000 burden hours. These burden estimates are based on the current OMB approved burdens (response volumes) associated with this information collection, which may be different from any forecasts mentioned in other parts of this rule.

The changes discussed in this rule do not affect the information collection requirements or burdens associated with 0651-0031, 0651-0032, and 0651-0064 listed above; therefore, the USPTO does not plan to take any additional actions on these information collections as a result of this rulemaking. Notwithstanding any other provision of law, no person is required to respond to, nor shall a person be subject to a penalty for failure to comply with, a collection of information subject to the requirements of the Paperwork Reduction Act unless that collection of information has a currently valid OMB control number.

P. E-Government Act Compliance: The USPTO is committed to compliance with the E-Government Act to promote the use of the internet and other information technologies, to provide increased opportunities for citizen access to Government information and services, and for other purposes.

List of Subjects in 37 CFR Part 1

Administrative practice and procedure, Biologics, Courts, Freedom of information, Incorporation by reference, Inventions and patents, Reporting and recordkeeping requirements, Small businesses.

For the reasons stated in the preamble and under the authority contained in 35 U.S.C. 2, as amended, the USPTO amends 37 CFR part 1 as follows:

PART 1—RULES OF PRACTICE IN PATENT CASES

1. The authority citation for part 1 continues to read as follows:

    Authority: 35 U.S.C. 2(b)(2), unless otherwise noted.

2. Section 1.52 is amended by:

a. Revising paragraphs (e)(1)(ii) and (e)(3)(ii) and (iii);

b. Adding paragraph (e)(3)(iv); and

c. Revising paragraphs (e)(7), (f)(1) introductory text, (f)(1)(i), (f)(2) introductory text, (f)(2)(i), and (f)(3).

The revisions and addition read as follows:

§ 1.52 Language, paper, writing, margins, read-only optical disc specifications.

* * * * *

(e) * * *

(1) * * *

(ii) A "Sequence Listing" (submitted under § 1.821(c) in compliance with §§ 1.822 through 1.824) or a "Sequence Listing XML" (submitted under § 1.831(a) in compliance with §§ 1.832 through 1.834); or

* * * * *

(3) * * *

(ii) Operating system compatibility: MS-DOS®, MS-Windows®, MacOS®, or Unix®/Linux®;

(iii) The contents of each read-only optical disc must be in American Standard Code for Information Interchange (ASCII) plain text and if compressed, must be

compressed in accordance with § 1.58 for "Large Tables," with § 1.96 for a "Computer Program Listing Appendix," or § 1.824 for a "Sequence Listing" or Computer Readable Form (CRF) of the "Sequence Listing," as applicable; and

(iv) The contents of each read-only optical disc for a "Sequence Listing XML" must be in eXtensible Markup Language (XML) file format, and if compressed, must be compressed in accordance with § 1.834.

* * * * *

(7) Any amendment to the information on a read-only optical disc must be by way of a replacement read-only optical disc, in compliance with § 1.58(g) for "Large Tables," § 1.96(c)(5) for a "Computer Program Listing Appendix," § 1.825(b) for a "Sequence Listing" or CRF of a "Sequence Listing," and § 1.835(b) for a "Sequence Listing XML."

* * * * *

(f) * * *

(1) *Submission on read-only optical discs*. The application size fee required by § 1.16(s) or § 1.492(j), for an application component submitted in part on a read-only optical disc in compliance with paragraph (e) of this section, shall be determined such that each three kilobytes of content submitted on a read-only optical disc shall be counted as a sheet of paper. Excluded from this determination is any ASCII plain text file or any XML file (as applicable) submitted on a read-only optical disc under paragraph (e) of this section containing:

(i) Any "Sequence Listing" or CRF of a "Sequence Listing" in compliance with § 1.821(c) or (e), or any "Sequence Listing XML" in compliance with § 1.831(a); or

