Chapter 2400 Biotechnology

2401 Introduction [R-01.2024]

This chapter provides guidance on the practices and procedures pertaining to the rules for deposits of biological materials for patent purposes (37 CFR 1.801 - 1.809) and the rules for the requirements for patent applications containing nucleotide sequence and/or amino acid sequence disclosures under World Intellectual Property Organization (WIPO) Standard ST.25 (37 CFR 1.821 - 1.825) and WIPO Standard ST.26 (37 CFR 1.831 - 1.835).

2402 The Deposit Rules [R-07.2015]

37 CFR 1.801 Biological material.

For the purposes of these regulations pertaining to the deposit of biological material for purposes of patents for inventions under 35 U.S.C. 101, the term biological material shall include material that is capable of self-replication either directly or indirectly. Representative examples include bacteria, fungi including yeast, algae, protozoa, eukaryotic cells, cell lines, hybridomas, plasmids, viruses, plant tissue cells, lichens and seeds. Viruses, vectors, cell organelles and other non-living material existing in and reproducible from a living cell may be deposited by deposit of the host cell capable of reproducing the non-living material.

Every patent must contain a written description of the invention sufficient to enable a person skilled in the art to which the invention pertains to make and use the invention. Where the invention involves a biological material and words alone cannot sufficiently describe how to make and use the invention in a reproducible manner, access to the biological material may be necessary for the satisfaction of the statutory requirements for patentability under 35 U.S.C. 112. Courts have recognized the necessity and desirability of permitting an applicant for a patent to supplement the written disclosure in an application with a deposit of biological material which is essential to meet some requirement of the statute with respect to the

claimed invention. See, e.g., *Ajinomoto Co. v. Archer-Daniels-Midland Co.*, 228 F.3d 1338, 1345-46, 56 USPQ2d 1332, 1337-38 (Fed. Cir. 2000), *cert. denied*, 121 S.Ct. 1957 (2001)(explaining how deposit may help satisfy enablement requirement); *Enzo Biochem, Inc. v. Gen-Probe, Inc.*, 323 F.3d 956, 63 USPQ2d 1609 (Fed. Cir. 2002)(deposit may satisfy the written description requirement); *In re Argoudelis*, 434 F.2d 666, 168 USPQ 99 (CCPA 1970). To facilitate the recognition of deposited biological material in patent applications throughout the world, the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure was established in 1977, and became operational in 1981. The Treaty requires signatory countries, like the United States, to recognize a deposit with any depository which has been approved by the World Intellectual Property Organization (WIPO).

The deposit rules (37 CFR 1.801 - 1.809) set forth examining procedures and conditions of deposit which must be satisfied in the event a deposit is required. The rules do not address the substantive issue of whether a deposit is required under any particular set of facts.

See PCT Rule 13bis and MPEP § 1823.01 for the requirements under the PCT for a reference to a deposited biological material in an international application.

2403 Deposit of Biological Material [R-07.2022]

37 CFR 1.801 indicates that the rules pertaining to deposits for purposes of patents for inventions under 35 U.S.C. 101 are intended to relate to biological material. For the purposes of these rules, the term "biological material" is defined in terms of a non-exhaustive list of representative materials which can be deposited in accordance with the procedures defined in these rules. These rules are intended to address procedural matters in the deposit of biological material for patent purposes. They are not designed to decide substantive issues such as whether a deposit of a particular organism or material would be recognized or necessary for the purposes of satisfying the statutory requirements for patentability under 35 U.S.C. 112. The issue of the need to make a deposit of biological material typically arises under 35 U.S.C. 112(a) with regard to the enablement requirement, although the issue may also arise under the written description or best mode requirement of the statute. Deposit issues may also arise under 35 U.S.C. 112(b) with respect to the claims.

37 CFR 1.801 does not attempt to identify what biological material either needs to be or may be deposited to comply with the requirements of 35 U.S.C. 112. For the most part, this issue must be addressed on a case-by-case basis. Thus, while the Office does not currently contemplate that there would be any situations where a material that is not capable of self-replication either directly or indirectly would be acceptable as a deposit, an applicant is clearly not precluded by these rules from attempting to show in any given application why the deposit of such a material should be acceptable to satisfy the requirements of 35 U.S.C. 112. The examiner is cautioned against requiring that a specific biological material be deposited where the applicant can show that a deposit of starting material that is currently not available to the public would allow the skilled artisan to make and use the claimed invention. For example, where a claimed invention is drawn to a plant having novel properties produced by the insertion of a proprietary gene at a specific locus, the plant per se need not be deposited if deposit of a vector or hybridoma containing the gene would enable one of ordinary skill in the art to make and use the claimed invention without undue experimentation.

2403.01 Material Capable of Self-Replication [R-08.2012]

Biological material includes material that is capable of self-replication either directly or indirectly. Direct self-replication includes those situations where the biological material reproduces by itself. Representative examples of materials capable of self-replication are defined in the rule. Indirect self-replication is meant to include those situations where the biological material is only capable of replication when another self-replicating biological material is present. Self-replication after insertion in a host is one example of indirect self-replication. Examples of indirect replicating biological materials

include viruses, phages, plasmids, symbionts, and replication defective cells. The list of representative examples of each type of replicating material includes viruses to demonstrate that the two lists in the rule are not intended to be mutually exclusive.

2403.02 Plant Material [R-10.2019]

Although plant material is included within the scope of the definition of biological material for purposes of patents for plant inventions under 35 U.S.C. 101, the rules on deposits are not applicable to applications filed under the Plant Patent Act (35 U.S.C. 161-164). The Office is of the view that a deposit is not required under the present provisions of 35 U.S.C. 162. Thus, a deposit is not necessary for the grant of a plant patent under the provisions of 35 U.S.C. 161-164. See also MPEP § 1605. As with other biological material deposited for purposes of patents for inventions under 35 U.S.C. 101, the deposit of plant material together with the written specification must enable those skilled in the art to make and use the claimed invention, in accordance with the requirements of 35 U.S.C. 112.

As with some types of reproducible biological material, seeds can be reproduced only after a growing season which may be relatively long. Although the rules do not specify a specific number of seeds to be deposited to meet the requirements of these rules, so long as the number of seeds deposited complies with the requirements of the Budapest Treaty International Depositary Authority (IDA) where the deposit is made, the USPTO would consider such a compliant submission as satisfying the rules under 37 CFR 1.801 through 1.809. Note that the American Type Culture Collection (ATCC), a Budapest IDA, requires a minimum deposit of 625 seeds; other IDAs may have different minimum requirements. Accordingly, any depositor should confirm that the number submitted to a specific IDA complies with that IDA's requirements for seed deposits.

2404 Need or Opportunity to Make a Deposit [R-08.2012]

37 CFR 1.802 Need or opportunity to make a deposit.

(a) Where an invention is, or relies on, a biological material, the disclosure may include reference to a deposit of such biological material.

(b) Biological material need not be deposited unless access to such material is necessary for the satisfaction of the statutory requirements for patentability under 35 U.S.C. 112. If a deposit is necessary, it shall be acceptable if made in accordance with these regulations. Biological material need not be deposited, *inter alia*, if it is known and readily available to the public or can be made or isolated without undue experimentation. Once deposited in a depository complying with these regulations, a biological material will be considered to be readily available even though some requirement of law or regulation of the United States or of the country in which the depository institution is located permits access to the material only under conditions imposed for safety, public health or similar reasons.

(c) The reference to a biological material in a specification disclosure or the actual deposit of such material by an applicant or patent owner does not create any presumption that such material is necessary to satisfy 35 U.S.C. 112 or that deposit in accordance with these regulations is or was required.

37 CFR 1.802(a) permits a deposit of a biological material to be referenced in a patent application where an invention is, or relies on, a biological material. The invention may rely on a biological material for the purposes of making or using the invention, either as a preferred mode or an alternative mode of operation. A reference to a deposit may be included in a specification even though the deposit is not required to satisfy the requirements of 35 U.S.C. 112.

There is no necessary implication or presumption that can or should be made about the need for a deposit simply because reference to a deposit is made in an application disclosure, as noted in paragraph (c). As noted in paragraph (b), biological material need not be deposited unless access to such material is necessary for the satisfaction of the statutory requirements for patentability under 35 U.S.C. 112 and that access is not otherwise available in the absence of a deposit. Where a deposit is required to provide the necessary access, a deposit is acceptable for patent purposes only where it is made in accordance with these regulations. Even where access to biological material is required to satisfy these statutory requirements, a deposit may not be necessary if access sufficient to satisfy these requirements is otherwise available.

2404.01 Biological Material That Is Known and Readily Available to the Public [R-07.2015]

In an application where the invention required access to specific biological material, an applicant could show that the biological material is accessible because it is known and readily available to the public. The concepts of "known and readily available" are considered to reflect a level of public accessibility to a necessary component of an invention disclosure that is consistent with an ability to make and use the invention. To avoid the need for a deposit on this basis, the biological material must be both known and readily available - neither concept alone is sufficient. A material may be known in the sense that its existence has been published, but is not available to those who wish to obtain that particular known biological material. Likewise, a biological material may be available in the sense that those having possession of it would make it available upon request, but no one has been informed of its existence.

The Board has held that a description of the precise geographic location of marine tunicates, as a biological material, used in a claimed invention was adequate to satisfy the enablement requirement of 35 U.S.C. 112. See *Ex Parte Rinehart,* 10 USPQ2d 1719 (Bd. Pat. App. & Int. 1985). The term "readily" used in the phrase "known and readily available" is considered appropriate to define that degree of availability which would be reasonable under the circumstances. If the biological material and its natural location can be adequately described so that one skilled in the art could obtain it using ordinary skill in the art, the disclosure would appear to be sufficient to meet the enablement requirement of 35 U.S.C. 112 without a deposit so long as its degree of availability is reasonable under the circumstances.

By showing that a biological material is known and readily available or by making a deposit in accordance with these rules, applicant does not guarantee that such biological material will be available forever. Public access during the term of the patent may affect the enforceability of the patent. Although there is a public interest in the availability of a deposited biological material during and after the period of enforceability of the patent, there should not be any undue concern about continued access to the public. See 37 CFR 1.806 (the term of deposit is "at least thirty (30) years and at least five (5) years after the most recent request" for a sample; the agreement sufficiently ensures that the deposit will be "available beyond the enforceable life of the patent"). Unless there is a reasonable basis to believe that the biological material will cease to be available during the enforceable life of the patent, current availability would satisfy the requirement. The incentives provided by the patent system should not be constrained by the mere possibility that a disclosure that was once enabling would become non-enabling over a period of time through no fault of the patentee. *In re Metcalfe,* 410 F.2d 1378, 161 USPQ 789 (CCPA 1969).

If an applicant has adequately established that a biological material is known and readily available, the Office will accept that showing. In those instances, however, the applicant takes the risk that the material may cease to be known and readily available. Such a defect cannot be cured by reissue after the grant of a patent.

On the other hand, *Ex parte Humphreys,* 24 USPQ2d 1255 (Bd. Pat. App. & Int. 1992), held that the only manner in which applicants could satisfy their burden of assuring public access to the needed biological material, and, thereby, compliance with the enablement requirement of 35 U.S.C. 112, was by making an appropriate deposit. The fact that applicants and other members of the public were able to obtain the material in question from a given depository prior to and after the filing date of the application in issue did not establish that upon issuance of a patent on the application that such material would continue to be accessible to the public. The applicants did not make of record any of the facts and circumstances surrounding their access to the material in issue from the depository, nor was there any evidence as to the depository's policy regarding the material if a patent would have been granted. Further, there was no assurance that the depository would have allowed unlimited access to the material if the application had matured into a patent.

There are many factors that may be used as indicia that a biological material is known and readily

available to the public. Relevant factors include commercial availability, references to the biological material in printed publications, declarations of accessibility by those working in the field, evidence of predictable isolation techniques, or an existing deposit made in accordance with these rules. Each factor alone may or may not be sufficient to demonstrate that the biological material is known and readily available. Those applicants that rely on evidence of accessibility other than a deposit take the risk that the patent may no longer be enforceable if the biological material necessary to satisfy the requirements of 35 U.S.C. 112 ceases to be accessible.

The Office will accept commercial availability as evidence that a biological material is known and readily available only when the evidence is clear and convincing that the public has access to the material. See the final rule entitled " *Deposit of Biological Materials for Patent Purposes*," 54 FR 34864, 34875 (August 22, 1989). A product could be commercially available but only at a price that effectively eliminates accessibility to those desiring to obtain a sample. The relationship between the applicant relying on a biological material and the commercial supplier is one factor that would be considered in determining whether the biological material was known and readily available. However, the mere fact that the biological material is commercially available only through the patent holder or the patent holder's agents or assigns shall not, by itself, justify a finding that the necessary material is not readily available, absent reason to believe that access to the biological material would later be improperly restricted.

The mere reference to a deposit or the biological material itself in any document or publication does not necessarily mean that the deposited biological material is readily available. Even a deposit made under the Budapest Treaty and referenced in a United States or foreign patent document would not necessarily meet the test for known and readily available unless the deposit was made under conditions that are consistent with those specified in these rules, including the provision that requires, with one possible exception (37 CFR 1.808(b)), that all restrictions on the accessibility be irrevocably removed by the applicant upon the granting of the patent. *Ex parte Hildebrand*, 15 USPQ2d 1662 (Bd. Pat. App. & Int. 1990).

A deposit of biological material cited in a U.S. patent need not be made available if it was not required to satisfy 35 U.S.C. 112. For this reason, 37 CFR 1.808(c) provides that upon request made to the Office, the Office will certify whether a deposit has been stated to have been made under conditions which make it available to the public as of the issue date. See 37 CFR 1.808(c) and MPEP § 2410.02 for the requirements of the request. The Office will not certify that the aforementioned statement has been made unless:

(A) the deposit was necessary to overcome a rejection under 35 U.S.C. 112,

(B) there is, in the record, a statement by the examiner that a rejection would have been made "but for" the deposit (assumes deposit information in record, as filed),

(C) the record otherwise clearly indicates that the deposit was made under Budapest Treaty, and that all restrictions imposed by the depositor on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent (with the possible exception of requiring the request for the deposit to be in the format specified in 37 CFR 1.808(b)), or

(D) the record otherwise clearly indicates that an acceptable non-Budapest Treaty deposit was made and that all restrictions imposed by the depositor on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent (with the possible exception of requiring the request for the deposit to be in the format specified in 37 CFR 1.808(b)).

If a deposit is not made under the conditions set forth in 37 CFR 1.808(a), the deposit cannot be relied upon for other purposes, e.g., the deposit cannot be relied upon by a third party to establish "known" and "readily available" in another application. See 37 CFR 1.808 and MPEP § 2410 and § 2410.02.

Once a deposit is made in a depository complying with these rules, and under conditions complying with these rules, a biological material will be considered to be readily available even though some

requirement of law or regulation in the United States or in the country where the depository institution is located permits access to the material only under conditions imposed for health, safety or similar reasons. This provision is consistent with the Budapest Treaty (Article 5) and is designed to permit the patenting of inventions involving materials having restricted distribution, where the restrictions are imposed for the public, as opposed to the private, welfare.

2404.02 Biological Material That Can Be Made or Isolated Without Undue Experimentation [R-08.2012]

Applicant may show that a deposit is not necessary even though specific biological materials are required to practice the invention if those biological materials can be made or isolated without undue experimentation. Deposits may be required to support the claims if an isolation procedure requires undue experimentation to obtain the desired biological material. *Ex Parte Jackson*, 217 USPQ 804 (Bd. App. 1982). No deposit is required, however, where the required biological materials can be obtained from publicly available material with only routine experimentation and a reliable screening test. *Tabuchi v. Nubel*, 559 F.2d 1183, 194 USPQ 521 (CCPA 1977); *Ex Parte Hata*, 6 USPQ2d 1652 (Bd. Pat. App. & Int. 1987).

2404.03 Reference to a Deposit in the Specification [R-08.2012]

37 CFR 1.802(c) specifically provides that the mere reference to a biological material in the specification disclosure or the actual deposit of such material does not create any presumption that such referenced or deposited material is necessary to satisfy 35 U.S.C. 112, or that a deposit in accordance with these regulations is or was required. It should be noted, however, that a reference to a biological material, present in an application upon filing, may form the basis for making a deposit, where required, after the filing date of a given application but that the reference to the biological material, itself, cannot be added after filing without risking the prohibited introduction of new matter (35 U.S.C. 132). See the discussion of the Lundak application in MPEP § 2406.01.

2405 Acceptable Depository [R-07.2015]

37 CFR 1.803 Acceptable depository.

(a) A deposit shall be recognized for the purposes of these regulations if made in

(1) any International Depositary Authority (IDA) as established under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure, or

(2) any other depository recognized to be suitable by the Office. Suitability will be determined by the Commissioner on the basis of the administrative and technical competence, and agreement of the depository to comply with the terms and conditions applicable to deposits for patent purposes. The Commissioner may seek the advice of impartial consultants on the suitability of a depository. The depository must:

(i) Have a continuous existence;

(ii) Exist independent of the control of the depositor;

(iii) Possess the staff and facilities sufficient to examine the viability of a deposit and store the deposit in a manner which ensures that it is kept viable and uncontaminated;

(iv) Provide for sufficient safety measures to minimize the risk of losing biological material deposited with it;

(v) Be impartial and objective;

(vi) Furnish samples of the deposited material in an expeditious and proper manner; and

(vii) Promptly notify depositors of its inability to furnish samples, and the reasons why.

(b) A depository seeking status under paragraph (a)(2) of this section must direct a communication to the Commissioner which shall:

(1) Indicate the name and address of the depository to which the communication relates;

(2) Contain detailed information as to the capacity of the depository to comply with the requirements of paragraph (a) (2) of this section, including information on its legal status, scientific standing, staff and facilities;

(3) Indicate that the depository intends to be available, for the purposes of deposit, to any depositor under these same conditions;

(4) Where the depository intends to accept for deposit only certain kinds of biological material, specify such kinds;

(5) Indicate the amount of any fees that the depository will, upon acquiring the status of suitable depository under paragraph (a) (2) of this section, charge for storage, viability statements and furnishings of samples of the deposit.

(c) A depository having status under paragraph (a)(2) of this section limited to certain kinds of biological material may

extend such status to additional kinds of biological material by directing a communication to the Commissioner in accordance with paragraph (b) of this section. If a previous communication under paragraph (b) of this section is of record, items in common with the previous communication may be incorporated by reference.