* * * * *

(2) *Submission via the USPTO patent electronic filing system*. The application size fee required by § 1.16(s) or § 1.492(j), for an application submitted in whole or in part via the USPTO patent electronic filing system, shall be determined such that the paper size equivalent will be considered to be 75% of the number of sheets of paper present in the specification and drawings for the application when entered into the Office records after being rendered by the USPTO patent electronic filing system. Excluded from this determination is any ASCII plain text file or any XML file (as applicable) submitted via the USPTO patent electronic filing system containing:

(i) Any "Sequence Listing" or CRF of a "Sequence Listing" in compliance with § 1.821(c)(1) or (e), or any "Sequence Listing XML" in compliance with § 1.831(a); or

* * * * *

(3) *Oversized submission.* Any submission of a "Sequence Listing" in electronic form or a "Sequence Listing XML" of 300 MB-800 MB filed in an application under 35 U.S.C. 111 or 371 will be subject to the fee set forth in § 1.21(o)(1). Any submission of a "Sequence Listing" in electronic form or a "Sequence Listing XML" that exceeds 800 MB filed in an application under 35 U.S.C. 111 or 371 will be subject to the fee set forth in § 1.21(o)(2).

3. Section 1.53 is amended by revising paragraph (c)(4) to read as follows:

§ 1.53 Application number, filing date, and completion of application.

* * * * *

(c) * * *

(4) A provisional application is not entitled to the right of priority under 35 U.S.C. 119, 365(a), or 386(a) or § 1.55, or to the benefit of an earlier filing date under

35 U.S.C. 120, 121, 365(c), or 386(c) or § 1.78 of any other application. No claim for priority under 35 U.S.C. 119(e) or § 1.78(a) may be made in a design application based on a provisional application. A provisional application disclosing nucleotide and/or amino acid sequences is not required to include a separate sequence listing; however, if submitted in a provisional application filed on or after July 1, 2022, any submission of nucleotide and/or amino acid sequence data must be by way of a "Sequence Listing XML" in compliance with §§ 1.831 through 1.834.

* * * * *

4. Section 1.77 is amended by revising paragraph (b)(5) to read as follows:

§ 1.77 Arrangement of application elements.

* * * * *

(b) * * *

(5) An incorporation by reference statement regarding the material in:

(i) One or more ASCII plain text files, submitted via the USPTO patent electronic filing system or on one or more read-only optical discs (*see* § 1.52(e)(8)), identifying the names of each file, the date of creation of each file, and the size of each file in bytes, for the following document types:

(A) A "Computer Program Listing Appendix" (*see* § 1.96(c));

(B) A "Sequence Listing" (*see* § 1.821(c)); or

(C) "Large Tables" (*see* § 1.58(c)).

(ii) An XML file for a "Sequence Listing XML" (*see* § 1.831(a)), submitted via the USPTO patent electronic filing system or on one or more read-only optical discs (*see*

§ 1.52(e)(8)), identifying the names of each file, the date of creation of each file, and the size of each file in bytes.

* * * * *

5. Section 1.121 is amended by revising paragraphs (b) introductory text and (b)(6) to read as follows:

§ 1.121 Manner of making amendments in applications.

* * * * *

(b) *Specification*. Amendments to the specification, other than the claims, "Large Tables" (§ 1.58(c)), a "Computer Program Listing Appendix" (§ 1.96(c)(5) and (7)), a "Sequence Listing" (§ 1.825), or a "Sequence Listing XML" (§ 1.835), must be made by adding, deleting, or replacing a paragraph; by replacing a section; or by providing a substitute specification, in the manner specified in this section.

* * * * *

(6) *Amendments to "Large Tables," a "Computer Program Listing Appendix," a "Sequence Listing," or a "Sequence Listing XML."* Changes to "Large Tables," a "Computer Program Listing Appendix," a "Sequence Listing," or a "Sequence Listing XML" must be made in accordance with § 1.58(g) for "Large Tables," § 1.96(c)(5) for a "Computer Program Listing Appendix," § 1.825 for a "Sequence Listing," or § 1.835 for a "Sequence Listing XML."