(d) Once a depository is recognized to be suitable by the Commissioner or has defaulted or discontinued its performance under this section, notice thereof will be published in the Official Gazette of the Patent and Trademark Office.

37 CFR 1.803 indicates that a depository will be recognized as acceptable for the purposes of these regulations if it is either an International Depositary Authority (IDA) established under the Budapest Treaty, or if it is a depository recognized as suitable by the Commissioner. After the effective date of these regulations, a deposit of biological material which is made in a depository which is not recognized as acceptable under this regulation will not be considered as satisfying the requirements of 35 U.S.C. 112. See *Ex parte Humphreys,* 24 USPQ2d 1255 (Bd. Pat. App. & Int. 1992). On the other hand, if a deposit is not required to satisfy the requirements of 35 U.S.C. 112, it is permissible to make reference to such a deposit even though it may not be in a depository or made under the conditions which are acceptable under these regulations. As new depositories are recognized as suitable by the Commissioner, their identity will be announced in the *Official Gazette.*

An organization may be recognized as suitable by the Office if the procedure and conditions specified in 37 CFR 1.803(a)(2) and 37 CFR 1.803(b) are followed. Generally, it is not the intention of the Office to recognize as suitable any organization where the need for a suitable depository for patent purposes is being met by depositories recognized as IDAs under the Budapest Treaty. Suitability will be judged by the Commissioner, based on need and the information supplied by the organization seeking status, and information obtained from other sources that may be consulted.

While there is a desire to provide flexibility to a patent applicant in selecting an appropriate depository, these rules are not intended to permit each patent applicant to become its own depository since both the patent owner and the public have an interest in the continued availability and accessibility of the deposit during the enforceable life of the patent, and the public has a continuing interest in its availability when the patent is no longer enforceable. The concept of a depository independent of the control of the depositor or an IDA as an acceptable depository is based on the need and desire to ensure the safe and reliable storage of a deposited biological material under circumstances that are substantially free of the opportunity for intentional mishandling or negligent handling of the deposited material. The use of an independent depository or internationally recognized depository will tend to preserve the integrity of the deposit process against those that may accidentally alter the deposited material, may wish to tamper with the deposited material or may wish to resume control of its availability when the patent is no longer enforceable, and will tend to preserve the interest of the public in the access to the biological material once the term of the patent expires.

When a depository having status under 37 CFR 1.803(a)(2) seeks to change the kinds of biological materials that it will accept and maintain for the purposes of these rules, a communication requesting such a change should be directed to the Commissioner containing the information requested in 37 CFR 1.803(b). When such a change is requested, the requesting depository should provide a complete list of the kinds of biological materials it will accept.

37 CFR 1.803(d) indicates that once a depository is recognized as suitable for the purposes of this rule, or has defaulted or discontinued its performance under this section, notice thereof will be published in the *Official Gazette* of the Patent and Trademark Office. The mere fact that a deposit has been made in an IDA recognized under the Budapest Treaty does not mean that the terms of the deposit meet either the requirements of the Budapest Treaty or the deposit regulations. Many of the depositories recognized under the Budapest Treaty have many different arrangements under which biological material may be stored.

The World Intellectual Property Organization (WIPO) publishes a Guide to the Deposit of Micro-organisms under the Budapest Treaty (WIPO Publication No. 661 (E)) on the procedures and requirements concerning the deposit of biological

material, including procedures for obtaining a sample of deposited material, in each of the international depository authorities. The Guide and a list of current IDAs under the Budapest Treaty are available at **www.wipo.int/treaties/en/registration/budapest/**.

2406 Time of Making an Original Deposit [R-07.2015]

37 CFR 1.804 Time of making an original deposit.

(a) Whenever a biological material is specifically identified in an application for patent as filed, an original deposit thereof may be made at any time before filing the application for patent or, subject to § 1.809, during pendency of the application for patent.

(b) When the original deposit is made after the effective filing date of an application for patent, the applicant must promptly submit a statement from a person in a position to corroborate the fact, stating that the biological material which is deposited is a biological material specifically identified in the application as filed.

37 CFR 1.804 specifies the time for making an original deposit to fulfill the requirements of 35 U.S.C. 112. For the reasons discussed throughout this section, it is recommended that a deposit be made before the filing date of the application. However, for the purposes of complying with the requirements of 35 U.S.C. 112, a deposit of a biological material may be made at any time before filing the application for patent or during the pendency of the application subject to the conditions of 37 CFR 1.809. Where a deposit is needed to satisfy the requirements of 35 U.S.C. 112 and it is made during the pendency of the application, it must be made no later than the time period set by the examiner at the time the Notice of Allowance and Issue Fee Due is mailed. However, a necessary deposit need not be made by an applicant until the application is in condition for allowance so long as the applicant provides a written assurance that an acceptable deposit will be made on or before the payment of the issue fee. This written assurance must provide sufficiently detailed information to convince the examiner that there is no outstanding issue regarding deposits that needs to be resolved. Note that while 37 CFR 1.804 permits making a deposit after the filing date of an application, in many countries the deposit must be made before the filing date of the application. See MPEP § 2406.03.

Insofar as the rules do not permit post-issuance original deposits, the failure to make an original deposit in an application cannot be cured by filing a reissue application or instituting a reexamination proceeding. However, if an amendment of claims in a reexamination proceeding raises the need for a deposit, an original deposit may be made during the reexamination proceeding.

2406.01 Description in Application Specification [R-07.2015]

37 CFR 1.804(a) specifies not only a permissible time frame for making an original deposit, but also specifies that the biological material deposited must be specifically identified in the application for patent as filed. The requirement for a specific identification is consistent with the description requirement of 35 U.S.C. 112 and provides an antecedent basis for the biological material which either has been or will be deposited before the patent is granted.

The description in the Lundak application as filed (now U.S. Patent No. 4,594,325) provides a suitable illustration of the specific identification and description which are required in an application as filed. In that application, an immortal B-cell line was disclosed and claimed. The cell line was referred to in the application, as filed, as WI-L2-729 HF2. The methods of obtaining and using this cell line were also described in the application as filed. A deposit of the cell line was made with the American Type Culture Collection (ATCC) about a week after the application was filed in the United States. The United States Court of Appeals for the Federal Circuit held that the requirements of access by the Office to a sample of the cell line during pendency, and public access after grant, were met by Lundak's procedures. The Court further held that the addition of information designating the depository, accession number, and deposit date of the deposited cell line in ATCC after the filing date did not violate the prohibition against new matter in 35 U.S.C. 132. *In re Lundak*, 773 F.2d 1216, 227 USPQ 90 (Fed. Cir. 1985). However, it must be clear from the application as filed that the invention claimed and described in the specification “was fully capable of being reduced to practice (i.e., no technological problems, the resolution of which would require more than ordinary skill and reasonable time,

remained in order to obtain an operative, useful process)." *Feldman v. Aunstrup*, 517 F.2d 1351, 1355, 186 USPQ 108, 113 (CCPA 1975), *cert. denied*, 424 U.S. 912 (1976).

2406.02 Deposit After Filing Date - Corroboration [R-08.2012]

When the original deposit is made after the effective filing date of an application for patent, an applicant is required to promptly submit a statement from a person in a position to corroborate that the biological material which is deposited is a biological material specifically identified in the application (the filing date of which is relied upon) as filed. The nature of this corroboration will depend on the circumstances in the particular application under consideration, including the length of time between the application filing date and the date of deposit. While few, if any, situations can be imagined where the description requirement of 35 U.S.C. 112 can be satisfied where the biological material was not in existence at the time of filing, the rules will not preclude such a situation as there is no requirement in the patent law that an actual reduction to practice occur as a condition precedent to filing a patent application.

2406.03 Possible Loss of U.S. Filing Date in Other Countries [R-07.2015]

Those applicants intending to file patent applications in a country other than the United States relying upon biological material that must be deposited to satisfy the requirements of 35 U.S.C. 112 when the application is filed in the United States are cautioned that in many countries the deposit must be made before the filing date of the priority application in order to obtain foreign priority rights. Thus, while the deposit of a biological material subsequent to the effective filing date of a United States application may be relied upon to comply with 35 U.S.C. 112, an applicant may not be able to rely on the filing date of such a U.S. application if a patent is sought in certain countries other than the United States.

2407 Replacement or Supplement of Deposit [R-08.2012]

37 CFR 1.805 Replacement or supplement of deposit.

(a) A depositor, after receiving notice during the pendency of an application for patent, application for reissue patent or reexamination proceeding, that the depository possessing a deposit either cannot furnish samples thereof or can furnish samples thereof but the deposit has become contaminated or has lost its capability to function as described in the specification, shall notify the Office in writing, in each application for patent or patent affected. In such a case, or where the Office otherwise learns, during the pendency of an application for patent, application for reissue patent or reexamination proceeding, that the depository possessing a deposit either cannot furnish samples thereof or can furnish samples thereof but the deposit has become contaminated or has lost its capability to function as described in the specification, the need for making a replacement or supplemental deposit will be governed by the same considerations governing the need for making an original deposit under the provisions set forth in § 1.802(b). A replacement or supplemental deposit made during the pendency of an application for patent shall not be accepted unless it meets the requirements for making an original deposit under these regulations, including the requirement set forth under § 1.804(b). A replacement or supplemental deposit made in connection with a patent, whether or not made during the pendency of an application for reissue patent or a reexamination proceeding or both, shall not be accepted unless a certificate of correction under § 1.323 is requested by the patent owner which meets the terms of paragraphs (b) and (c) of this section.

(b) A request for certificate of correction under this section shall not be granted unless the certificate identifies:

(1) The accession number for the replacement or supplemental deposit;

(2) The date of the deposit; and

(3) The name and address of the depository.

(c) A request for a certificate of correction under this section shall not be granted unless the request is made promptly after the replacement or supplemental deposit has been made and the request:

(1) Includes a statement of the reason for making the replacement or supplemental deposit;

(2) Includes a statement from a person in a position to corroborate the fact, and stating that the replacement or supplemental deposit is of a biological material which is identical to that originally deposited;

(3) Includes a showing that the patent owner acted diligently —

(i) In the case of a replacement deposit, in making the deposit after receiving notice that samples could no longer be furnished from an earlier deposit; or

(ii) In the case of a supplemental deposit, in making the deposit after receiving notice that the earlier deposit had become contaminated or had lost its capability to function as described in the specification;

(4) Includes a statement that the term of the replacement or supplemental deposit expires no earlier than the term of the deposit being replaced or supplemented; and

(5) Otherwise establishes compliance with these regulations.

(d) A depositor's failure to replace a deposit, or in the case of a patent, to diligently replace a deposit and promptly thereafter request a certificate of correction which meets the terms of paragraphs (b) and (c) of this section, after being notified that the depository possessing the deposit cannot furnish samples thereof, shall cause the application or patent involved to be treated in any Office proceeding as if no deposit were made.

(e) In the event a deposit is replaced according to these regulations, the Office will apply a rebuttable presumption of identity between the original and the replacement deposit where a patent making reference to the deposit is relied upon during any Office proceeding.

(f) A replacement or supplemental deposit made during the pendency of an application for patent may be made for any reason.

(g) In no case is a replacement or supplemental deposit of a biological material necessary where the biological material, in accordance with § 1.802(b), need not be deposited.

(h) No replacement deposit of a biological material is necessary where a depository can furnish samples thereof but the depository for national security, health or environmental safety reasons is unable to provide samples to requesters outside of the jurisdiction where the depository is located.

(i) The Office will not recognize in any Office proceeding a replacement deposit of a biological material made by a patent owner where the depository could furnish samples of the deposit being replaced.

37 CFR 1.805 relates to the deposit of a biological material to replace or supplement a previous deposit. The term "replacement" is directed to those situations where one deposit is being substituted for another. An applicant may have greater latitude in replacing a deposit during the pendency of an application than after the patent is granted. Replacement will typically take place where the earlier deposit is no longer viable. The term "supplement" is directed to those situations where the earlier deposit is still viable in the sense that it is alive and capable of replication either directly or indirectly, but has lost a quality (e.g., purity, functionality) it allegedly possessed at the time the application was filed.

2407.01 In a Pending Application [R-07.2015]

Pursuant to 37 CFR 1.805(a), an applicant is required to notify the Office when it obtains information that the depository possessing a deposit referenced in an application cannot furnish samples of the deposit. When the Office is so informed or otherwise becomes aware that samples of the deposited material cannot be furnished by the depository, the examiner will treat the application as if no deposit existed. A replacement or supplemental deposit must be made if access to the deposited material is necessary to satisfy the requirements for patentability under 35 U.S.C. 112. A replacement or supplemental deposit will be accepted if it meets all the requirements for making an original deposit.

It should be noted that in a pending application, an applicant need not replace the identical material previously deposited, but may make an original deposit of a biological material which is specifically identified and described in the application as filed. Whether this alternative deposit will meet the requirements of 35 U.S.C. 112 with respect to the claimed subject matter must be resolved by the examiner on a case-by-case basis. The conditions in 37 CFR 1.802(b) and 37 CFR 1.804(b) must be satisfied.

2407.02 After a Patent Has Issued [R-07.2015]

A patent owner is required to notify the Office when it obtains information that a depository possessing a deposit referenced in a patent cannot furnish samples of the deposit. Failure to diligently replace the deposit and promptly thereafter request a certificate of correction which meets the terms of 37 CFR 1.805(b) and 37 CFR 1.805(c) will cause the patent involved to be treated in any Office proceeding as if no deposit were made.

A replacement or supplemental deposit made in connection with a patent, whether or not made during the pendency of an application for reissue patent or a reexamination proceeding or both, shall not be accepted unless a certificate of correction under 37 CFR 1.323 is requested which meets the terms of 37 CFR 1.805(b) and 37 CFR 1.805(c) for replacement or supplemental deposits. See MPEP § 1411.01 for including changes that were made by a certificate of correction to the original patent grant in a reissue application, and MPEP § 2219 for including a copy of any certificate of correction to the original patent grant as part of a request for *ex parte* reexamination.

37 CFR 1.805(b) and 37 CFR 1.805(c) specify the procedures that a patent owner may follow to ensure that a patent contains the appropriate information about a deposited biological material in the event that a replacement or supplemental deposit is made after the patent is granted. 37 CFR 1.805(b) describes the information which must be contained in the certificate of correction, whereas 37 CFR 1.805(c) describes the information which must be provided in the request to make the correction.

A request for a certificate of correction of a patent under 37 CFR 1.805(b) and 37 CFR 1.805(c) will not be granted where no original deposit was made before or during the pendency of the application which matured into the patent. A patent defective because of lack of a necessary deposit is necessarily fatally defective for failure to comply with the first paragraph of 35 U.S.C. 112. Reissue is not available in such cases. See *In re Hay*, 534 F.2d 917, 189 USPQ 790 (CCPA 1976).

2407.03 Failure to Replace [R-08.2012]

37 CFR 1.805(d) sets forth the Office position that the failure to make a replacement deposit in a case pending before the Office, for example a reissue or reexamination proceeding, where a deposit is considered to be necessary to satisfy the requirements of 35 U.S.C. 112, shall cause the application or patent involved to be treated in any Office proceeding as if no deposit were made. The provisions of 37 CFR 1.805(g) indicate that a replacement need not be made where, at the point in time when replacement would otherwise be necessary, access to the necessary biological material was otherwise available. For example, a replacement deposit would not be required under the circumstances where access to the necessary biological material was established through commercial suppliers.

2407.04 Treatment of Replacement [R-08.2012]

37 CFR 1.805(e) indicates that the Office will apply a rebuttable presumption of identity between the replacement deposit and an original deposit where a patent making reference to the deposit is relied on during any Office proceeding. This means that where a replacement deposit is permitted and made, the examiner will assume that the same material as described in the patent is accessible from the identified depository unless evidence to the contrary comes to the attention of the Office.

An applicant for patent may make a replacement deposit during the pendency of the application for any reason. The provisions of 37 CFR 1.805(f) recognize that since an original deposit may be made during the pendency of the application subject to the conditions of 37 CFR 1.809, a replacement deposit logically cannot be held to any higher standard or any further requirements.

2407.05 Exemption From Replacement [R-08.2012]

The provisions of 37 CFR 1.805(h) indicate that a replacement deposit is not required even though the depository cannot furnish samples, under certain conditions, to those requesting a sample outside of the jurisdiction where the depository is located. The conditions are specified in this paragraph as being limited to national security, health or environmental safety reasons. See also Article 5 of the Budapest Treaty.

2407.06 Replacement May Not Be Recognized [R-07.2015]

37 CFR 1.805(i) indicates that the Office will not recognize in any Office proceeding a replacement deposit made by the patent owner where the depository could furnish samples of the original deposit being replaced. The best evidence of what was originally deposited should not be lost through destruction or replacement if made in association with an existing patent. A supplemental deposit may be accepted in an Office proceeding, however, depending on the circumstances in each case.

2408 Term of Deposit [R-07.2022]

37 CFR 1.806 Term of deposit.

A deposit made before or during pendency of an application for patent shall be made for a term of at least thirty (30) years and at least five (5) years after the most recent request for the furnishing of a sample of the deposit was received by the depository. In any case, samples must be stored under

agreements that would make them available beyond the enforceable life of the patent for which the deposit was made.

The term of deposit must satisfy the requirements of the Budapest Treaty which sets a term of at least 30 years from the date of deposit and at least 5 years after the most recent request for the furnishing of a sample of the deposit was received by the depository. In the event that the 30-year term of deposit covers the 17-year term or 20-year term of the patent (see MPEP § 2701) plus 6 years to include the Statute of Limitations, no further requirement is necessary. Unless applicant indicates that the deposit has been made under the Budapest Treaty, applicant must indicate the term for which the deposit has been made. The mere possibility of patent term extension or extended litigation involving the patent should not be considered in this analysis.

In the event that the 30-year term of deposit measured from the date of deposit would necessarily terminate within the period of enforceability of the patent (the patent term (see MPEP § 2701) plus 6 years to include the Statute of Limitations), samples must be stored under agreements that would make them available beyond the enforceable life of the patent for which the deposit was made. No requirement should be made as to any particular period of time beyond the enforceable life of the patent. The purpose of the requirement is to ensure that a deposited biological material necessary for the practice of a patented invention would be available to the public after expiration of the patent for which the deposit was made. The term of the deposit must comply with the requirements of each sentence of 37 CFR 1.806 whether or not the deposit is made under the Budapest Treaty. A specific statement that the deposit would be stored under agreements that would make them available beyond the enforceable life of the patent for which the deposit was made is required only where the 30-year term of deposit would terminate within the enforceable life of the patent.

2409 Viability of Deposit [R-07.2015]

37 CFR 1.807 Viability of deposit.

(a) A deposit of biological material that is capable of self-replication either directly or indirectly must be viable at the time of deposit and during the term of deposit. Viability may be tested by the depository. The test must conclude only that the deposited material is capable of reproduction. No evidence is necessarily required regarding the ability of the deposited material to perform any function described in the patent application.

(b) A viability statement for each deposit of a biological material defined in paragraph (a) of this section not made under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure must be filed in the application and must contain:

(1) The name and address of the depository;

(2) The name and address of the depositor;

(3) The date of deposit;

(4) The identity of the deposit and the accession number given by the depository;

(5) The date of the viability test;

(6) The procedures used to obtain a sample if the test is not done by the depository; and

(7) A statement that the deposit is capable of reproduction.

(c) If a viability test indicates that the deposit is not viable upon receipt, or the examiner cannot, for scientific or other valid reasons, accept the statement of viability received from the applicant, the examiner shall proceed as if no deposit has been made. The examiner will accept the conclusion set forth in a viability statement issued by a depository recognized under § 1.803(a).

37 CFR 1.807 requires that the deposit of biological material that is capable of self-replication either directly or indirectly must be viable at the time of deposit and during the term of deposit. This requirement for viability is essentially a requirement that the deposited material is capable of reproduction. For the purpose of making a deposit under these rules, there is no requirement that evidence be provided that the deposited material is capable or has the ability to perform any function described in the patent application. However, as with any other issue of description or enablement, if the examiner has evidence or reason to question the objective statements made in the patent application, applicants may be required to demonstrate that the deposited biological material will perform in the manner described.

Under the Budapest Treaty, there is a requirement that the deposit be tested for viability before it is accepted. Thus, a mere statement by an applicant, an authorized representative of applicant or the assignee that the deposit has been accepted under the Budapest Treaty would satisfy 37 CFR 1.807. The examiner should note the clear distinction between a statement by the applicant that the deposit

has been made under the Budapest Treaty and one in which the deposit has been made and accepted under the Budapest Treaty. Where a statement is merely an indication that a deposit has been made (with no indication as to whether it has been accepted), there is no assurance that the requirements under 35 U.S.C. 112 have been satisfied.

For each deposit which is not made under the Budapest Treaty, a viability statement must be filed in the patent application and contain the information listed in paragraph (b) of this section. Under 37 CFR 1.807(c), the examiner will accept the conclusion set forth in a viability statement which is issued by a depository recognized under 37 CFR 1.803(a). If the viability test indicates that the deposit is not viable upon receipt, or the examiner cannot, for scientific or other valid reasons, accept the statement of viability received from the applicant, the examiner shall so notify the applicant stating the reasons for not accepting the statement and proceed with the examination process as if no deposit had been made.

2410 Furnishing of Samples [R-08.2012]

37 CFR 1.808 Furnishing of samples.

(a) A deposit must be made under conditions that assure that:

(1) Access to the deposit will be available during pendency of the patent application making reference to the deposit to one determined by the Commissioner to be entitled thereto under § 1.14 and 35 U.S.C. 122, and

(2) Subject to paragraph (b) of this section, all restrictions imposed by the depositor on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent.

(b) The depositor may contract with the depository to require that samples of a deposited biological material shall be furnished only if a request for a sample, during the term of the patent:

(1) Is in writing or other tangible form and dated;

(2) Contains the name and address of the requesting party and the accession number of the deposit; and

(3) Is communicated in writing by the depository to the depositor along with the date on which the sample was furnished and the name and address of the party to whom the sample was furnished.

(c) Upon request made to the Office, the Office will certify whether a deposit has been stated to have been made under conditions which make it available to the public as of the issue date of the patent grant provided the request contains:

(1) The name and address of the depository;

(2) The accession number given to the deposit;

(3) The patent number and issue date of the patent referring to the deposit; and

(4) The name and address of the requesting party.

2410.01 Conditions of Deposit [R-07.2015]

37 CFR 1.808 requires that the deposit of biological material be made under two conditions:

(A) access to the deposit will be available during pendency of the patent application making reference to the deposit to one determined by the Commissioner to be entitled thereto under 37 CFR 1.14 and 35 U.S.C. 122, and

(B) with one exception, that all restrictions imposed by the depositor on the availability to the public of the deposited biological material be irrevocably removed upon the granting of the patent.

The one exception that is permitted is specified in 37 CFR 1.808(b) which permits the depositor to contract with the depository to require that samples of a deposited biological material shall be furnished only if a request for a sample, during the term of the patent, meets any one or all of the three conditions specified in this paragraph. These conditions are:

(A) the request is in writing or other tangible form and dated; and/or

(B) the request contains the name and address of the requesting party and the accession number of the deposit; and/or

(C) the request is communicated in writing by the depository to the depositor along with the date on which the sample was furnished and the name and address of the party to whom the sample was furnished.

It should be noted that this exception to the general rule that all restrictions will be removed must be strictly followed and that no variations of this explicit exception will be accepted as meeting the conditions of this section. Although this exception is consistent with the provisions in the Budapest Treaty and its implementing regulations (Rule 11.4), other conditions on accessibility are permitted under the Budapest Treaty as prescribed by national law. Consequently, the mere indication that a deposit has been made and accepted under conditions prescribed

by the Budapest Treaty would satisfy all conditions of these regulations except the requirement that all restrictions on access be removed on grant of the patent. *Ex parte Hildebrand*, 15 USPQ2d 1662 (Bd. Pat. App. & Int. 1990).

2410.02 Certification of Statement of Availability of Deposit [R-01.2024]

Since the mere description of a deposit or identity of a deposit in a patent specification is not necessarily an indication that a requirement for deposit was made or that a deposit which complies with these rules has been made, accessibility to a deposited material referenced in a patent may depend on the satisfaction of conditions not apparent on the face of the patent. For these reasons, and upon request made to the U.S. Patent and Trademark Office, the Office will certify whether a deposit has been stated to have been made under conditions which make it available to the public as of the issue date of the patent grant provided the request is made to a Director of Technology Center (TC) 1600, and contains the following information:

(A) the name and address of the depository;

(B) the accession number given to the deposit;

(C) the patent number and issue date of the patent referring to the deposit; and

(D) the name and address of the requesting party.

Persons requesting a certificate of statement of availability of deposit should contact the TC 1600 Director's office directly, and should not submit a request via the examiner of record.

The Office will not certify whether a deposit has been made under conditions which would make it available to the public until the issuance of a U.S. Patent referencing the deposit. For example, a request for such certification will not be considered where the request is filed with regard to a deposit referenced in an international application filed under the PCT, in an international application that entered the U.S. national stage, or any other U.S. national application, that has not yet issued as a patent.

See also MPEP § 2404.01.

For those deposits made pursuant to the Budapest Treaty, the World Intellectual Property Organization provides a form (Form BP-12) for requesting a certification of legal entitlement to receive samples of deposited microorganisms pursuant to Rule 11.3(a) of the Regulations under the Budapest Treaty. The form is available at **www.wipo.int/budapest/en/guide/**.

2411 Examination Procedures [R-07.2015]

37 CFR 1.809 Examination procedures.

(a) The examiner shall determine pursuant to § 1.104 in each application for patent, application for reissue patent or reexamination proceeding if a deposit is needed, and if needed, if a deposit actually made is acceptable for patent purposes. If a deposit is needed and has not been made or replaced or supplemented in accordance with these regulations, the examiner, where appropriate, shall reject the affected claims under the appropriate provision of 35 U.S.C. 112, explaining why a deposit is needed and/or why a deposit actually made cannot be accepted.

(b) The applicant for patent or patent owner shall reply to a rejection under paragraph (a) of this section by—

(1) In the case of an applicant for patent, either making an acceptable original, replacement, or supplemental deposit, or assuring the Office in writing that an acceptable deposit will be made; or, in the case of a patent owner, requesting a certificate of correction of the patent which meets the terms of paragraphs (b) and (c) of § 1.805, or

(2) Arguing why a deposit is not needed under the circumstances of the application or patent considered and/or why a deposit actually made should be accepted. Other replies to the examiner's action shall be considered nonresponsive. The rejection will be repeated until either paragraph (b)(1) of this section is satisfied or the examiner is convinced that a deposit is not needed.

(c) If an application for patent is otherwise in condition for allowance except for a needed deposit and the Office has received a written assurance that an acceptable deposit will be made, the Office may notify the applicant in a notice of allowability and set a three month period of time from the mailing date of the notice of allowability within which the deposit must be made in order to avoid abandonment. This time period is not extendable under § 1.136 (see § 1.136(c)).

(d) For each deposit made pursuant to these regulations, the specification shall contain:

(1) The accession number for the deposit;

(2) The date of the deposit;

(3) A description of the deposited biological material sufficient to specifically identify it and to permit examination; and

(4) The name and address of the depository.

(e) Any amendment required by paragraphs (d)(1), (d)(2) or (d)(4) of this section must be filed before or with the payment of the issue fee (see § 1.312).

37 CFR 1.809 sets forth procedures that will be used by the examiner to address a deposit issue. The burden is initially on the Office to establish that access to a biological material is necessary for the satisfaction of the statutory requirements for patentability under 35 U.S.C. 112. Once the Office has met this burden, the burden shifts to the applicant or patent owner to demonstrate that access to such biological material either is not necessary, or is already available, or that a deposit of such material will be made in accordance with these regulations.

2411.01 Rejections Based on Deposit Issue [R-07.2015]

Under 37 CFR 1.809(a), once the examiner has determined that access to a biological material is necessary, and there is no information that would support the conclusion that access is currently available in accordance with these regulations, the examiner should make an appropriate rejection under 35 U.S.C. 112 until such time as a deposit in accordance with these regulations is actually made or a written assurance is received in the patent application that such a deposit will be made upon an indication of allowability of the application. The examiner should clearly indicate the statutory basis for the rejection and the reasons that are relied upon by the examiner to conclude that the application does not comply with some requirement of 35 U.S.C. 112. Although not exhaustive, the following grounds of rejection may be applicable in appropriate circumstances:

(A) *35 U.S.C. 112(a), enablement requirement.* Rejection for lack of an enabling disclosure without access to a specific biological material. This ground of rejection should be accompanied by evidence of scientific reasoning to support the conclusion that a person skilled in the art could not make or use the invention defined in and commensurate with the claims without access to the specific biological material. See MPEP § 2164.06(a).

(B) *35 U.S.C. 112(a), written description requirement.* A rejection for lack of written description typically arises in the context that the application as filed does not contain a description to support an amendment to the specification or claims, although it can arise in the context of original claims. See MPEP § 2163 *et seq.* An amendment to the claims that is not described in the application as filed would justify a rejection of the affected claims under 35 U.S.C. 112(a). If an amendment is made to the application, other than the claims, that is not described in the application as filed, this would justify a rejection under 35 U.S.C. 112(a) and/or an objection under 35 U.S.C. 132 (prohibition against the introduction of new matter) and a requirement that the amendment be canceled.

(C) *35 U.S.C. 112(a), best mode requirement.* A rejection for lack of a best mode will be rare in the *ex parte* examination process because it requires (1) a finding by the examiner that, at the time the application was filed, the inventor(s) knew of a specific material that was considered by the inventor(s) to be better than any other, and (2) if a best mode was contemplated at that time, that the inventor(s) concealed the best mode (accidentally or intentionally) by failing to adequately describe that best mode. See *Chemcast Corp. v. Arco Industries Corp.*, 913 F.2d 923, 16 USPQ2d 1033 (Fed. Cir. 1990). See also MPEP § 2165 *et seq.*

(D) *35 U.S.C. 112(b), requirement to particularly point out and distinctly claim invention.* A rejection for indefiniteness, as applied to a deposit issue, requires the examiner to provide reasons why the terms in the claims and/or scope of the invention are unclear because of an incomplete or inaccurate description or the absence of a reference to a biological material.

(E) *35 U.S.C. 112(b), requirement to claim what the inventor or a joint inventor regards as the invention.* A rejection on the basis that the claims do not set forth the subject matter that the inventor or a joint inventor regards as the invention requires the citation of some evidence, not contained in the application as filed, to support the rejection. *In re Prater,* 415 F.2d 1393, 162 USPQ 541 (CCPA 1969). Any disagreement between the content of the application disclosure and the scope of the claims should be addressed under 35 U.S.C. 112(a). See *In re Ehrreich,* 590 F.2d 902, 200 USPQ 504 (CCPA 1979).

Where a deposit is required to satisfy 35 U.S.C. 112, a deposit must be made in accordance with these regulations. See 37 CFR 1.802. A deposit accepted

in any International Depositary Authority (IDA) under the Budapest Treaty shall be accepted for patent purposes if made under conditions which comply with 37 CFR 1.806 and 37 CFR 1.808(a) concerning term of deposit and permissible conditions on access once the patent is granted.

2411.02 Replies to Rejections Based on Deposit Issue [R-07.2015]

Once a rejection under 35 U.S.C. 112 has been made by the examiner directed to the absence of access to a biological material, applicant may reply, pursuant to 37 CFR 1.809(b)(1), by either making an acceptable original, replacement, or supplemental deposit in accordance with these regulations, or assuring the Office in writing that an acceptable deposit will be made on or before the date of payment of the issue fee, or by submitting an argument of why a deposit is not required under the circumstances of the application being considered. Other replies to such a rejection by the examiner shall be considered nonresponsive and may result in abandonment of the application. The rejection will be repeated and made final until the requirements of 37 CFR 1.809(b)(1) are satisfied or the examiner is convinced that a deposit is not required for the claimed subject matter. Once the rejection is made final, the requirements of 37 CFR 1.116 apply to further submissions. The written assurance will be accepted by the Office if it clearly states that an acceptable deposit will be made within the required time and under conditions which satisfy these rules. In the case that an acceptable written assurance has been made by the applicant, the rejection under 35 U.S.C. 112 directed to the absence of access to the biological material should be removed.

2411.03 Application in Condition for Allowance Except for Deposit [R-07.2015]

As set forth in 37 CFR 1.809(c), in the event that an application for patent is otherwise in condition for allowance except for a required deposit and the Office has received a written assurance that an acceptable deposit will be made, the Office may notify the applicant in a notice of allowability and set a three month period of time from the mailing date of the notice of allowability within which the deposit must be made in order to avoid abandonment. This time period is not extendable under 37 CFR 1.136 (see 37 CFR 1.136(c)). Failure to make the needed deposit in accordance with this requirement will be considered a failure to prosecute the application under 35 U.S.C. 133 and result in abandonment of the application.

Once the deposit has been made, information regarding the deposit, such as the name and address of the depository, the accession number and the date of the deposit, that is to be added to the specification must be added by means of filing an amendment under the provisions of 37 CFR 1.312. Such an amendment must be filed before or with the payment of the issue fee. Therefore, applicants need to make any necessary deposit of biological material well prior to payment of the issue fee such that the accession number is received with sufficient time remaining to amend the specification as required by 37 CFR 1.809(d) on or before the date the issue fee is paid. See 37 CFR 1.809(e).

2411.04 [Reserved]

2411.05 Content of Application with Respect to Deposited Material [R-08.2012]

37 CFR 1.809(d) sets forth the requirements for the content of the specification with respect to a deposited biological material. Specifically, the specification shall contain the accession number for the deposit, the date of the deposit, the name and address of the depository, and a description of the deposited biological material sufficient to specifically identify it and to permit examination. The description also must be sufficient to permit verification that the deposited biological material is in fact that disclosed. Once the patent issues, the description must be sufficient to aid in the resolution of questions of infringement. As a general rule, the more information that is provided about a particular deposited biological material, the better the examiner will be able to compare the identity and characteristics of the deposited biological material with the prior art.

2412 The Requirements for Patent Applications Containing Nucleotide Sequence

and/or Amino Acid Sequence Disclosures to Include a Sequence Listing in XML file format [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

Patent applications that disclose a nucleotide and/or amino acid sequence(s) by enumeration of its residues, as defined in 37 CFR 1.831(b), must present each sequence and associated sequence data in a standardized electronic eXtensible Markup Language (XML) format, (a "Sequence Listing XML") as a separate part of the specification. This standardized format is set forth in the World Intellectual Property Organization (WIPO) Standard ST.26, and applies to sequence listings in international applications filed under the Patent Cooperation Treaty (PCT) and in national and regional applications filed 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 all IPOs of WIPO member states. The regulatory provisions found at 37 CFR 1.831 - 1.835 implement WIPO Standard ST.26 in the USPTO and set forth requirements for presenting sequence data in patent applications filed on or after July 1, 2022, that disclose nucleotide sequences and/or amino acid sequences.

WIPO Standard ST.26 is incorporated by reference into the USPTO regulations by including new regulatory text at 37 CFR 1.839:

37 CFR 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 Intellectual 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 1.6, approved November 25, 2022; IBR approved for §§ 1.831 through 1.834.

(2) [Reserved]

For ease of access, WIPO Standard ST.26 can be found at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**

A link to WIPO Standard ST.26 is also found on the USPTO's Sequence Listing Resource Center:

www.uspto.gov/patents/apply/sequence-listing-resource-center

2412.01 Overview of the Sequence Rules [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

Under the sequence listing rules, 37 CFR 1.831 - 1.834, an application having a filing date on or after July 1, 2022, that discloses one or more nucleotide and/or amino acid sequences by enumeration of its residues, as defined in 37 CFR 1.831(a), must contain, as a part of the description, a sequence listing in eXtensible Markup Language (XML) format, where the XML file of the sequence information conforms to the requirements of 37 CFR 1.831 - 1.834, which specify requirements of particular paragraphs of WIPO Standard ST.26. For U.S. applications, the sequence rules do not require that all sequences and associated sequence information contained in the required sequence listing part of the description be disclosed elsewhere in the application; the content of a sequence listing is considered part of the disclosure of the invention. However, for international applications, all

sequences and associated sequence information contained in the required sequence listing part of the description is considered an application part if present on the filing date without an incorporation by reference.