* * * * *

6. Section 1.173 is amended by revising paragraphs (b)(1) and (d) introductory text to read as follows:

§ 1.173 Reissue specification, drawings, and amendments.

* * * * *

(b) * * *

(1) *Specification other than the claims, "Large Tables" (§ 1.58(c)), a "Computer Program Listing Appendix" (§ 1.96(c)), a "Sequence Listing" (§ 1.821(c)), or a "Sequence Listing XML" (§ 1.831(a)).* (i) Changes to the specification, other than to the claims, "Large Tables" (§ 1.58(c)), a "Computer Program Listing Appendix" (§ 1.96(c)), a "Sequence Listing" (§ 1.821(c)), or a "Sequence Listing XML" (§ 1.831(a)), must be made by submission of the entire text of an added or rewritten paragraph, including markings pursuant to paragraph (d) of this section, except that an entire paragraph may be deleted by a statement deleting the paragraph, without presentation of the text of the paragraph. The precise point in the specification where any added or rewritten paragraph is located must be identified.

(ii) Changes to "Large Tables," a "Computer Program Listing Appendix," a "Sequence Listing," or a "Sequence Listing XML" must be made in accordance with § 1.58(g) for "Large Tables," § 1.96(c)(5) for a "Computer Program Listing Appendix," § 1.825 for a "Sequence Listing," and § 1.835 for a "Sequence Listing XML."

* * * * *

(d) *Changes shown by markings*. Any changes relative to the patent being reissued that are made to the specification, including the claims but excluding "Large Tables" (§ 1.58(c)), a "Computer Program Listing Appendix" (§ 1.96(c)), a "Sequence Listing" (§ 1.821(c)), and a "Sequence Listing XML" (§ 1.831(a)) upon filing or by an amendment paper in the reissue application, must include the following markings:

* * * * *

7. Section 1.211 is amended by revising paragraph (c) to read as follows:

§ 1.211 Publication of applications.

* * * * *

(c) An application filed under 35 U.S.C. 111(a) will not be published until it includes

the basic filing fee (§ 1.16(a) or (c)) and any English translation required by

§ 1.52(d). The Office may delay publishing any application until it includes any

application size fee required by the Office under § 1.16(s) or § 1.492(j), a

specification having papers in compliance with § 1.52 and an abstract (§ 1.72(b)),

drawings in compliance with § 1.84, a "Sequence Listing" in compliance with

§§ 1.821 through 1.825 (if applicable) for an application filed before July 1, 2022, a

"Sequence Listing XML" in compliance with §§ 1.831 through 1.835 (if applicable)

for an application filed on or after July 1, 2022, and the inventor's oath or declaration

or application data sheet containing the information specified in § 1.63(b).

* * * * *

8. Section 1.495 is amended by revising paragraph (c)(5) to read as follows:

§ 1.495 Entering the national stage in the United States of America.

* * * * *

(c) * * *

(5) For international applications having an international filing date before July 1,

2022, a sequence listing need not be translated if the sequence listing complies with

PCT Rule 12.1(d) and the description complies with PCT Rule 5.2(b). For

international applications having an international filing date on or after July 1, 2022,

for purposes of paragraph (c)(1)(i) of this section, an English translation is required

for any sequence listing in XML format ("Sequence Listing XML") containing non-

English language values for any language-dependent free text qualifiers in accordance with §§ 1.831 through 1.834.

* * * * *

9. Section 1.530 is amended by revising paragraph (d)(1) to read as follows:

§ 1.530 Statement by patent owner in *ex parte* reexamination; amendment by patent owner in *ex parte* or *inter partes* reexamination; inventorship change in *ex parte* or *inter partes* reexamination.