2412.02 Definition of "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

37 CFR 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 a nucleotide and/or amino acid sequences by enumeration of its 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 sequence(s) disclosed in the patent application using the symbols and format in accordance with the requirements of §§ 1.832 through 1.834.

*****

For 35 U.S.C. 111 applications and international applications filed on or after July 1, 2022, that disclose one or more nucleotide and/or amino acid sequences as enumerated by its residues, each nucleotide and/or amino acid sequence and associated sequence data must be presented as a separate part of the disclosure that comprises the sequences in computer readable XML format in accordance with WIPO Standard ST.26 as implemented by 37 CFR 1.831 - 1.834. This sequence listing is referred to as a "Sequence Listing XML" in order to distinguish it from a "Sequence Listing" submitted in an application having a filing date BEFORE July 1, 2022. For 35 U.S.C. 111 applications and international applications having a filing date before July 1, 2022, that disclose of nucleotide and/or amino acid sequences, see 37 CFR 1.821(c) and 1.821(e)(1). See also MPEP §§ 2421.01 and 2421.02.

2412.02(a) "Enumeration of its residues" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

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

*****

(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).

*****

WIPO Standard ST.26 specifies that "enumeration of its residues" means "disclosure of a sequence in a patent application by listing, in order, each residue of the sequence, wherein either (i) the residue is represented by a name, abbreviation, symbol, or structure (e.g., HHHHHHQ or HisHisHisHisHisHisGln); or (ii) multiple residues are represented by a shorthand formula (e.g., His6Gln) (WIPO Standard ST.26, paragraph 3(c)). See also the Guidance portion of the Standard, Annex VI, page 3.26.vi.2 for a discussion of "enumeration of its residues" and page 3.26.vi.18-19, 51, 53 for examples.

2412.03 Nucleotides and Amino Acids Included and Excluded From a "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

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

*****

(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).

*****

(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.

Generally, nucleotide sequences that are an unbranched sequence or constitute a linear portion of a branched sequence of 10 or more specifically defined nucleotides are required to be listed in a "Sequence Listing XML." See MPEP §§ 2412.03(d) and (e) for definitions of nucleotide and modified nucleotide, respectively. See MPEP § 2412.03(a) for definition of "specifically defined."

Similarly, amino acid sequences that are an unbranched sequence or constitute a linear portion of a branched sequence of 4 or more specifically defined amino acids are required to be listed in a "Sequence Listing XML." See MPEP §§ 2412.03(b) and (c) for definitions of amino acid and modified amino acid, respectively. See MPEP § 2412.03(a) for definition of "specifically defined."

37 CFR 1.831(b) sets forth the nucleotide and amino acid sequences which must be included in a "Sequence Listing XML". 37 CFR 1.831(j) specifies that any sequence having fewer than 10 specifically defined nucleotides, or fewer than 4 specifically defined amino acids must be excluded from any "Sequence listing XML."

2412.03(a) "Specifically Defined" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

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

*****

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

*****

WIPO Standard ST.26, paragraph 3(k), provides that "specifically defined" means any nucleotide other than those represented by the symbol "n" and any amino acid other than those represented by the symbol "X,", wherein "n" and "X" are used in a conventional manner as shown below in Table 1 for nucleotide symbols and Table 3 for amino acids symbols.

Table 1: List of Nucleotides Symbols

Symbol	Definition
a	adenine
c	cytosine
g	guanine
t	thymine in DNA/uracil in RNA (t/u)
m	a or c
r	a or g
w	a or t/u
s	c or g
y	c or t/u
k	g or t/u
v	a or c or g; not t/u
h	a or c or t/u; not g
d	a or g or t/u; not c
b	c or g or t/u; not a
n	a or c or g or t/u; "unknown" or "other"

Reproduced from WIPO Standard ST.26, Annex I, Section 1.

Table 3: List of Amino Acids Symbols

Symbol	Definition
A	Alanine
R	Arginine
N	Asparagine

Symbol	Definition
D	Aspartic acid (Aspartate)
C	Cysteine
Q	Glutamine
E	Glutamic acid (Glutamate)
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
O	Pyrrolysine
S	Serine
U	Selenocysteine
T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine
B	Aspartic acid or Asparagine
Z	Glutamine or Glutamic acid
J	Leucine or Isoleucine
X	A or R or N or D or C or Q or E or G or H or I or L or K or M or F or P or O or S or U or T or W or Y or V; “unknown” or “other”

Reproduced from WIPO Standard ST.26, Annex I, Section 3.

2412.03(b) “Amino Acid” [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

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

*****

(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.

*****

WIPO Standard ST.26, paragraph 3(a), defines “amino acid” to mean any amino acid that can be represented using any of the symbols shown in Annex I, Table 3: List of Amino Acids Symbols of WIPO Standard ST.26 (reproduced in MPEP § 2412.03(a)). Such amino acids include, *inter alia*, D-amino acids and amino acids containing modified or synthetic side chains. Amino acids will be construed as unmodified L-amino acids unless further described in a “feature table”. See MPEP § 2413.01(g), subsection I, for discussion of a “feature table”. A peptide nucleic acid (PNA) residue is not considered an amino acid, but is considered a nucleotide.

2412.03(c) “Modified Amino Acid” [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**.]*

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

*****

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

*****

WIPO Standard ST.26, paragraph 3(e), defines “modified amino acid” to mean any amino acid as described in the definition of “amino acid”, other than L-alanine, L-arginine, L-asparagine, L-aspartic acid, L-cysteine, L-glutamine, L-glutamic acid, L-glycine, L-histidine, L-isoleucine, L-leucine, L-lysine, L-methionine, L-phenylalanine, L-proline, L-pyrrolysine, L-serine, L-selenocysteine, L-threonine, L-tryptophan, L-tyrosine, or L-valine. See MPEP § 2412.03(b).

Modified amino acids, including D-amino acids, should be represented in the sequence as the corresponding unmodified amino acids whenever possible. Any modified amino acid in a sequence that cannot otherwise be represented by a symbol set forth in Table 3 must be represented by "X", i.e., an "other" amino acid.

A modified amino acid must be further described in a feature table. See MPEP § 2413.01(g), subsection I, for discussion of a "feature table". Where applicable, the feature keys "CARBOHYD" or "LIPID" should be used together with the qualifier "note". The feature key "MOD_RES" should be used for other post-translationally modified amino acids together with the qualifier "note". The feature key "SITE" together with the qualifier "note" should be used when the modified amino acid is not a post-translationally modified amino acid. See MPEP § 2413.01(g), subsections II and III, for discussion of a "feature key". The value for the qualifier "note" must either be an abbreviation set forth in Table 4 below or the complete, unabbreviated name of the modified amino acid. The abbreviations set forth in Table 4 or the complete, unabbreviated names must not be used in the sequence itself.

Table 4: List of Modified Amino Acids

Abbreviation	Modified Amino acid
Aad	2-Aminoadipic acid
bAad	3-Aminoadipic acid
bAla	beta-Alanine, beta-Aminopropionic acid
Abu	2-Aminobutyric acid
4Abu	4-Aminobutyric acid, piperidinic acid
Acp	6-Aminocaproic acid
Ahe	2-Aminoheptanoic acid
Aib	2-Aminoisobutyric acid
bAib	3-Aminoisobutyric acid
Apm	2-Aminopimelic acid
Dbu	2,4 Diaminobutyric acid
Des	Desmosine
Dpm	2,2'-Diaminopimelic acid
Dpr	2,3-Diaminopropionic acid
EtGly	N-Ethylglycine
EtAsn	N-Ethylasparagine
Hyl	Hydroxylysine
aHyl	allo-Hydroxylysine
3Hyp	3-Hydroxyproline
4Hyp	4-Hydroxyproline
Ide	Isodesmosine
alle	allo-Isoleucine
MeGly	N-Methylglycine, sarcosine
Melle	N-Methylisoleucine
MeLys	6-N-Methyllysine
MeVal	N-Methylvaline
Nva	Norvaline
Nle	Norleucine
Orn	Ornithine

Reproduced from WIPO Standard ST.26, Annex I, Section 4.

2412.03(d) "Nucleotide" [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/ standards/en/pdf/03-26-01.pdf**.]*

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

*****

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

*****

WIPO Standard ST.26, paragraphs 3(f) and (g), identify a "nucleotide" to encompass any nucleotide or nucleotide analogue or "modified nucleotide" (see MPEP § 2412.03(e)) that can be represented using any of the symbols set forth in Table 1: List of Nucleotides Symbols (see MPEP § 2412.03(a)), wherein the nucleotide, or nucleotide analogue, or modified nucleotide contains:

(i) a backbone moiety selected from:

(1) 2’ deoxyribose 5’ monophosphate (the backbone moiety of a deoxyribonucleotide) or ribose 5’ monophosphate (the backbone moiety of a ribonucleotide); or

(2) an analogue of a 2’ deoxyribose 5’ monophosphate or ribose 5’ monophosphate, which when forming the backbone of a nucleic acid analogue, results in an arrangement of nucleobases that mimics the arrangement of nucleobases in nucleic acids containing a 2’ deoxyribose 5’ monophosphate or ribose 5’ monophosphate backbone, wherein the nucleic acid analogue is capable of base pairing with a complementary nucleic acid; examples of backbone moieties include amino acids as in peptide nucleic acids, glycol molecules as in glycol nucleic acids, threofuranosyl sugar molecules as in threose nucleic acids, morpholine rings and phosphorodiamidate groups as in morpholinos, and cyclohexenyl molecules as in cyclohexenyl nucleic acids; and

(ii) the backbone moiety is either:

(1) joined to a nucleobase, including a modified or synthetic purine or pyrimidine nucleobase; or

(2) lacking a purine or pyrimidine nucleobase when the nucleotide is part of a nucleotide sequence, referred to as an “AP site” or an “abasic site”.

2412.03(e) “Modified Nucleotide” [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int /export/sites/www/ standards/en/pdf/03-26-01.pdf**.]*

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

*****

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

*****

WIPO Standard ST.26, paragraph 3(f), provides that a “modified nucleotide” means any “nucleotide” as explained in MPEP § 2412.03(d) other than deoxyadenosine 3’-monophosphate, deoxyguanosine 3’-monophosphate, deoxycytidine 3’-monophosphate, deoxythymidine 3’-monophosphate, adenosine 3’-monophosphate, guanosine 3’-monophosphate, cytidine 3’-monophosphate, or uridine 3’-monophosphate.

2412.04 Use of Sequence Identifiers to Denote Sequences Disclosed in the Description, Drawings or Claims [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int /export/sites/www/ standards/en/pdf/03-26-01.pdf**.]*

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

*****

(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.

*****

37 CFR 1.831(c) requires that each nucleotide and/or amino acid sequence set forth in a “Sequence Listing XML” in accordance with 37 CFR § 1.831(a) must be referenced by a sequence identifier as defined in 37 CFR 1.832(a) (see MPEP § 2412.05(a)), preceded by the notation “SEQ ID NO:” or the like, when the sequence appears in the description or claims. Additionally, where a sequence set forth in a “Sequence Listing XML” is presented in a drawing,

reference must be made using the sequence identifier preceded by the notation "SEQ ID NO:" or the like, either in the drawing or in the Brief Description of the Drawings. The sequence identifiers in the disclosure must correspond to sequence identifiers set forth in the "Sequence Listing XML" as defined in 37 CFR 1.832(a).

37 CFR 1.831(c) is also intended to permit references in the application (e.g., specification, claims, or drawings) to sequences set forth in the "Sequence Listing XML" by the use of assigned sequence identifiers without enumerating the sequence. Sequence identifiers can also be used to discuss and/or claim parts or fragments of a properly presented sequence. For example, language such as "residues 14 to 243 of SEQ ID NO:23" is permissible and the fragment need not be separately presented in the "Sequence Listing XML."

37 CFR 1.831(c) does not alter, in any way, the requirements of 35 U.S.C. 112. The implementation of this rule has had no effect on disclosure and/or claiming requirements. 37 CFR 1.831 - 1.835, in general, or the use of sequence identifiers throughout the specification and claims, specifically, should not raise any issues under 35 U.S.C. 112(a) or 35 U.S.C. 112(b) because the of sequence identifiers (preceded by "SEQ ID NO: or the like") only provides a shorthand way for applicants to discuss and claim their inventions. These identifiers do not in any way restrict the manner in which an invention can be claimed.

2412.05 Representation and Symbols for Nucleotide and/or Amino Acid Sequences [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

WIPO Standard ST.26 sets forth specific symbols for representing nucleotide and/or amino acid residues in a sequence. The USPTO rules incorporate those specific symbols and XML formatting representations.

2412.05(a) Use of Sequentially Numbered Sequence Identifiers in the "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf.]

37 CFR 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).

*****

In accordance with 37 CFR 1.832(a), the sequence identifiers in the "Sequence Listing XML" must begin with 1 and increase sequentially by integers. The requirement for sequence identifiers, at a minimum, requires that each sequence be assigned a different number for purposes of identification. However, where practical and for ease of reference, sequences should be presented in the "Sequence Listing XML" in numerical order and in the order in which they are discussed in the application.

Each nucleotide and/or amino acid sequence that meets the definition in 37 CFR 1.831(b) and is enumerated by its residues must be assigned a separate sequence identifier, including a sequence which is identical to a region of a longer sequence. See MPEP § 2412.02 for further description of a "sequence".

Where no sequence is present for a sequence identifier, i.e. an intentionally skipped sequence, "000" must be used in place of a sequence in a "Sequence Listing XML". The total number of sequences indicated in the "Sequence Listing XML"

must equal the total number of sequence identifiers, whether followed by a sequence or by “000”.

WIPO Standard ST.26 paragraphs 58 and 59 require that an intentionally skipped sequences in a “Sequence Listing XML” must be represented as follows:

(a) the value of the element SequenceData and its attribute sequenceIDNumber, is the sequence identifier of the skipped sequence;

(b) no value is provided for the elements INSDSeq _length, INSDSeq _moltype, and INSDSeq _division;

(c) the element INSDSeq _feature-table is not included; and

(d) the value of the element INSDSeq _sequence is the string “000”.

2412.05(b) Representation and Symbols of Nucleotide Sequence Data [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**.]*

37 CFR 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.

*****

(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.

*****

I. REPRESENTATION OF NUCLEOTIDE SEQUENCE

WIPO Standard ST.26, paragraph 11, provides that a nucleotide sequence must be represented only by a single strand, in the 5’ to 3’ direction from left to right, or in the direction from left to right that mimics the 5’ to 3’ direction. The designations 5’ and 3’ or any other similar designations must not be included in the sequence. A double-stranded nucleotide sequence disclosed by enumeration of the residues of both strands must be represented as:

(a) a single sequence or as two separate sequences, each assigned its own sequence identifier, where the two separate strands are fully complementary to each other, or

(b) two separate sequences, each assigned its own sequence identifier, where the two strands are not fully complementary to each other.

WIPO Standard ST.26, paragraph 12, provides that the first nucleotide presented in the sequence is residue position number 1. When nucleotide sequences are circular in configuration, applicant must choose the nucleotide in residue position number 1. Numbering is continuous throughout the entire sequence in the 5’ to 3’ direction, or in the direction that mimics the 5’ to 3’ direction. The last residue position number must equal the number of nucleotides in the sequence.

II. SYMBOLS FOR A NUCLEOTIDE SEQUENCE

WIPO Standard ST.26, paragraph 13, provides that all nucleotides in a sequence must be represented using the symbols as listed in Table 1: List of Nucleotides Symbols (see MPEP § 2412.03(a)). Only lower-case letters must be used. Any symbol used to represent a nucleotide is the equivalent of only one residue.

WIPO Standard ST.26, paragraph 14, sets forth that the symbol “t” will be construed as thymine in deoxyribonucleic acid (DNA) and uracil in ribonucleic acid (RNA). Uracil in DNA or thymine in RNA is considered a modified nucleotide and

must be further described in a feature table. See MPEP § 2413.01(g), subsection I for more detail regarding a "feature table."

WIPO Standard ST.26, paragraph 15, provides that where an ambiguity symbol (representing two or more alternative nucleotides) is appropriate, the most restrictive symbol should be used, as listed in Table 1: List of Nucleotides Symbols (see MPEP § 2412.03(a)). For example, if a nucleotide in a given position could be "a" or "g", then "r" should be used, rather than "n". The symbol "n" will be construed as any one of "a", "c", "g", or "t/u" except where it is used with a further description in a feature table. The symbol "n" must not be used to represent anything other than a nucleotide. A single modified or "unknown" nucleotide may be represented by the symbol "n", together with a further description in a feature table. See MPEP § 2413.01(g), subsection I, for more detail regarding a "feature table." For representation of sequence variants, i.e., alternatives, deletions, insertions or substitutions relative to a primary sequence, see MPEP § 2412.05(c); and also MPEP § 2413.01(g), subsection XII for information on variants.

WIPO Standard ST.26, paragraph 16, sets forth that modified nucleotides should be represented in the sequence as the corresponding unmodified nucleotides, i.e., "a", "c", "g" or "t" whenever possible. Any modified nucleotide in a sequence that cannot otherwise be represented by any other symbol in Table 1: List of Nucleotides Symbols (see MPEP § 2412.03(a)), i.e., an "other" nucleotide, such as a non-naturally occurring nucleotide, must be represented by the symbol "n". The symbol "n" is the equivalent of only one residue.

WIPO Standard ST.26, paragraph 19, specifies that uracil in DNA or thymine in RNA are considered modified nucleotides and must be represented in the sequence as "t" and be further described in a feature table using the feature key "modified_base", the qualifier "mod_base" with "OTHER" as the qualifier value and the qualifier "note" with "uracil" or "thymine", respectively, as the qualifier value. See MPEP § 2413.01(g), subsection I for more detail regarding a "feature table."

WIPO Standard ST.26, paragraph 21, provides that any "unknown" nucleotide must be represented by the symbol "n" in the sequence. An "unknown" nucleotide should be further described in a feature table using the feature key "unsure". The symbol "n" is the equivalent of only one residue. See MPEP § 2413.01(g), subsection I, for more detail regarding a "feature table."

III. DESCRIPTION OF MODIFIED NUCLEOTIDES WITHIN A NUCLEOTIDE SEQUENCE

WIPO Standard ST.26, paragraph 17, specifies that a modified nucleotide must be further described in a feature table (see MPEP § 2413.01(g), subsection I, for more detail regarding a "feature table") using the feature key "modified_base" and the mandatory qualifier "mod_base" in conjunction with a single abbreviation from Table 2: List of Modified Nucleotides, below, as the qualifier value. See MPEP § 2413.01(g) subsections II and III, for more information regarding use of a feature key; and MPEP § 2413.01(g) subsections V and VI, for more information regarding use of a qualifier. If the abbreviation is "OTHER", the complete unabbreviated name of the modified nucleotide must be provided as the value in a "note" qualifier. For a listing of alternative modified nucleotides, the qualifier value "OTHER" may be used in conjunction with a further "note" qualifier. The abbreviations (or full names) provided in Table 2 must not be used in the sequence itself.

WIPO Standard ST.26, paragraph 18, describes that a nucleotide sequence including one or more regions of consecutive modified nucleotides that share the same backbone moiety must be further described in a feature table as required for a modified nucleotide. See MPEP § 2413.01(g), subsection I, for information regarding a feature table and MPEP § 2412.03(e) regarding modified nucleotides. The modified nucleotides of each such region may be jointly described in a single INSDFeature element of a "feature table" as described below. See MPEP § 2413.01(g), subsection I, for information regarding INSDFeature elements of a feature table. The most restrictive unabbreviated chemical name that encompasses all of the modified nucleotides in the range or a list of the chemical names of all the nucleotides in the range must be provided as the

value in the “note” qualifier. For example, a glycol nucleic acid sequence containing “a”, “c”, “g”, or “t” nucleobases may be described in the “note” qualifier as “2,3-dihydroxypropyl nucleosides.” Alternatively, the same sequence may be described in the “note” qualifier as “2,3-dihydroxypropyladenine, 2,3-dihydroxypropylthymine, 2,3-dihydroxypropylguanine, or 2,3-dihydroxypropylcytosine.” Where an individual modified nucleotide in the region includes an additional modification, then the modified nucleotide must also be further described in the feature table as required for a modified nucleotide.

WIPO Standard ST.26, paragraph 19, provides that uracil in DNA or thymine in RNA are considered modified nucleotides and must be represented in the sequence as “t” and be further described in a feature table using the feature key “modified_base”, the qualifier “mod_base” with “OTHER” as the qualifier value and the qualifier “note” with “uracil” or “thymine”, respectively, as the qualifier value.See MPEP § 2413.01(g), subsections II and III, for more information regarding use of a feature key; and MPEP § 2413.01(g), subsections V and VI, for more information regarding use of a qualifier.

Table 2: List of Modified Nucleotides

Abbreviation	Definition
ac4c	4-acetylcytidine
chm5u	5-(carboxyhydroxymethyl)uridine
cm	2'-O-methylcytidine
cmnm5s2u	5-carboxymethylaminomethyl-2-thiouridine
cmnm5u	5-carboxymethylaminomethyluridine
dhu	dihydrouridine
fm	2'-O-methylpseudouridine
gal q	beta, D-galactosylqueuosine
gm	2'-O-methylguanosine
i	inosine
i6a	N6-isopentenyladenosine
m1a	1-methyladenosine
m1f	1-methylpseudouridine
m1g	1-methylguanosine
m1i	1-methylinosine
m22g	2,2-dimethylguanosine
m2a	2-methyladenosine
m2g	2-methylguanosine
m3c	3-methylcytidine
m4c	N4-methylcytosine
m5c	5-methylcytidine
m6a	N6-methyladenosine
m7g	7-methylguanosine
mam5u	5-methylaminomethyluridine
mam5s2u	5-methoxyaminomethyl-2-thiouridine
man q	beta, D-mannosylqueuosine
mcm5s2u	5-methoxycarbonylmethyl-2-thiouridine
mcm5u	5-methoxycarbonylmethyluridine
mo5u	5-methoxyuridine
ms2i6a	2-methylthio-N6-isopentenyladenosine
ms2t6a	N-((9-beta-D-ribofuranosyl-2-methylthiopurine-6-yl)carbamoyl)threonine
mt6a	N-((9-beta-D-ribofuranosylpurine-6-yl)N-methylcarbamoyl)threonine
mv	uridine-5-oxyacetic acid-methylester
o5u	uridine-5-oxyacetic acid
osyw	wybutoxosine
p	pseudouridine
q	queuosine
s2c	2-thiocytidine
s2t	5-methyl-2-thiouridine
s2u	2-thiouridine
s4u	4-thiouridine
m5u	5-methyluridine
t6a	N-((9-beta-D-ribofuranosylpurine-6-yl)-carbamoyl)threonine
tm	2'-O-methyl-5-methyluridine
um	2'-O-methyluridine
yw	wybutosine
x	3-(3-amino-3-carboxy-propyl)uridine, (acp3)u
OTHER	(requires note qualifier)

Reproduced from WIPO Standard ST. 26, Annex I, Section 2.

IV. JOINTLY DESCRIBING A REGION OF A NUCLEOTIDE SEQUENCE

WIPO Standard ST.26, paragraph 22, specifies that 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 using a single INSDFeature element with the syntax "x..y" as the location descriptor in the element INSDFeature_location. See MPEP § 2413.01(g) subsection IV, for information regarding INSDFeature_location. For representation of sequence variants, i.e., alternatives, deletions, insertions or substitutions, see MPEP § 2412.05(c) and MPEP § 2413.01(g), subsection XI.

2412.05(c) Representation and Inclusion of Variants [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**.]*

A primary sequence and any variant of that sequence, each disclosed by enumeration of its residues and meeting the definition in 37 CFR 1.831(a) and 1.831(b), must each be included in the "Sequence Listing XML" and assigned its own sequence identifier. Where a variant sequence is disclosed as a single sequence with enumerated alternative residues at one or more positions, it must be included in the "Sequence Listing XML" and should be represented by a single sequence, wherein the enumerated alternative residues are represented by the most restrictive ambiguity symbol. Any variant sequence, disclosed only by reference to deletion(s), insertion(s), or substitution(s) in a primary sequence, should be included in the "Sequence Listing XML". The table below indicates the proper use of feature keys and qualifiers for nucleic acid and amino acid sequence variants:

List of Feature Keys and Qualifiers

Type of sequence	Feature Key	Qualifier	Use
Nucleic acid	variation	replace or note	Naturally occurring mutations and polymorphisms, e.g., alleles, RFLPs.
Nucleic acid	misc_difference	replace or note	Variability introduced artificially, e.g., by genetic manipulation or by chemical synthesis.
Amino acid	VAR_SEQ	note	Variant produced by alternative splicing, alternative promoter usage, alternative initiation and ribosomal frameshifting.
Amino acid	VARIANT	note	Any type of variant for which VAR_SEQ is not applicable.

Reproduced from paragraph 96 of WIPO Standard ST.26.

For additional information about feature keys and qualifiers, see MPEP § 2413.01(g), subsections II, III and V.

For additional information about the representation of sequence variants in a "Sequence Listing XML," see MPEP § 2413.01(g), subsection XI.

2412.05(d) Representation and Symbols of Amino Acid Sequence Data [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase

applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf.]

37 CFR 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.

*****

(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.

*****

I. REPRESENTATION OF AN AMINO ACID SEQUENCE

WIPO Standard ST.26, paragraph 24, specifies that the amino acids in an amino acid sequence must be represented in the amino to carboxy direction from left to right. The amino and carboxy groups must not be represented in the sequence.

WIPO Standard ST.26, paragraph 25, indicates that the first amino acid in the sequence is residue position number 1, including amino acids preceding the mature protein, for example, pre-sequences, pro-sequences, pre-pro-sequences and signal sequences. When an amino acid sequence is circular in configuration and the ring consists solely of amino acid residues linked by peptide bonds, i.e., the sequence has no amino and carboxy termini, applicant must choose the amino acid in residue position number 1. Numbering is continuous through the entire sequence in the amino to carboxy direction.

II. SYMBOLS FOR AN AMINO ACID SEQUENCE

WIPO Standard ST.26, paragraph 26, specifies that all amino acids in a sequence must be represented using the symbols set forth in Table 3: List of Amino Acids Symbols, in MPEP § 2412.03(a) above. Only uppercase letters must be used. Any symbol used to represent an amino acid is the equivalent of only one residue.

WIPO Standard ST.26, paragraph 27, indicates that where an ambiguity symbol (representing two or more amino acids in the alternative) is appropriate, the most restrictive symbol should be used, as listed in Table 3: List of Amino Acids Symbols (MPEP § 2412.03(a)). For example, if an amino acid in a given position could be aspartic acid or asparagine, the symbol "B" should be used, rather than "X". The symbol "X" will be construed as any one of "A", "R", "N", "D", "C", "Q", "E", "G", "H", "I", "L", "K", "M", "F", "P", "O", "S", "U", "T", "W", "Y", or "V", except where it is used with a further description in the feature table. The symbol "X" must not be used to represent anything other than an amino acid. A single modified or "unknown" amino acid may be represented by the symbol "X", together with a further description in a feature table (see MPEP § 2413.01(g), subsection I or MPEP § 2412.03(c), for more detail regarding a "feature table"). For representation and inclusion of sequence variants, see MPEP § 2412.05(c). For details of how to represent variants in a "Sequence Listing XML," see MPEP § 2413.01(g), subsection XI.

WIPO Standard ST.26, paragraph 28, specifies that disclosed amino acid sequences separated by internal terminator symbols, represented for example by "Ter" or asterisk "*" or period "." or a blank space, must be included as separate sequences for each enumerated amino acid sequence that contains at least four specifically defined amino acids and is encompassed by the description of sequences found in MPEP § 2412.03. Each such separate sequence must be assigned its own sequence identifier (see MPEP § 2412.05(a)). Terminator symbols and spaces must not be included in a sequence contained in a "Sequence Listing XML".

Any "unknown" amino acid must be represented by the symbol "X" in the sequence. An "unknown" amino acid designated as "X" must be further described in a feature table (see MPEP § 2413.01(g), subsection I, for more detail regarding a "feature table") using the feature key "UNSURE" and optionally the qualifier "note." The symbol "X" is the equivalent of only one residue (WIPO Standard ST.26, paragraph 32).

III. DESCRIPTION OF MODIFIED AMINO ACIDS WITHIN AN AMINO ACID SEQUENCE

WIPO Standard ST.26, paragraph 29, specifies that modified amino acids, including D-amino acids, should be represented in the sequence as the corresponding unmodified amino acids whenever possible. Any modified amino acid in a sequence that cannot otherwise be represented by any other symbol in Table 3: List of Amino Acids Symbols (see MPEP § 2412.03(a)), i.e., an "other" amino acid, must be represented by "X". The symbol "X" is the equivalent of only one residue.

WIPO Standard ST.26, paragraph 30, provides that a modified amino acid must be further described in a feature table (see MPEP § 2413.01(g), subsection I, for more detail regarding a "feature table"). Where applicable, the feature keys "CARBOHYD" or "LIPID" should be used together with the qualifier "note". The feature key "MOD_RES" should be used for other post-translationally modified amino acids together with the qualifier "note". The feature key "SITE" together with the qualifier "note" should be used when the modified amino acid is not a post-translationally modified amino acid. The value for the qualifier "note" must either be an abbreviation set forth in Table 4: List of Modified Amino Acids (see MPEP § 2412.03(c)), above, or the complete, unabbreviated name of the modified amino acid. The abbreviations set forth in Table 4, or the complete, unabbreviated names must not be used in the sequence itself.

IV. JOINTLY DESCRIBING A REGION OF AN AMINO ACID SEQUENCE

WIPO Standard ST.26, paragraph 34, provides that a region containing a known number of contiguous "X" residues for which the same description applies may be jointly described in one feature key using the syntax "x..y" as the location descriptor in the element INSDFeature_location (see MPEP § 2413.01(g) subsections II-III for information regarding "feature keys" and subsection IV, for information regarding INSDFeature_location). For representation and inclusion of sequence variants, see MPEP § 2412.05(c). For details of how to represent variants in a "Sequence Listing XML," see MPEP § 2413.01(g), subsection XII.

2412.05(e) Presentation of Special Situations [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b)].

37 CFR 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.

*****

(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.

WIPO Standard ST.26, paragraph 35, describes that a sequence disclosed by enumeration of its residues that is constructed as a single continuous sequence from one or more non-contiguous segments of a larger sequence or of segments from different sequences must be included in the "Sequence Listing XML" and assigned its own sequence identifier, as defined in 37 CFR 1.832(a). See MPEP § 2412.05(a).

WIPO Standard ST.26, paragraph 36, describes that a 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 included in the "Sequence Listing XML" as one sequence and assigned its own sequence identifier.

WIPO Standard ST.26, paragraph 37, describes that a sequence that contains regions of specifically defined residues separated by one or more gaps of an unknown or undisclosed number of residues must not be represented in the "Sequence Listing XML" as a single sequence. Each region of specifically defined residues (as encompassed by the definition in 37 CFR 1.831(a)) must be included in the "Sequence Listing XML" as a separate sequence and assigned its own sequence identifier.

2412.06 The Requirement for Exclusive Conformance; Sequences Presented in Drawing Figures [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

For all applications that disclose a nucleotide sequence and/or amino acid sequence(s) by enumeration of its residues, as defined in 37 CFR 1.831(b), 37 CFR 1.831(a) requires conformance to the requirements of 37 CFR 1.832 through 37 CFR 1.834 with regard to the manner in which the disclosed nucleotide and/or amino acid sequences are presented and described in the "Sequence Listing XML." This requirement is necessary to minimize any confusion that could result if more than one format for representing sequence data was employed in a given application.

Pursuant to 37 CFR 1.83(a), sequences that are included in the "Sequence Listing XML" should not be duplicated in the drawings. With the use of feature keys and qualifiers in a "Sequence Listing XML" to represent and describe features of a nucleotide or amino acid sequence, the need to re-present a sequence in a drawing is less critical. However, in certain instances, a significant sequence characteristic may not be readily conveyed by the sequence-associated data of the "Sequence Listing XML" and may need to be depicted in a figure. For example, in view of the fact that the representation of double stranded nucleic acids is not permitted in the "Sequence Listing XML," many significant nucleic acid features, such as "sticky ends" and the like, may only be shown effectively by reference to a drawing figure. Further, the similarity or homology between/among sequences may only be depicted in an effective manner in a drawing figure. Similarly, drawing figures are recommended for use with amino acid sequences to depict structural features of the corresponding protein, such as epitopes and interaction domains. The situations discussed herein are given by way of example only and there may be many other reasons for including a sequence in a drawing. However, when an enumerated sequence is presented in a drawing, the sequence must still be included in the "Sequence Listing XML" if the sequence falls within the definition set forth in 37 CFR 1.831(b), and a sequence identifier (preceded by "SEQ ID NO:X" or the like) must be used, either in the drawing itself or in the Brief Description of the Drawings.

2412.07 Examination of Patent Applications Claiming Large Numbers of Nucleotide Sequences [R-07.2022]

Content regarding the examination of patent applications claiming large numbers of nucleotide sequences is located in MPEP § 2434.

2413 Content of a "Sequence Listing XML" and Form and Format of the "Sequence Listing XML" File [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

In order to ensure consistency and uniformity of a "Sequence Listing XML," 37 CFR 1.833 defines the

content of the "Sequence Listing XML." 37 CFR 1.834 details the specifics on the technical form and format of the .xml file containing the "Sequence Listing XML" XML file.

2413.01 Parts of the "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

The constituent parts of the "Sequence Listing XML" are identified in multiple paragraphs of WIPO Standard ST.26 and have been specifically incorporated into the USPTO rules of practice at 37 CFR 1.839.

2413.01(a) The "Sequence Listing XML" is a Single File Encoded Using Unicode UTF-8 [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf.]

37 CFR 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).

*****

According to 37 CFR 1.833, the entire "Sequence Listing XML" must be presented as a single file. WIPO Standard ST.26 specifies that the file must be encoded using Unicode UTF-8, with the following restrictions:

(1) the information contained in the elements ApplicantName, InventorName and InventionTitle of the general information part, and the NonEnglishQualifier_value of the sequence data part, may be composed of any valid Unicode characters indicated in the XML 1.0 specification except the Unicode Control code points 0000-001F and 007F-009F. The reserved characters ", &, ', <, and > (Unicode code points 0022, 0026, 0027, 003C and 003E respectively), must be replaced as set forth the table below; and

(2) the information contained in all other elements and attributes of the general information part and in all other elements and attributes of the sequence data part must be composed of printable characters (including the space character) from the Unicode Basic Latin code table (i.e., limited to Unicode code points 0020 through 007E). The reserved characters ", &, ', <, and > (Unicode code points 0022, 0026, 0027, 003C and 003E respectively), must be replaced as set forth in the table below (WIPO Standard ST.26, paragraph 40).See MPEP § 2413.01(f) for details about the "general information part" and MPEP § 2413.01(g) for details about the "sequence data part."

WIPO Standard ST.26 specifies that in an XML instance of a "Sequence Listing XML", numeric character references must not be used and the following reserved characters must be replaced by the corresponding predefined entities when used in a value of an attribute or content of an element:

List of Reserved Characters and Predefined Entities

Reserved Character	Predefined Entities
<	<
>	>
&	&
"	"
'	'

Reproduced from WIPO Standard ST.26, paragraph 41. See also WIPO Standard ST.26, paragraph 41, footnote 1 of the WIPO standard for details about "numeric character references."

The only character entity references permitted are the predefined entities set forth above (WIPO Standard ST.26, paragraph 41).

2413.01(b) The "Sequence Listing XML" Must Be Valid According To the DTD [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

37 CFR 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.

*****

(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.

*****

37 CFR 1.833(b)(1) incorporates the DTD requirement of WIPO Standard ST.26 such that a "Sequence Listing XML" must conform to the Document Type Definition (the document that sets out the structure and use of elements and attributes in an XML compliant document) presented in Annex II of WIPO Standard ST.26. Use of the WIPO Sequence tool will enable a user to generate a "Sequence Listing XML" that conforms to the DTD in Annex II of WIPO Standard ST.26. See MPEP § 2418.

2413.01(c) The "Sequence Listing XML" Must Contain an XML Declaration [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf.]

37 CFR 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.

*****

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

*****

(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;

*****

WIPO Standard ST.26, paragraph 39(a), specifies that the first line of the XML instance must contain the XML declaration, which is:

<?xml version="1.0" encoding="UTF-8"?>

2413.01(d) The "Sequence Listing XML" Must Contain a DOCTYPE Declaration [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf.]