* * * * *

(d) * * *

(1) *Specification other than the claims, "Large Tables" (§ 1.58(c)), a "Computer Program Listing Appendix" (§ 1.96(c)), a "Sequence Listing" (§ 1.821(c)), or a "Sequence Listing XML (§ 1.831(a)).* (i) Changes to the specification, other than to the claims, "Large Tables" (§ 1.58(c)), a "Computer Program Listing Appendix" (§ 1.96(c)), a "Sequence Listing" (§ 1.821(c)), or a "Sequence Listing XML" (§ 1.831(a)), must be made by submission of the entire text of an added or rewritten paragraph, including markings pursuant to paragraph (f) of this section, except that an entire paragraph may be deleted by a statement deleting the paragraph, without presentation of the text of the paragraph. The precise point in the specification where any added or rewritten paragraph is located must be identified.

(ii) Changes to "Large Tables," a "Computer Program Listing Appendix," a "Sequence Listing," or a "Sequence Listing XML" must be made in accordance with § 1.58(g) for "Large Tables," § 1.96(c)(5) for a "Computer Program Listing Appendix," § 1.825 for a "Sequence Listing," or § 1.835 for a "Sequence Listing XML."

* * * * *

10. Section 1.704 is amended by revising paragraph (f) to read as follows:

§ 1.704 Reduction of period of adjustment of patent term.

* * * * *

(f) An application filed under 35 U.S.C. 111(a) is in condition for examination when it includes a specification, including at least one claim and an abstract (§ 1.72(b)), and has papers in compliance with § 1.52, drawings (if any) in compliance with § 1.84, any English translation required by § 1.52(d) or § 1.57(a), a "Sequence Listing" in compliance with §§ 1.821 through 1.825 (if applicable), a "Sequence Listing XML" in compliance with §§ 1.831 through 1.835 (if applicable), an inventor's oath or declaration or an application data sheet containing the information specified in § 1.63(b), the basic filing fee (§ 1.16(a) or (c)), the search fee (§ 1.16(k) or (m)), the examination fee (§ 1.16(o) or (q)), any certified copy of the previously filed application required by § 1.57(a), and any application size fee required by the Office under § 1.16(s). An international application is in condition for examination when it has entered the national stage as defined in § 1.491(b), and includes a specification, including at least one claim and an abstract (§ 1.72(b)), and has papers in compliance with § 1.52, drawings (if any) in compliance with § 1.84, a "Sequence Listing" in compliance with §§ 1.821 through 1.825 (if applicable), a "Sequence Listing XML" in compliance with §§ 1.831 through 1.835 (if applicable), an inventor's oath or declaration or an application data sheet containing the information specified in § 1.63(b), the search fee (§ 1.492(b)), the examination fee (§ 1.492(c)), and any application size fee required by the Office under § 1.492(j). An application shall be considered as having papers in compliance with § 1.52, drawings (if any) in compliance with § 1.84, and a "Sequence Listing" in compliance with §§ 1.821

through 1.825 (if applicable), or a "Sequence Listing XML" in compliance with §§ 1.831 through 1.835 (if applicable), for purposes of this paragraph (f) on the filing date of the latest reply (if any) correcting the papers, drawings, "Sequence Listing," or "Sequence Listing XML" that is prior to the date of mailing of either an action under 35 U.S.C. 132 or a notice of allowance under 35 U.S.C. 151, whichever occurs first.

11. Sections 1.831 through 1.835 and 1.839 are added to read as follows:

Sec.

* * * * *

1.831 Requirements for patent applications filed on or after July 1, 2022, having nucleotide and/or amino acid sequence disclosures.

1.832 Representation of nucleotide and/or amino acid sequence data in the "Sequence Listing XML" part of a patent application filed on or after July 1, 2022.

1.833 Requirements for a "Sequence Listing XML" for nucleotide and/or amino acid sequences as part of a patent application filed on or after July 1, 2022.

1.834 Form and format for nucleotide and/or amino acid sequence submissions as the "Sequence Listing XML" in patent applications filed on or after July 1, 2022.

1.835 Amendment to add or replace a "Sequence Listing XML" in patent applications filed on or after July 1, 2022.

1.839 Incorporation by reference.

* * * * *

§ 1.831 Requirements for patent applications filed on or after July 1, 2022, having nucleotide and/or amino acid sequence disclosures.