37 CFR 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.

*****

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

*****

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

*****

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

*****

WIPO Standard ST.26, paragraph 39(b), specifies that the second line of the XML instance must contain the document type (DOCTYPE) declaration, which is:

<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">

2413.01(e) The "Sequence Listing XML" Must Contain a Root Element [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf.]

37 CFR 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.

*****

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

*****

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

*****

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

*****

WIPO Standard ST.26, paragraph 43, prescribes the root element of an XML instance to have the following attributes:

List of File Attributes

Attribute	Description	Mandatory/Optional
dtdVersion	Version of the DTD used to create this file in the format "V#_#", e.g., "V1_3".	Mandatory
fileName	Name of the sequence listing file.	Optional
softwareName	Name of the software that generated this file.	Optional
softwareVersion	Version of the software that generated this file.	Optional
productionDate	Date of production of the sequence listing file (format "CCYY-MM-DD").	Optional
originalFreeTextLanguageCode	The language code (see reference in paragraph 9 to ISO 639-1:2002) for the single original language in which the language-dependent free text qualifiers were prepared.	Optional
nonEnglishFreeTextLanguageCode	The language code (see reference in paragraph 9 to ISO 63901:2002) for the NonEnglishQualifier_value elements.	Mandatory when a NonEnglishQualifier_value element is present in the sequence listing

Reproduced from paragraph 43 of WIPO Standard ST.26.

2413.01(f) The "Sequence Listing XML" Must Contain a General Information Part [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**.]*

37 CFR 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.

*****

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

*****

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

*****

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

*****

WIPO Standard ST.26 prescribes that the general information part of the sequence listing must contain the following bibliographic elements:

List of Patent Application Information Elements

Element	Description	Mandatory/Optional
ApplicationIdentification The ApplicationIdentification is composed of: IPOfficeCode ApplicationNumberText	The application identification for which the sequence listing is submitted ST.3 Code of the office of filing The application number as provided by the office of filing (e.g., PCT/IB2013/099999)	Mandatory when a sequence listing is furnished at any time following the assignment of the application number Mandatory Mandatory
FilingDate	The date of filing of the patent application for which the sequence listing is submitted (ST.2 format "CCYY-MM-DD", using a 4-digit calendar year, a 2-digit calendar month and a 2-digit day within the calendar month, e.g., 2015-01-31).	Mandatory when a sequence listing is furnished at any time following the assignment of a filing date
ApplicantFileReference	A single unique identifier assigned by applicant to identify a particular application, typed in the characters as set forth in paragraph 40 (b).	Mandatory when a sequence listing is furnished at any time prior to assignment of the application number; otherwise, Optional
EarliestPriorityApplicationIdentification	The identification of the earliest priority application (also contains IPOfficeCode, ApplicationNumberText and FilingDate, see ApplicationIdentification above).	Mandatory where priority is claimed
ApplicantName	Name of the first mentioned applicant typed in the characters as set forth in paragraph 40 (a).	Mandatory

Element	Description	Mandatory/Optional
	This element includes the mandatory attribute languageCode as set forth in paragraph 47.	
ApplicantNameLatin	Where ApplicantName is typed in characters other than those as set forth in paragraph 40 (b), a translation or transliteration of the name of the first mentioned applicant must also be typed in characters as set forth in paragraph 40 (b).	Mandatory where ApplicantName contains non-Latin characters
InventorName	Name of the first mentioned inventor typed in the characters as set forth in paragraph 40 (a). This element includes the mandatory attribute languageCode as set forth in paragraph 47.	Optional
InventorNameLatin	Where InventorName is typed in characters other than those as set forth in paragraph 40 (b), a translation or transliteration of the first mentioned inventor may also be typed in characters as set forth in paragraph 40 (b).	Optional
InventionTitle	Title of the invention typed in the characters as set forth in paragraph 40 (a) in the language of filing. A translation of the title of the invention into additional languages may be typed in the characters as set forth in paragraph 40 (a) using additional InventionTitle elements. This element includes the mandatory attribute languageCode as set forth in paragraph 48. The title of invention should be between two to seven words.	Mandatory in the language of filing. Optional for additional languages.
SequenceTotalQuantity	The total number of all sequences in the sequence listing including intentionally skipped sequences (also known as empty sequences) (see paragraph 10).	Mandatory

Reproduced from paragraph 45 of WIPO Standard ST.26.

WIPO Standard ST.26, paragraph 47, specifies that the name of the applicant and, optionally, the name

of the inventor must be indicated in the element ApplicantName and InventorName, respectively, as they are generally referred to in the language in which the application is filed. The appropriate language code (referenced in WIPO Standard ST.26, paragraph 9, which references the International Standard ISO 639-1:2002 for the codes for the representation of names of languages) must be indicated in the languageCode attribute for each element. Where the applicant name indicated contains characters other than those of the Latin alphabet permitted as set forth in paragraph 40(b), reproduced in MPEP § 2413.01(a) item (2), a transliteration or translation of the applicant name must also be indicated in characters of the Latin alphabet in the element ApplicantNameLatin. Where the inventor name indicated contains characters other than those of the Latin alphabet, a transliteration or a translation of the inventor name may also be indicated in characters of the Latin alphabet in the element InventorNameLatin.

WIPO Standard ST.26, paragraph 48, provides that the title of the invention must be indicated in the element InventionTitle in the language of filing and may also be indicated in additional languages using multiple InventionTitle elements. The appropriate language code (see WIPO Standard ST.26, paragraph 9, which references the International Standard ISO 639-1:2002) must be indicated in the languageCode attribute of the element. See MPEP § 2413.01(i) for more information about the InventionTitle element in a "Sequence Listing XML."

2413.01(g) The "Sequence Listing XML" Must Contain a Sequence Data Part [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**.]*

37 CFR 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.

*****

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

*****

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

*****

(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.

*****

The sequence data part is the part of the "Sequence Listing XML" that contains each individual nucleotide or amino acid sequence that meets the definition for inclusion in a "Sequence Listing XML" together with sequence-associated data. WIPO Standard ST.26, paragraph 50, specifies that the sequence data part must be composed of one or more SequenceData elements, each element containing information about one sequence.

WIPO Standard ST.26, paragraph 51, specifies that each SequenceData element must have a mandatory attribute sequenceIDNumber, in which the sequence identifier (see MPEP § 2412.05(a)) for each sequence is contained.

WIPO Standard ST.26 specifies that the SequenceData element must contain a dependent element INSDSeq, consisting of further dependent elements as follows:

List of INSDSeq Dependent Elements

Element	Description	Mandatory	Mandatory/Not Included
		Sequences	**Intentionally Skipped Sequences**
INSDSeq_length	Length of the sequence	Mandatory	Mandatory with no value
INSDSeq_moltype	Molecule type	Mandatory	Mandatory with no value
INSDSeq_division	Indication that a sequence is related to a patent application	Mandatory with the value “PAT”	Mandatory with no value
INSDSeq_feature-table	List of annotations of the sequence	Mandatory	Must NOT be included
INSDSeq_sequence	Sequence	Mandatory	Mandatory with the value “000”

Reproduced from paragraph 52 of WIPO Standard ST.26.

See MPEP § 2412.05(a) for information about intentionally skipped sequences.

WIPO Standard ST.26, paragraph 53, specifies that the element INSDSeq_length must disclose the number of nucleotides or amino acids of the sequence contained in the INSDSeq_sequence element.

WIPO Standard ST.26, paragraph 54, specifies that the element INSDSeq_moltype must disclose the type of molecule that is being represented. For nucleotide sequences, including nucleotide analogue sequences, the molecule type must be indicated as DNA or RNA. For amino acid sequences, the molecule type must be indicated as AA.

WIPO Standard ST.26, paragraph 55, specifies that for a nucleotide sequence that contains both DNA and RNA segments of one or more nucleotides, the molecule type must be indicated as DNA. The combined DNA/RNA molecule must be further described in the feature table, using the feature key “source” and the mandatory qualifier “organism” with the value “synthetic construct” and the mandatory qualifier “mol_type” with the value “other DNA.” Each DNA and RNA segment of the combined DNA/RNA molecule must be further described with the feature key “misc_feature” and the qualifier “note,” wherein the qualifier value indicates whether the segment is DNA or RNA.

WIPO Standard ST.26, paragraph 57, specifies that the element INSDSeq_sequence must disclose the sequence. Only the appropriate symbols set forth in Table 1: List of Nucleotides Symbols and Table 3: List of Amino Acids Symbols (see MPEP § 2412.03(a)) must be included in the sequence. The sequence must not include numbers, punctuation or whitespace characters.

I. FEATURE TABLE

According to WIPO Standard ST.26, a “feature table” “contains information on the location and roles of various regions within a particular sequence. A feature table is required for every sequence, except for any intentionally skipped sequence, in which case it must not be included. The feature table is contained in the element INSDSeq_feature-table, which consists of one or more INSDFeature elements.” (WIPO Standard ST.26, paragraph 60).

WIPO Standard ST.26 specifies that each INSDFeature element that comprises the feature table describes one feature, and consists of dependent elements as follows:

List of INSDFeature Dependent Elements

Element	Description	Mandatory/Optional
INSDFeature_key	A word or abbreviation indicating a feature	Mandatory
INSDFeature_location	Region of the sequence which corresponds to the feature	Mandatory
INSDFeature_quals	Qualifier containing auxiliary information about a feature	Mandatory where the feature key requires one or more qualifiers, e.g., source; otherwise, Optional

Reproduced from paragraph 61 of WIPO Standard ST.26.

II. FEATURE KEYS

WIPO Standard ST.26, paragraph 62, specifies that Annex I contains the exclusive listing of feature keys that must be used when preparing and submitting a "Sequence Listing XML," along with an exclusive listing of associated qualifiers and an indication as to whether those qualifiers are mandatory or optional. Section 5 of Annex I of WIPO Standard ST.26 provides the exclusive listing of feature keys for nucleotide sequences and Section 7 of Annex I of WIPO Standard ST.26 provides the exclusive listing of feature keys for amino acid sequences.

III. MANDATORY FEATURE KEYS

WIPO Standard ST.26, paragraph 63, specifies that the "source" feature key is mandatory for all nucleotide sequences and for all amino acid sequences, except for any intentionally skipped sequence. Each sequence must have a single "source" feature key spanning the entire sequence. Where a sequence originates from multiple sources, those sources may be further described in the feature table, using the feature key "misc_feature" and the qualifier "note" for nucleotide sequences, and the feature key "REGION" and the qualifier "note" for amino acid sequences.

IV. FEATURE LOCATION

WIPO Standard ST.26, paragraph 64, specifies that the mandatory element INSDFeature_location must contain at least one location descriptor, which defines a site or a region corresponding to a feature of the sequence in the INSDSeq_sequence element. Amino acid sequences must contain one and only one location descriptor in the mandatory INSDFeature_location element. Nucleotide sequences may have more than one location descriptor in the mandatory INSDFeature_location element when used in conjunction with one or more location operator(s) (more information about location descriptors is discussed below).

WIPO Standard ST.26, paragraph 65, specifies that the location descriptor can be a single residue number, a region delimiting a contiguous span of residue numbers, or a site or region that extends beyond the specified residue or span of residues. The location descriptor must not include numbering for residues beyond the range of the sequence in the INSDSeq_sequence element. For nucleotide sequences only, a location descriptor can be a site between two adjacent residue numbers. Multiple location descriptors must be used in conjunction with a location operator when a feature corresponds to discontinuous sites or regions of a nucleotide sequence (more information about location descriptors and operators is discussed below).

WIPO Standard ST.26, paragraph 66, specifies that the syntax for each type of location descriptor is indicated in Tables (a)-(c) below, where x and y are residue numbers, indicated as positive integers, not greater than the length of the sequence in the INSDSeq_sequence element, and x is less than y.

(a) Location descriptors for nucleotide and amino acid sequences:

Location Descriptors for Nucleotide and Amino Acid Sequences

Location descriptor type	Syntax	Description
Single residue number	x	Points to a single residue in a sequence.
Residue numbers delimitating a sequence span	x. .y	Points to a continuous range of residues bounded by and including the starting and ending residues.
Residues before the first or beyond the last specified residue number	<x >x <x. .y x. .>y <x. .>y	Points to a region including a specified residue or span of residues and extending beyond a specified residue. The '<' and '>' symbols may be used with a single residue or the starting and ending residue numbers of a span of residues to indicate that a feature extends beyond the specified residue number.

Reproduced from paragraph 66 of WIPO Standard ST.26.

(b) Location descriptors for nucleotide sequences only:

Location Descriptors for Nucleotide Sequence Only

Location descriptor type	Syntax	Description
A site between two adjoining nucleotides	x^y	Points to a site between two adjoining nucleotides, e.g., endonucleolytic cleavage site. The position numbers for the adjacent nucleotides are separated by a carat (^). The permitted formats for this descriptor are x^x+1 (for example 55^56), or, for circular nucleotides, x^1, where "x" is the full length of the molecule, i.e. 1000^1 for circular molecule with length 1000.

Reproduced from paragraph 66 of WIPO Standard ST.26.

(c) Location descriptors for amino acid sequences only:

Location Descriptors for Amino Acid Sequences Only

Location descriptor type	Syntax	Description
Residue numbers joined by an intrachain cross-link	x. .y	Points to amino acids joined by an intrachain linkage when used with a feature that indicates an intrachain cross-link, such as "CROSSLNK" or "DISULFID".

Reproduced from paragraph 66 of WIPO Standard ST.26.

WIPO Standard ST.26 specifies that the INSDFeature_location element of nucleotide sequences may contain one or more location operators. A location operator is a prefix to either one location descriptor or a combination of location descriptors corresponding to a single but discontinuous feature, and specifies where the location corresponding to the feature on the indicated sequence is found or how the feature is constructed. A list of location operators is provided in the table below with their descriptions. Location operators can be used for nucleotides only.

Location Operators

Location syntax	Location description
join (location, location ,..., location)	The indicated locations are joined (placed end-to-end) to form one contiguous sequence.
order (location, location,...,location)	The elements are found in the specified order but nothing is

Location syntax	Location description
	implied about whether joining those elements is reasonable.
complement (location)	Indicates that the feature is located on the strand complementary to the sequence span specified by the location descriptor, when read in the 5' to 3' direction or in the direction that mimics 5' to 3' direction.

Reproduced from paragraph 67 of WIPO Standard ST.26.

WIPO Standard ST.26, paragraph 68, specifies that the join and order location operators require that at least two comma-separated location descriptors be provided. Location descriptors involving sites between two adjacent residues, i.e. x^y, must not be used within a join or order combination of locations. Use of the join location operator implies that the residues described by the location descriptors are physically brought into contact by biological processes (for example, the exons that contribute to a coding region feature).

WIPO Standard ST.26, paragraph 69, specifies that the location operator "complement" can be used in combination with either "join" or "order" within the same location. Combinations of "join" and "order" within the same location must not be used. See paragraph 70, examples of WIPO Standard ST.26.

WIPO Standard ST.26, paragraph 71, specifies that in an XML instance of a "Sequence Listing XML" , the characters "<" and ">" in a location descriptor must be replaced by the appropriate predefined entities, "<" and ">", respectively (see MPEP § 2413.01(a) regarding the predefined entities).

V. FEATURE QUALIFIERS

WIPO Standard ST.26, paragraph 72, specifies that qualifiers are used to supply information about features in addition to that conveyed by the feature key and feature location. There are three types of value formats to accommodate different types of information conveyed by qualifiers, namely:

(a) free text (see MPEP §§ 2413.01(g), subsection IX and 2413.01(h), for more detail about "free text");

(b) controlled vocabulary or enumerated values (e.g., a number or date); and

(c) sequences.

WIPO Standard ST.26, paragraph 73, specifies that Section 6 of Annex I contains the exclusive listing of qualifiers and their specified value formats, if any, for each nucleotide sequence feature key and Section 8 of Annex I contains the exclusive listing of qualifiers and their specified value formats, if any, for each amino acid sequence feature key.

WIPO Standard ST.26, paragraph 74, specifies that any sequence encompassed by 37 CFR 1.831(b) (see MPEP § 2412.03) that is provided as a qualifier value must be separately included in the "Sequence Listing XML" and assigned its own sequence identifier as described in MPEP § 2412.05(a).

VI. MANDATORY FEATURE QUALIFIERS

WIPO Standard ST.26, paragraph 75, specifies that one mandatory feature key, i.e., "source" requires two mandatory qualifiers, "organism" and "mol_type." Some optional feature keys also require mandatory qualifiers. See Annex I of WIPO Standard ST.26, Sections 5 and 7, for listings of feature keys with mandatory qualifiers.

VII. QUALIFIER ELEMENTS

WIPO Standard ST.26 specifies that the element INSDFeature_quals contains one or more INSDQualifier elements. Each INSDQualifier element represents a single qualifier and consists of three dependent elements and one optional attribute, as shown below:

List of INSDQualifier Dependent Elements

Element/Attribute	Description	Mandatory/Optional
INSDQualifier_name	Name of the qualifier (see Annex I, Sections 6 and 8).	Mandatory
INSDQualifier_value	Value of the qualifier, if any, in the specified format (see Annex I, Sections 6 and 8) and composed in the characters as set forth in paragraph 40(b).	Mandatory, when specified (see paragraph 87 and Annex I, Sections 6 and 8)
NonEnglishQualifier_value	Value of the qualifier, if any, in the specified format (see Annex I, Sections 6 and 8) and composed in the characters as set forth in paragraph 40(a).	Mandatory, when specified (see paragraph 87 and Annex I, Sections 6 and 8)
NonEnglishQualifier_value	Value of the qualifier, if any, in the specified format (see Annex I, Sections 6 and 8) and composed in the characters as set forth in paragraph 40(a).	Mandatory, when specified (see paragraph 87 and Annex I, Sections 6 and 8)
id	A qualifier with a language-dependent free text value may be uniquely identified by using the optional XML attribute 'id' in the element INSDQualifier (see paragraph 87(d)). The value of the 'id' attribute must start with the letter 'q' and continue with any positive integer. The value of an 'id' attribute must be unique to one INSDQualifier element, i.e. the attribute value must only be used once in a sequence listing file.	Optional

Reproduced from paragraph 76 of WIPO Standard ST.26.

VIII. ORGANISM AND MOL_TYPE QUALIFIERS

WIPO Standard ST.26, paragraph 77, specifies that the organism qualifier, i.e., “organism” for nucleotide sequences (See Table 5: List of Qualifier Values for Nucleotide Sequences with Language-Dependent Free-Text Values reproduced in MPEP § 2413.01(h), Annex I, section 6, of WIPO Standard ST.26) and “organism” for amino acid sequences (see Table 6: List of Qualifiers for Amino Acid Sequences with Language-Dependent Free Text Values reproduced in MPEP § 2413.01(h), Annex I, section 6, of WIPO Standard ST.26) must disclose the source, i.e., a single organism or origin, of the sequence. Organism designations should be selected from a taxonomy database.

WIPO Standard ST.26, paragraph 78, specifies that if the sequence is naturally occurring and the source organism has a Latin genus and species designation, that designation must be used as the qualifier value. The preferred English common name may be specified using the qualifier “note” for nucleotide sequences and amino acid sequences, but must not be used in the organism qualifier value.

WIPO Standard ST.26, paragraph 80, specifies that if the sequence is naturally occurring and the source organism has a known Latin genus, but the species

is unspecified or unidentified, then the organism qualifier value must indicate the Latin genus followed by "sp".

WIPO Standard ST.26, paragraph 81, specifies that if the sequence is naturally occurring, but the Latin organism genus and species designation is unknown, then the organism qualifier value must be indicated as "unidentified". Any known taxonomic information should be indicated in the qualifier "note" for nucleotide sequences and the qualifier "note" for amino acid sequences.

WIPO Standard ST.26, paragraph 82, specifies that if the sequence is naturally occurring and the source organism does not have a Latin genus and species designation, such as a virus, then another acceptable scientific name (e.g., "Canine adenovirus type 2") must be used as the organism qualifier value.

WIPO Standard ST.26, paragraph 83, specifies that if the sequence is not naturally occurring, the organism qualifier value must be indicated as "synthetic construct." Further information with respect to the way the sequence was generated may be specified using the qualifier "note" for nucleotide sequences and the qualifier "note" for amino acid sequences.

WIPO Standard ST.26, paragraph 84, specifies that the "mol_type" qualifier for nucleotide sequences and "mol_type" qualifier for amino acid sequences must disclose the type of molecule represented in the sequence. These qualifiers are distinct from the element INSDSeq_moltype discussed above where INSDSeq_moltype for nucleotide sequences, including nucleotide analogue sequences must be indicated as DNA or RNA, and for amino acid sequences, must be indicated as AA:

(1) For a nucleotide sequence, the "mol_type" qualifier value must be one of the following: "genomic DNA", "genomic RNA", "mRNA", "tRNA", "rRNA", "other RNA", "other DNA", "transcribed RNA", "viral cRNA", "unassigned DNA", or "unassigned RNA". If the sequence is not naturally occurring, i.e. the value of the "organism" qualifier is "synthetic construct", the "mol_type" qualifier value must be either "other RNA" or "other DNA";

(2) For an amino acid sequence, the "mol_type" qualifier value is "protein."

IX. FREE TEXT

WIPO Standard ST.26, paragraph 85, specifies that "free text" is a type of value format for certain qualifiers presented in the form of a descriptive text phrase or other specified format (see MPEP § 2413.01(h) for the definition of "free text" and see Annex I of WIPO Standard ST.26 for controlled vocabulary).

WIPO Standard ST.26, paragraph 86, specifies that the use of free text must be limited to a few short terms indispensable for the understanding of a characteristic of the sequence. For each qualifier other than the "translation" qualifier, the free text must not exceed 1000 characters.

WIPO Standard ST.26, paragraph 87, specifies that language-dependent free text is the free text value of certain qualifiers that is language-dependent in that it may require translation for international, national, or regional procedures. Qualifiers for nucleotide sequences with a language-dependent free text value format are identified in Annex I, Table 5: List of Qualifiers with Language-Dependent FreeText Values for Nucleotide Sequences (reproduced in MPEP § 2413.01(h)). Qualifiers for amino acid sequences with a language-dependent free text value format are identified in Annex I, Table 6: List of Qualifiers with Language-Dependent Free Text Values for Amino Acid Sequences (reproduced in MPEP § 2413.01(h)).

X. CODING SEQUENCES

WIPO Standard ST.26, paragraph 89, specifies that the "CDS" feature key may be used to identify coding sequences, i.e., sequences of nucleotides which correspond to the sequence of amino acids in a protein and the stop codon. The location of the "CDS" feature in the mandatory element INSDFeature_location must include the stop codon.

WIPO Standard ST.26, paragraph 90, specifies that the "transl_table" and "translation" qualifiers may be used with the "CDS" feature key (see Annex I of WIPO Standard ST.26). Where the "transl_table"

qualifier is not used, the use of the Standard Code Table (see Annex I, Section 9, Table 7 of WIPO Standard ST.26) is assumed.

WIPO Standard ST.26, paragraph 91, specifies that the "transl_except" qualifier must be used with the "CDS" feature key and the "translation" qualifier to identify a codon that encodes either pyrrolysine or selenocysteine.

WIPO Standard ST.26, paragraph 92, specifies that an amino acid sequence encoded by the coding sequence and disclosed in a "translation" qualifier that is encompassed by the description of sequences found in MPEP § 2412.03 must be included in the sequence listing and assigned its own sequence identifier. The sequence identifier assigned to the amino acid sequence must be provided as the value in the qualifier "protein_id" with the "CDS" feature key. The "organism" qualifier of the "source" feature key for the amino acid sequence must be identical to that of its coding sequence.

XI. VARIANTS

MPEP § 2412.05(c) provides information about representation and inclusion of variants

WIPO Standard ST.26, paragraph 93, specifies that a primary sequence and any variant of that sequence, each disclosed by enumeration of its residues and encompassed by the description of sequences found in MPEP § 2412.03 must each be included in the sequence listing and assigned its own sequence identifier.

WIPO Standard ST.26, paragraph 94, specifies that any variant sequence, disclosed as a single sequence with enumerated alternative residues at one or more positions, must be included in the sequence listing and should be represented by a single sequence, wherein the enumerated alternative residues are represented by the most restrictive ambiguity symbol. See MPEP § 2412.05(b), subsection II, for more information regarding representing alternative nucleotide residues and MPEP § 2412.05(d), subsection II, for more information regarding representing alternative amino acid residues.

WIPO Standard ST.26, paragraph 95, specifies that any variant sequence, disclosed only by reference to deletion(s), insertion(s), or substitution(s) in a primary sequence in the sequence listing, should be included in the sequence listing. Where included in the sequence listing, such a variant sequence:

(a) may be represented by annotation of the primary sequence, where it contains variation(s) at a single location or multiple distinct locations and the occurrence of those variations are independent;

(b) should be represented as a separate sequence and assigned its own sequence identifier, where it contains variations at multiple distinct locations and the occurrence of those variations are interdependent; and

(c) must be represented as a separate sequence and assigned its own sequence identifier, where it contains an inserted or substituted sequence that contains in excess of 1000 residues (see WIPO Standard ST.26, paragraph 86).

WIPO Standard ST.26, paragraph 96, specifies the proper use of feature keys and qualifiers for nucleic acid and amino acid sequence variants from the table List of Feature Keys and Qualifiers (reproduced in MPEP § 2412.05(c)).

WIPO Standard ST.26, paragraph 97, specifies that annotation of a sequence for a specific variant must include a feature key and qualifier, as indicated in the table in MPEP § 2412.05(c), and the feature location. The value for the "replace" qualifier must be only a single alternative nucleotide or nucleotide sequence using only the symbols in set forth Table 1: List of Nucleotides Symbols (see MPEP § 2413.01(a)), or empty. A listing of alternative residues may be provided as the value in the "note" qualifier. In particular, a listing of alternative amino acids must be provided as the value in the "note" qualifier where "X" is used in a sequence, and represents a value other than "any one of 'A', 'R', 'N', 'D', 'C', 'Q', 'E', 'G', 'H', 'I', 'L', 'K', 'M', 'F', 'P', 'O', 'S', 'U', 'T', 'W', 'Y', or 'V.'" A deletion must be represented by an empty qualifier value for the "replace" qualifier or by an indication in the "note" qualifier that the residue may be deleted. An inserted or substituted residue(s) must be provided in the "replace" or "note" qualifier. The value format for the "replace" and "note" qualifiers

is free text and must not exceed 1000 characters. See below for sequences encompassed by the definition in MPEP § 2412.03 that are provided as an insertion or a substitution in a qualifier value.

WIPO Standard ST.26, paragraph 98, specifies that the symbols set forth in Tables 1 to 4 of Annex I, reproduced in MPEP §§ 2412.03(a), 2412.03(c), and 2412.05(b), subsection III, should be used to represent variant residues where appropriate. For the "note" qualifier, where the variant residue is a modified residue not set forth in Tables 2 or 4 the complete unabbreviated name of the modified residue must be provided as the qualifier value. Modified residues must be further described in a feature table as described in MPEP § 2412.05(b), subsection III for modified nucleotides and MPEP § 2412.05(d), subsection III, for modified amino acids.

WIPO Standard ST.26, paragraph 100, specifies that a sequence encompassed by the description of sequences found in MPEP § 2412.03 that is provided as an insertion or a substitution in a qualifier value for a primary sequence annotation must also be included in the sequence listing and assigned its own sequence identifier.

2413.01(h) Language Dependent Free Text Qualifier Values in the English Language [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**.]*

37 CFR 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.

*****

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

*****

(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).

WIPO Standard ST.26, paragraph 3(n), uses the term "free text" to mean a type of value format for certain qualifiers, presented in the form of a descriptive text phrase or other specified format.

WIPO Standard ST.26, paragraph 3(o), uses the term "language-dependent free text" to mean the free text value of certain qualifiers that may require translation for international, national or regional procedures.

WIPO Standard ST.26, paragraph 85, provides that free text is a type of value format for certain qualifiers presented in the form of a descriptive text phrase or other specified format with reference to the controlled vocabulary found in Annex I of WIPO Standard ST.26. See MPEP § 2413.01(g), subsection IX.

WIPO Standard ST.26, paragraph 86, requires that the use of free text be limited to short terms indispensable for the understanding of a characteristic of the sequence. For each qualifier other than the "translation" qualifier, the free text must not exceed 1000 characters.

WIPO Standard ST.26, paragraph 87, specifies that language-dependent free text is the free text value of certain qualifiers that is language-dependent in that it may require translation for international, national, regional procedures. Qualifiers with language-dependent free text values for nucleotide sequences are identified below:

Table 5: List of Qualifiers with Language-Dependent FreeText Values for Nucleotide Sequences

Section	Language-Dependent Free Text Qualifier
6.3	bound_moiety
6.5	cell_type
6.8	clone
6.9	clone_lib
6.11	collected_by
6.14	cultivar

Section	Language-Dependent Free Text Qualifier
6.15	dev_stage
6.18	ecotype
6.21	frequency
6.22	function
6.24	gene_synonym
6.26	haplogroup
6.28	host
6.29	identified_by
6.30	isolate
6.31	isolation_source
6.32	lab_host
6.36	mating_type
6.41	note
6.45	organism
6.47	phenotype
6.49	pop_variant
6.50	product
6.66	serotype
6.67	serovar
6.68	sex
6.69	standard_name
6.70	strain
6.71	sub_clone
6.72	sub_species
6.73	sub_strain
6.75	tissue_lib
6.76	tissue_type
6.81	variety

Reproduced from Table 5 of WIPO Standard ST.26, Annex I, Section 6.

Qualifiers with a language-dependent free text values for amino acid sequences are identified below:

Table 6: List of Qualifiers with Language-Dependent Free Text Values for Amino Acid Sequences

Section	Language-Dependent Free Text Qualifier
8.2	note
8.3	organism

Reproduced from Table 6 of WIPO Standard ST.26, Annex I, Section 8.

(a) Language-dependent free text must be presented in the INSDQualifier_value element in English, or in the NonEnglishQualifier_value element in a language other than English, or in both elements. Note that if an organism name is a Latin genus and species name, no translation is required. Technical terms and proper names originating from non-English words that are used internationally are considered English for the purpose of the value of the INSDQualifier_value element (e.g., ‘in vitro’, ‘in vivo’).

(b) If a NonEnglishQualifier_value element is present in a sequence listing, the appropriate language code (see WIPO Standard ST.26 at paragraph 9, which references the International Standard ISO 639-1:2002) must be indicated in the nonEnglishFreeTextLanguageCode attribute in the root element (see List of File Attributes reproduced in MPEP § 2413.01(e)). All NonEnglishQualifier_value elements in a single sequence listing must have values in the language indicated in the nonEnglishFreeTextLanguageCode attribute. The NonEnglishQualifier_value element is permitted only for qualifiers that have a language-dependent free text value format.

(c) Where NonEnglishQualifier_value and INSDQualifier_value are both present for a single qualifier, the information contained in the two elements must be equivalent. One of the following conditions must be true: NonEnglishQualifier_value contains a translation of the value of INSDQualifier_value; or, INSDQualifier_value contains a translation of the value of NonEnglishQualifier_value; or, both elements contain a translation of the qualifier value from the language specified in the originalFreeTextLanguageCode attribute (see List of File Attributes reproduced in MPEP § 2413.01(e)).

(d) For qualifiers with a language-dependent free text value, the INSDQualifier element may include an optional attribute *id*. The value of this attribute must be in the format “q” followed by a positive integer, e.g. “q23”, and must be unique to one INSDQualifier element, i.e. the attribute value must only be used once in a “Sequence Listing XML” file (WIPO Standard ST.26, paragraph 87).

2413.01(i) Title Element in "Sequence Listing XML" [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**.]*

The element InventionTitle described in the table of elements of the general information part (reproduced in MPEP § 2413.01(f) from WIPO Standard ST.26, paragraph 45) as a required element of the "general information" part of the "Sequence Listing XML" must be in the language of filing. It is possible for an applicant to provide more than one InventionTitle element of the general information part in more than one language. However, when the USPTO is the Receiving Office (RO/US) for an international application, the English language is required for the InventionTitle element and also for language-dependent free text of the "Sequence Listing XML." (See MPEP § 2413.01(h) for more information about language-dependent free text.) If non-English language is presented to RO/US for a "Sequence Listing XML," then RO/US will transfer the international application to the International Bureau, under PCT Rule 19.4.

2413.02 Form and Format of the XML file containing the "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

37 CFR 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 nonprintable (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.

*****

In order for the USPTO to be able to process the "Sequence Listing XML" .xml file, all characters must be encoded using Unicode UTF-8. The file must be compatible with PC or Mac® computers using one of the following operating systems, MS–DOS®, MS-Windows®, Mac OS®, or Unix®/Linux®. The printable and non-printable characters in the .xml file are defined in paragraph 40 and 41 of WIPO Standard ST.26 (see MPEP § 2413.01(a)) where Annex IV of WIPO Standard ST.26 provides a table of the CHARACTER SUBSET FROM THE UNICODE BASIC LATIN CODE TABLE FOR USE IN AN XML INSTANCE OF A SEQUENCE LISTING.

2413.03 How to Submit the "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

37 CFR 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.

*****

(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);

*****

For submission of a "Sequence Listing XML" that is 100MB in size or less, the preferred method of submission is via the USPTO patent electronic filing system. Only the Patent Center system of the USPTO patent electronic filing system is capable of accepting .xml files. If the "Sequence Listing XML" is larger than 100MB in size, then applicant may submit the .xml file on physical media using read-only optical discs (CDs or DVDs). "Sequence Listing XML" submissions on discs may be compressed if the resulting compressed file is non-self-extracting.

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 a 35 U.S.C. 111(a) application having a filing date on or after July 1, 2022, will result in a notice from the Office of Patent Application Processing (OPAP) 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." See MPEP § 2415.03 for addressing improper submission of a "Sequence Listing" that complies with 37 CFR 1.821-1.824 where a "Sequence Listing XML" that complies with 37 CFR 1.831-1.834 is required in a national nonprovisional application.

To facilitate administrative processing of all papers and read-only optical discs associated with sequence rule compliance, all read-only optical discs, fees, and papers accompanying them filed in the Office should be marked "Mail Stop SEQUENCE." Correspondence relating to the submission of a "Sequence Listing XML" may also be hand-delivered to the Customer Service Window. In cases of hand delivery to the Customer Service Window, the read-only optical disc enclosed in a hard case should be placed in a protective mailer. The use of staples and clips, if any, should be confined to carefully attaching the protective mailer to the submitted papers without contact or compression of the media. The labeling requirements of 37 CFR 1.52(e) and including the application number (if known), apply to all read-only optical disc submissions. For submission of a "Sequence Listing XML" larger than 100MB in size in a new application, it is recommended that the user file the application without the "Sequence Listing XML" using the USPTO patent electronic filing system to obtain the application number, and then file the "Sequence Listing XML" on read-only optical disc(s) in accordance with 37 CFR 1.52(e) on the same day by using Priority Mail Express® from the USPS in accordance with 37 CFR 1.10, or hand delivery, in order to secure the same filing date for all parts of the application. In no situations should additional or complimentary electronic copies be delivered to examiners or other Office personnel.

2413.04 Requirements Regarding Incorporation By Reference of the "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

37 CFR 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.

*****

(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.

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

*****

(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.

*****

Since the "Sequence Listing XML" is not the text of the specification, but rather is sequence data in an XML file format, an incorporation by reference statement is needed to ensure that the content of the "Sequence Listing XML," submitted to the USPTO as an XML file, is considered part of the disclosure capable of providing 35 U.S.C. 112(a) support for the disclosure and any claims relating to nucleotide and/or amino acid sequences. The incorporation by reference statement identifies: (i) the name of the file; (ii) the date of creation of the file; and (iii) the size of the file in bytes Note that this requirement pertaining to applicant submission of a "Sequence Listing XML" does not apply to a sequence listing that is part of an international application and communicated to the USPTO under PCT Article 20, for a national phase application.

The incorporation by reference statement of the material in an .xml file is required to be part of the specification so it is clear to the Office, the printer, and the public that the application as originally filed includes material on an .xml file. However, for international applications during the international phase, where a "Sequence Listing XML" is submitted via the USPTO patent electronic filing system or on read-only optical disc, no such incorporation by reference statement is required.

If an applicant submits a "Sequence Listing XML" with the complete specification but fails to include the required incorporation by reference paragraph, a notice can be issued by pre-examination requiring the incorporation by reference paragraph in the context of a substitute specification under 37 CFR 1.125. Similarly, the examiner could require applicant to amend the specification to include an incorporation by reference paragraph of the "Sequence Listing XML" file. See also form paragraph 24.24.26.