(a) Patent applications disclosing nucleotide and/or amino acid sequences by enumeration of their residues, as defined in paragraph (b) of this section, must contain, as a separate part of the disclosure, a computer readable Sequence Listing in XML format (a "Sequence Listing XML"). Disclosed nucleotide or amino acid sequences that do not meet the definition in paragraph (b) of this section must not be included in the "Sequence Listing XML." The "Sequence Listing XML" contains the information of the nucleotide and/or amino acid sequences disclosed in the patent application using the symbols and format in accordance with the requirements of §§ 1.832 through 1.834.

(b) Nucleotide and/or amino acid sequences, as used in this section and §§ 1.832 through 1.835, encompass:

(1) An unbranched sequence or linear region of a branched sequence containing 4 or more specifically defined amino acids, wherein the amino acids form a single peptide backbone; or

(2) An unbranched sequence or linear region of a branched sequence of 10 or more specifically defined nucleotides, wherein adjacent nucleotides are joined by:

(i) A 3' to 5' (or 5' to 3') phosphodiester linkage; or

(ii) Any chemical bond that results in an arrangement of adjacent nucleobases that mimics the arrangement of nucleobases in naturally occurring nucleic acids (i.e., nucleotide analogs).

(c) Where the description or claims of a patent application discuss a sequence that is set forth in the "Sequence Listing XML" in accordance with paragraph (a) of this section, reference must be made to the sequence by use of the sequence identifier, preceded by "SEQ ID NO:" or the like in the text of the description or claims, even if the sequence is also embedded in the text of the description or claims of the patent

application. Where a sequence is presented in a drawing, reference must be made to the sequence by use of the sequence identifier (§ 1.832(a)), either in the drawing or in the Brief Description of the Drawings, where the correlation between multiple sequences in the drawing and their sequence identifiers (§ 1.832(a)) in the Brief Description is clear.

(d) "Enumeration of its residues" means disclosure of a nucleotide or amino acid sequence in a patent application by listing, in order, each residue of the sequence, where the residues are represented in the manner as defined in paragraph 3(c)(i) or (ii) of WIPO Standard ST.26 (incorporated by reference, *see* § 1.839).

(e) "Specifically defined" means any amino acid or nucleotide as defined in paragraph 3(k) of WIPO Standard ST.26.

(f) "Amino acid" includes any D- or L-amino acid or modified amino acid as defined in paragraph 3(a) of WIPO Standard ST.26.

(g) "Modified amino acid" includes any amino acid as described in paragraph 3(e) of WIPO Standard ST.26.

(h) "Nucleotide" includes any nucleotide, nucleotide analog, or modified nucleotide as defined in paragraphs 3(f) and 3(g) of WIPO Standard ST.26.

(i) "Modified nucleotide" includes any nucleotide as described in paragraph 3(f) of WIPO Standard ST.26.

(j) A "Sequence listing XML" must not include any sequences having fewer than 10 specifically defined nucleotides, or fewer than 4 specifically defined amino acids.

§ 1.832 Representation of nucleotide and/or amino acid sequence data in the "Sequence Listing XML" part of a patent application filed on or after July 1, 2022.

(a) Each disclosed nucleotide or amino acid sequence that meets the requirements of § 1.831(b) must appear separately in the "Sequence Listing XML." Each sequence set forth in the "Sequence Listing XML" must be assigned a separate sequence identifier. The sequence identifiers must begin with 1 and increase sequentially by integers as defined in paragraph 10 of WIPO Standard ST.26 (incorporated by reference, *see* § 1.839).

(b) The representation and symbols for nucleotide sequence data shall conform to the requirements of paragraphs (b)(1) through (4) of this section.

(1) A nucleotide sequence must be represented in the manner described in paragraphs 11-12 of WIPO Standard ST.26.

(2) All nucleotides, including nucleotide analogs, modified nucleotides, and "unknown" nucleotides, within a nucleotide sequence must be represented using the symbols set forth in paragraphs 13-16, 19, and 21 of WIPO Standard ST.26.

(3) Modified nucleotides within a nucleotide sequence must be described in the manner discussed in paragraphs 17, 18, and 19 of WIPO Standard ST.26.

(4) A region containing a known number of contiguous "a," "c," "g," "t," or "n" residues for which the same description applies may be jointly described in the manner described in paragraph 22 of WIPO Standard ST.26.

(c) The representation and symbols for amino acid sequence data shall conform to the requirements of paragraphs (c)(1) through (4) of this section.

(1) The amino acids in an amino acid sequence must be represented in the manner described in paragraphs 24 and 25 of WIPO Standard ST.26.

(2) All amino acids, including modified amino acids and "unknown" amino acids, within an amino acid sequence must be represented using the symbols set forth in paragraphs 26-29 and 32 of WIPO Standard ST.26.

(3) Modified amino acids within an amino acid sequence must be described in the manner discussed in paragraphs 29 and 30 of WIPO Standard ST.26.

(4) A region containing a known number of contiguous "X" residues for which the same description applies may be jointly described in the manner described in paragraph 34 of WIPO Standard ST.26.

(d) A nucleotide and/or amino acid sequence that is constructed as a single continuous sequence derived from one or more non-contiguous segments of a larger sequence or of segments from different sequences must be listed in the "Sequence Listing XML" in the manner described in paragraph 35 of WIPO Standard ST.26.

(e) A nucleotide and/or amino acid sequence that contains regions of specifically defined residues separated by one or more regions of contiguous "n" or "X" residues, wherein the exact number of "n" or "X" residues in each region is disclosed, must be listed in the "Sequence Listing XML" in the manner described in paragraph 36 of WIPO Standard ST.26.

(f) A nucleotide and/or amino acid sequence that contains regions of specifically defined residues separated by one or more gaps of an unknown or undisclosed number of residues must be listed in the "Sequence Listing XML" in the manner described in paragraph 37 of WIPO Standard ST.26.

§ 1.833 Requirements for a "Sequence Listing XML" for nucleotide and/or amino acid sequences as part of a patent application filed on or after July 1, 2022.

(a) The "Sequence Listing XML" as required by § 1.831(a) must be presented as a single file in XML 1.0 encoded using Unicode UTF-8, where the character set complies with paragraphs 40 and 41 and Annex IV of WIPO Standard ST.26 (incorporated by reference, *see* § 1.839).

(b) The "Sequence Listing XML" presented in accordance with paragraph (a) of this section must further:

(1) Be valid according to the Document Type Definition (DTD) as presented in WIPO Standard ST.26, Annex II.

(2) Comply with the requirements of WIPO Standard ST.26 to include:

(i) An XML declaration as defined in paragraph 39(a) of WIPO Standard ST.26;

(ii) A document type (DOCTYPE) declaration as defined in paragraph 39(b) of WIPO Standard ST.26;

(iii) A root element as defined in paragraph 43 of WIPO Standard ST.26;

(iv) A general information part that complies with the requirements of paragraphs 45, 47, and 48, as applicable, of WIPO Standard ST.26; and

(v) A sequence data part that complies with the requirements of paragraphs 50-55, 57, 58, 60-69, 71-78, 80-87, 89-98, and 100, as applicable, of WIPO Standard ST.26 representing the nucleotide and/or amino acid sequences according to § 1.832.

(3) Include an INSDQualifier_value element with a value in English for any language-dependent free text qualifier as defined by paragraphs 76 and 85-87 of WIPO Standard ST.26, and as required by § 1.52(b)(1)(ii).

§ 1.834 Form and format for nucleotide and/or amino acid sequence submissions as the "Sequence Listing XML" in patent applications filed on or after July 1, 2022.

(a) A "Sequence Listing XML" encoded using Unicode UTF-8, created by any means (e.g., text editors, nucleotide/amino acid sequence editors, or other custom computer programs) in accordance with §§ 1.831 through 1.833, must:

(1) Have the following compatibilities:

(i) Computer compatibility: PC or Mac®; and

(ii) Operating system compatibility: MS-DOS®, MS-Windows®, Mac OS®, or Unix®/Linux®.

(2) Be in XML format, where all permitted printable characters (including the space character) and non-printable (control) characters are defined in paragraph 40 of WIPO Standard ST.26 (incorporated by reference, *see* § 1.839).

(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.

(b) The "Sequence Listing XML" must be in a single file containing the sequence information and be submitted either:

(1) Electronically via the USPTO patent electronic filing system, where the file size must not exceed 100 MB, and file compression is not permitted; or

(2) On read-only optical disc(s) in compliance with § 1.52(e), where:

(i) A file that is not compressed must be contained on a single read-only optical disc;

(ii) The file may be compressed using WinZip®, 7-Zip, or Unix®/Linux® Zip;

(iii) A compressed file must not be self-extracting; or

(iv) A compressed XML file that does not fit on a single read-only optical disc may be split into multiple file parts, in accordance with the target read-only optical disc size, and labeled in compliance with § 1.52(e)(5)(vi);

(c)(1) Unless paragraph (c)(2) of this section applies, when the "Sequence Listing XML" required by § 1.831(a) is submitted in XML file format via the USPTO patent

electronic filing system or on a read-only optical disc (in compliance with § 1.52(e)), then the specification must contain a statement in a separate paragraph (*see* § 1.77(b)(5)) that incorporates by reference the material in the XML file identifying:

(i) The name of the file;

(ii) The date of creation; and

(iii) The size of the file in bytes; or

(2) If the "Sequence Listing XML" required by § 1.831(a) is submitted in XML file format via the USPTO patent electronic filing system or on a read-only optical disc (in compliance with § 1.52(e)) for an international application during the international stage, then an incorporation by reference statement of the material in the XML file is not required.

§ 1.835 Amendment to add or replace a "Sequence Listing XML" in patent applications filed on or after July 1, 2022.

(a) Any amendment to a patent application adding an initial submission of a "Sequence Listing XML" as required by § 1.831(a) after the application filing date must include:

(1) A "Sequence Listing XML" in accordance with §§ 1.831 through 1.834, submitted as an XML file:

(i) Via the USPTO patent electronic filing system; or

(ii) On a read-only optical disc, in compliance with § 1.52(e);

(2) A request to amend the specification to include an incorporation by reference statement of the material in the "Sequence Listing XML" file, identifying the name of the file, the date of creation, and the size of the file in bytes (*see* § 1.77(b)(5)(ii)),

except when submitted to the United States International Preliminary Examining Authority for an international application;

(3) A statement that indicates the basis for the amendment, with specific references to particular parts of the application as originally filed (specification, claims, drawings) for all sequence data in the "Sequence Listing XML"; and

(4) A statement that the "Sequence Listing XML" includes no new matter.

(b) Any amendment adding to, deleting from, or replacing sequence information in a "Sequence Listing XML" submitted as required by § 1.831(a) must include:

(1) A replacement "Sequence Listing XML" in accordance with the requirements of §§ 1.831 through 1.834 containing the entire "Sequence Listing XML," including any additions, deletions, or replacements of sequence information, which shall be submitted:

(i) Via the USPTO patent electronic filing system; or

(ii) On a read-only optical disc, in compliance with § 1.52(e), labeled as "REPLACEMENT MM/DD/YYYY" (with the month, day, and year of creation indicated);

(2) A request to amend the specification to include an incorporation by reference statement of the material in the replacement "Sequence Listing XML" file that identifies the name of the file, the date of creation, and the size of the file in bytes (*see* § 1.77(b)(5)(ii)), except when the replacement "Sequence Listing XML" is submitted to the United States International Preliminary Examining Authority for an international application;

(3) A statement that identifies the location of all additions, deletions, or replacements of sequence information relative to the replaced "Sequence Listing XML";

(4) A statement that indicates the support for the additions, deletions, or replacements of the sequence information, with specific references to particular parts of the application as originally filed (specification, claims, drawings) for all amended sequence data in the replacement "Sequence Listing XML"; and

(5) A statement that the replacement "Sequence Listing XML" includes no new matter.

(c) The specification of a complete application, filed on the application filing date, with a "Sequence Listing XML" as required under § 1.831(a), without an incorporation by reference of the material contained in the "Sequence Listing XML" file, must be amended to include a separate paragraph incorporating by reference the material contained in the "Sequence Listing XML" file, in accordance with § 1.77(b)(5)(ii), except for international applications.

(d)(1) If any of the requirements of §§ 1.831 through 1.834 are not satisfied in an application under 35 U.S.C. 111(a) or in a national stage application under 35 U.S.C. 371, the applicant will be notified and given a period of time within which to comply with such requirements in order to prevent abandonment of the application. Subject to paragraph (d)(2) of this section, any amendment to add or replace a "Sequence Listing XML" or add an incorporation by reference of the material contained in the "Sequence Listing XML" in response to a requirement under this paragraph (d)(1) must be submitted in accordance with the requirements of paragraphs (a) through (c) of this section.

(2) Compliance with paragraphs (a) through (c) of this section is not required for submission of a "Sequence Listing XML" that is solely an English translation of a previously submitted "Sequence Listing XML" that contains non-English values for any language-dependent free text elements (as per § 1.833(b)(3)). The required

submission will be a translated "Sequence Listing XML" in compliance with §§ 1.831 through 1.834. Updated values for attributes in the root element (§ 1.833(b)(2)(iii)) or elements of the general information part (§ 1.833(b)(2)(iv)) are not considered amendments for purposes of complying with paragraphs (a) through (c) of this section.

(e) If any of the requirements of §§ 1.831 through 1.834 are not satisfied at the time of filing an international application under the PCT, where the application is to be searched by the United States International Searching Authority or examined by the United States International Preliminary Examining Authority, the applicant may be sent a notice necessitating compliance with the requirements within a prescribed time period. Under PCT Rule 13*ter*, the applicant can provide, in response to such a requirement or otherwise, a sequence listing that is a "Sequence Listing XML" in accordance with § 1.831(a). The "Sequence Listing XML" must be accompanied by a statement that the information recorded does not go beyond the disclosure in the international application as filed. In response to such a requirement, the late furnishing fee set forth in § 1.445(a)(5) is also required. If the applicant fails to timely provide the required "Sequence Listing XML," the United States International Searching Authority shall search only to the extent that a meaningful search can be performed without the "Sequence Listing XML," and the United States International Preliminary Examining Authority shall examine only to the extent that a meaningful examination can be performed without the "Sequence Listing XML."

(f) Any appropriate amendments to the "Sequence Listing XML" in a patent (e.g., by reason of reissue, reexamination, or certificate of correction) must comply with the requirements of paragraph (b) of this section.

§ 1.839 Incorporation by reference.

(a) Certain material is incorporated by reference into this subpart with the approval of the Director of the Federal Register under 5 U.S.C. 552(a) and 1 CFR part 51. All approved incorporation by reference (IBR) material is available for inspection at the USPTO and at the National Archives and Records Administration (NARA). Contact the USPTO's Office of Patent Legal Administration at 571-272-7701. For information on the availability of this material at NARA, email fr.inspection@nara.gov or go to www.archives.gov/federal-register/cfr/ibr-locations.html. The material may be obtained from the source(s) in paragraph (b) of this section.

(b) World Intellectual Property Organization (WIPO), 34 chemin des Colombettes, 1211 Geneva 20 Switzerland, www.wipo.int.

(1) WIPO Standard ST.26. WIPO Handbook on Industrial Property Information and Documentation, Standard ST.26: Recommended Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings Using XML (eXtensible Markup Language) including Annexes I-VII, version 1.5, approved November 5, 2021; IBR approved for §§ 1.831 through 1.834.

(2) [Reserved]

Katherine K. Vidal,
Under Secretary of Commerce for Intellectual Property and Director of the United States Patent and Trademark Office.