2413.05 Presumptions Regarding Compliance [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

Neither the presence nor absence of information which is not required under the sequence rules will create a presumption that such information is necessary to satisfy any of the requirements of 35 U.S.C. 112. Further, the grant of a patent on an application that is subject to 37 CFR 1.831 through 37 CFR 1.835 constitutes a presumption that the granted patent complies with the requirements of these rules.

2414 Notification of a Failure to Comply with Sequence Listing Requirements and Amendments Relating to "Sequence Listing XML" Files Under 37 CFR 1.835 [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

When an application filed under 35 U.S.C. 111(a) having a filing date on or after July 1, 2022, or an application which entered the national stage, having an international filing date on or after July 1, 2022, contains disclosure of nucleotide and/or amino acid

sequences, a "Sequence Listing XML" is required. When no "Sequence Listing XML" or a defective "Sequence Listing XML" is submitted in an application where a compliant "Sequence Listing XML" is required, the pre-examination staff will issue a notification that compliance with 37 CFR 1.831 - 1.835 is required and an added initial submission of a "Sequence Listing XML" (37 CFR 1.835(a)) or replacement "Sequence Listing XML" (37 CFR 1.835(b)) will be required. Additionally, the examiner can require the filing of a "Sequence Listing XML" if an application fails to comply with 37 CFR 1.831 - 1.834. For example, if there are sequences in the disclosure which are not part of the "Sequence Listing XML" of record, then the examiner can require a replacement "Sequence Listing XML" which includes the omitted sequences. See form paragraphs 24.18.26 - 24.20.26, 24.23.26, and 24.27.26-24.28.26.

2414.01 Issuance of a Notice Relating to the Requirements For Patent Applications Containing Nucleotide and/or Amino Acid Sequence Disclosures [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

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

*****

(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.

*****

Initial review by the pre-examination staff at the USPTO checks for compliance with formal matters. In order to ensure that an application is ready for examination, pre-examination staff check that applications containing disclosures of nucleotides and/or amino acid sequences have complied with the sequence listing rules. This means that for applications having a filing date on or after July 1, 2022, containing such sequence disclosures, a "Sequence Listing XML" must be of record and if not, applicant will be notified of such deficiency and provided instructions for responding. Adding a "Sequence Listing XML" where one was not previously filed must conform with the requirements of 37 CFR 1.835(a). See MPEP § 2414.02 for details. To amend a "Sequence Listing XML" due to errors or omissions, the procedure is outlined in 37 CFR 1.835(b). See MPEP § 2414.03 for details.

2414.02 Amendment to Add an Initial "Sequence Listing XML" under 37 CFR 1.835(a) [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

37 CFR 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.

*****

Part of the pre-examination process as well as initial review by the examiner involves identifying whether an application that discloses nucleotide and/or amino acid sequences requires submission of a "Sequence Listing XML," and if so, if it contains a compliant "Sequence Listing XML." If pre-examination staff identifies that a "Sequence Listing XML" is required but missing or that compliance with 37 CFR 1.831 - 1.835 is lacking, a notice indicating the deficiencies will be issued to the applicant. Similarly, upon having a new application placed on the examiner's docket, the examiner can assess whether a disclosure of nucleotide and/or amino acid sequences in the application requires a "Sequence Listing XML" to be submitted where one has not been submitted. The examiner can identify the requirement for a "Sequence Listing XML" on an "OPAP Routing Sheet" (Document Code "SEQREQ") and return the application to pre-examination for the mailing of a notice indicating the deficiencies.

Applicant's response to a notice or an Office Action requiring an initial "Sequence Listing XML" would involve providing the following:

(1) A compliant "Sequence Listing XML" submitted as an XML file using either the USPTO patent electronic filing system (currently, XML files can only be submitted via Patent Center) or a read-only optical disc that is mailed or hand-delivered into the Office. If the size of "Sequence Listing XML" file exceeds the upload limit of Patent Center, then it must be submitted on read-only optical disc. (37 CFR 1.835(a)(1));

(2) An amendment to the specification to incorporate by reference the material in the "Sequence Listing XML" by reciting in a separate paragraph of the specification the name of the file, the date of creation, and the size of the file in bytes (37 CFR 1.835(a)(2) and 37 CFR 1.835(c));

(3) A statement identifying the basis for the added "Sequence Listing XML," which would include, e.g., paragraphs of the originally filed specification or sheets of the drawings where a sequence is disclosed, identifying each sequence that is part of the "Sequence Listing XML" (37 CFR 1.835(a)(3)); and

(4) A statement that no new matter is introduced by the content of the "Sequence Listing XML" (37 CFR 1.835(a)(4)).

2414.03 Amendment Adding to, Deleting from, or Replacing Sequence Information in a "Sequence Listing XML" under 37 CFR 1.835(b) [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

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

*****

(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.

*****

Part of the pre-examination process as well as initial review by the examiner involves identifying whether an application that discloses nucleotide and/or amino acid sequences contains a compliant "Sequence Listing XML." If pre-examination staff identifies that a previously submitted "Sequence Listing XML" contains errors and/or the application fails to comply with 37 CFR 1.831 - 1.834, a notice indicating the deficiencies will be issued to the applicant. Similarly, upon having a new application placed on the examiner's docket, the examiner can identify errors in any "Sequence Listing XML" and/or instances of noncompliance with 37 CFR 1.831 - 1.834 that might not have been identified by pre-examination staff. The examiner can identify the outstanding requirements on a "OPAP Routing Sheet" (Document Code "SEQREQ") and return the application to pre-examination status for the mailing of a notice indicating the deficiencies.

Applicant response to a notice or an Office action requiring a replacement "Sequence Listing XML" would involve submitting the following:

(1) A compliant replacement "Sequence Listing XML" submitted as an XML file containing the entire "Sequence Listing XML," including any additions, deletions, or replacements of sequence information, using either the USPTO patent electronic filing system (currently, XML files can only be submitted via Patent Center) or a read-only optical disc that is mailed or hand-delivered into the Office. If the size of "Sequence Listing XML" file exceeds the upload limit of Patent Center, then it must be submitted on read-only optical disc (37 CFR 1.835(b)(1));

(2) An amendment to the specification to incorporate by reference the material in the replacement "Sequence Listing XML" by reciting in a separate paragraph of the specification the name of the file, the date of creation, and the size of the file in bytes (37 CFR 1.835(b)(1) and 37 CFR 1.835(c));

(3) A statement that identifies the location of all additions, deletions, or replacements of sequence information relative to the replaced "Sequence Listing XML" (37 CFR 1.835(b)(3));

(4) A statement identifying the basis for the additions, deletions, or replacements of the sequence information in the application as originally filed for all amended sequence data in the replacement "Sequence Listing XML," which would include, e.g., paragraphs of specification or sheets of the drawings where a sequence is disclosed, identifying the sequence which is part of the replacement "Sequence Listing XML" (37 CFR 1.835(b)(4)); and

(5) A statement that no new matter is introduced by the content of the replacement "Sequence Listing XML" (37 CFR 1.835(b)(5)).

If applicant recognizes a deficiency in a previously submitted "Sequence Listing XML" or otherwise chooses to amend a previously submitted "Sequence Listing XML," a replacement "Sequence Listing XML" must be filed including all of the above-described items.

2414.04 Translation of "Sequence Listing XML" is Not an Amended "Sequence Listing XML" under 37 CFR 1.835(b) [R-01.2024]

*[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b). Formatting representations of XML (eXtensible Markup Language) elements in this section appear different than shown in Standard ST.26, which may be accessed at: **www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf**.]*

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

*****

(d)(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.

*****

In accordance with 37 CFR 1.52(b)(1)(ii), any:

> "application or proceeding and any amendments or corrections to the application (including any translation submitted pursuant to paragraph (d) of this section) or proceeding, except as provided for in § 1.69 and paragraph (d) of this section, must:
> ***
> (ii) Be in the English language or be accompanied by a translation of the application and a translation of any corrections or amendments into the English language together with a statement that the translation is accurate."

Thus, any "Sequence Listing XML" that contains one or more qualifiers having a free text value that is language-dependent must comprise an "INSDQualifier_value" element containing text in the English language, otherwise a translation of the "Sequence Listing XML" is required.

The provisions of 37 CFR 1.835(d)(2) indicate that providing a translation of such language-dependent free text values into English in a new "Sequence Listing XML" does not constitute an added or amended "Sequence Listing XML" under 37 CFR 1.835(a) or 1.835(b).

2414.05 Notifications for Failure to Submit a "Sequence Listing XML" in an International Application under the PCT to be Searched by the United States International Searching Authority or Examined by International Preliminary Examination Authority [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

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

*****

(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."

*****

Where an international application, filed under the PCT, contains a disclosure of nucleotide and/or amino acid sequences, PCT Rule 5.2 requires that the description shall include a sequence listing part of the description complying with the standard provided for in the Administrative Instructions (AI). In accordance with Section 208, any sequence listing, whether forming part of the international application or not forming part of the international application, shall comply with Annex C of the AI. Where applicant has not provided a sequence listing in accordance with Annex C of the AI and the USPTO acts as International Searching Authority (ISA) or International Preliminary Examining Authority (IPEA), the ISA/IPEA may invite applicant to furnish a sequence listing, with a late furnishing fee, under PCT Rule 13 *ter* . This invitation will specify a time limit for a proper response. A proper response to such an invitation would include, the submission of a sequence listing accordance with Annex C of the AI, the late furnishing fee specified in 37 CFR 1.445(a)(5), and a statement to the effect that the sequence listing does not go beyond the disclosure of the international application as filed. If a sequence listing compliant with WIPO Standard ST.26 has not been furnished to the ISA/IPEA within the time limit set forth in the invitation, the ISA/IPEA will only search/examine the international application to the extent that a meaningful search or examination can be performed without the sequence listing.

See MPEP §§ 1823.02 and 1848.

2414.06 Amendment of a "Sequence Listing XML" in a Patent [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

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

*****

(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.

Where an amendment is made to a "Sequence Listing XML" contained in a patent, the correction must be made by way of an amendment to the "Sequence Listing XML" contained in a patent, pursuant to 1.835(b).

2415 Applicability date of "Sequence Listing XML" Requirements to Applications Based on Filing Date of Application and Procedures for Failure to Submit Correct Format of Sequence Listing [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

Patent applications that contain disclosures of nucleotide and/or amino acid sequences must present the sequences and associated biological sequence data in a standardized electronic format. For those applications having a filing date on or after July 1, 2022, a "Sequence Listing XML" (see 37 CFR 1.831(a)) as a separate part of the specification is required. Note that for national phase applications submitted under 35 U.S.C. 371, the relevant date is the international filing date.

2415.01 Determining if a "Sequence Listing XML" Must be Submitted [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

An application filed under 35 U.S.C. 111(a) that has a filing date on or after July 1, 2022, or an application which entered the national stage, having an international 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 filing date or 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), 37 CFR 1.121 (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 transform that "Sequence Listing" into 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. See also MPEP § 1893.03(b).

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 37 CFR 1.831-1.835.

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 WIPO Standard ST.25) or 37 CFR 1.831 through 1.835 (rules based on WIPO Standard ST.26) simplifies the application of the sequence rules, both for the USPTO and the applicant. Though 37 CFR 1.821 through 1.825 remain applicable for applications with a filing date before July 1, 2022, those provisions will not be applicable to applications filed on or after July 1, 2022.

2415.02 Provisional Applications Containing Disclosures of Nucleotides and/or Amino Acids, Compliance with 37 CFR 1.831-1.834 [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

37 CFR 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.

*****

Similar to the rules of practice regarding provisional applications filed before July 1, 2022, that contain disclosures of nucleotide and/or amino acid sequences, provisional applications filed on or after July 1, 2022, are not required to contain a separate compliant sequence listing. However, if an applicant chooses to file a separate sequence listing in an application filed on or after July 1, 2022, the submission must be by way of a "Sequence Listing XML" that is compliant with 37 CFR 1.831 - 1.834. Provisional applications cannot be amended, so if a provisional application contains a non-compliant "Sequence Listing XML" no notice relating to the requirements will be mailed in the application.

2415.03 Improper Submissions of "Sequence Listing" under 1.821(c) When a "Sequence Listing XML" was Required [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

While implementing regulations and procedures for WIPO Standard ST.26, the USPTO recognized that an applicant might erroneously provide a "Sequence Listing" in ASCII plain text file format even though a "Sequence Listing XML" is required. Therefore, in the rare circumstance in which a "Sequence Listing" .txt file 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 MPEP § 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 file, or 37 CFR 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 file, 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" erroneously submitted as an ASCII plain text file that is relied on for submission of a compliant "Sequence Listing XML," (i.e., identifying the nucleotide and/or amino acid sequence information found in the "Sequence Listing XML" within the erroneously submitted ASCII "Sequence Listing") 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.

2416 Form Paragraphs [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

¶ 24.17.26 Heading for ST.26 Sequence Requirements

Summary of Requirements for Patent Applications Filed On Or After July 1, 2022, That Have Sequence Disclosures

37 CFR 1.831(a) requires that patent applications which contain disclosures of nucleotide and/or amino acid sequences that fall within the definitions of 37 CFR 1.831(b) must contain a "Sequence Listing XML", *as a separate part of the disclosure*, which presents the nucleotide and/or amino acid sequences and associated information using the symbols and format in accordance with the requirements of 37 CFR 1.831-1.835. This "Sequence Listing XML" part of the disclosure may be submitted:

1. In accordance with 37 CFR 1.831(a) using the symbols and format requirements of 37 CFR 1.832 through 1.834 via the USPTO patent electronic filing system (see Section I.1 of the Legal Framework for Patent Electronic System (**https://www.uspto.gov/PatentLegalFramework**), hereinafter "Legal Framework") in XML format, together with an incorporation by reference statement of the material in the XML file in a separate paragraph of the specification (an incorporation by reference paragraph) as required by 37 CFR 1.835(a)(2) or 1.835(b)(2) identifying:

    a. the name of the XML file

    b. the date of creation; and

    c. the size of the XML file in bytes; or

2. In accordance with 37 CFR 1.831(a) using the symbols and format requirements of 37 CFR 1.832 through 1.834 on read-only optical disc(s) as permitted by 37 CFR 1.52(e)(1)(ii), labeled according to 37 CFR 1.52(e)(5), with an incorporation by reference statement of the material in the XML format according to 37 CFR 1.52(e)(8) and 37 CFR 1.835(a)(2) or

1.835(b)(2) in a separate paragraph of the specification identifying:

a. the name of the XML file;

b. the date of creation; and

c. the size of the XML file in bytes.

SPECIFIC DEFICIENCIES AND THE REQUIRED RESPONSE TO THIS NOTICE ARE AS FOLLOWS:

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be followed by one or more of form paragraphs 24.18.26 - 24.29.26.

¶ 24.18.26 No "Sequence Listing XML" part of the disclosure

Specific deficiency - This application fails to comply with the requirements of 37 CFR 1.831-1.834 because it does not contain a "Sequence Listing XML" as a separate part of the disclosure. A "Sequence Listing XML" is required because <1>.

Required response - Applicant must provide:

• A "Sequence Listing XML" part of the disclosure, as described above in item 1. or 2.; **together with**

o A statement that indicates **the basis for the amendment,** with specific references to particular parts of the application as originally filed, as required by 37 CFR 1.835(a)(3);

o A statement that the "Sequence Listing XML" includes no new matter as required by 37 CFR 1.835(a)(4)

AND

• A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 inserting the required incorporation by reference paragraph as required by 37 CFR 1.835(a)(2), consisting of:

A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

A copy of the amended specification without markings (clean version); and

A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application that has no "Sequence Listing XML" in compliance with 37 CFR 1.831-1.834. The examiner should explain why a "Sequence Listing XML" is required in <1>.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.19.26 Defective "Sequence Listing XML"

Specific deficiency - This application fails to comply with the requirements of 37 CFR 1.831-1.834 because the "Sequence Listing XML," as a separate part of the disclosure, is defective, damaged or unreadable. Refer to document "Sequence Listing in Computer Readable Format is Defective " dated **[1]**.

Required response - Applicant must provide:

• A replacement "Sequence Listing XML" part of the disclosure, as described above submitted in accordance with either item 1. or 2.; together with

o A statement that identifies the location of all additions, deletions or replacements of sequence information relative to the replaced "Sequence Listing XML" as required by 37 CFR 1.835(b)(3);

o A statement that indicates support for the replacement "Sequence Listing XML" in the application, as filed, as required by 37 CFR 1.835(b)(4); and

o A statement that the replacement "Sequence Listing XML" includes no new matter as required by 37 CFR 1.835(b)(5).

AND

• A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125, inserting the required incorporation-by-reference paragraph as required by 37 CFR 1.835(b)(2), consisting of:

o A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

o A copy of the amended specification without markings (clean version); and

o A statement that the substitute specification contains no new matter.

Examiner Note:

1 This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2 This form paragraph must be preceded by form paragraph 24.17.26.

3 This form paragraph should be used for an application where a defective "Sequence Listing XML" was filed and a notice was sent to the applicant including the errors identified according to STIC.

4 This form paragraph may be followed by one or more deficiency form paragraphs.

5 This form paragraph would only be used when an applicant provides a disc of a "Sequence Listing XML" in reply to an examiner requirement for a "Sequence Listing XML."

¶ 24.20.26 "Sequence Listing XML" contains errors according to STIC report

The "Sequence Listing XML" part of the disclosure filed with this application in accordance with 37 CFR 1.831-1.834 has been found to contain an error or errors as indicated on the document "Computer Readable Form (CRF) for Sequence Listing - Defective" dated **[1]**. Applicant must provide:

• A replacement "Sequence Listing XML" part of the disclosure, as described above submitted in accordance with either item 1. or 2., together with

o A statement that identifies the location of all additions, deletions, or replacements of sequence information in the replacement "Sequence Listing XML" as required by 37 CFR 1.835(b)(3);

o A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.835(b)(4);

o A statement that the replacement "Sequence Listing XML" includes no new matter as required by 37 CFR 1.835(b)(5); and

o A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 inserting the required incorporation by reference paragraph as required by 37 CFR 1.835(b)(2), consisting of:

A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

A copy of the amended specification without markings (clean version); and

A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where a defective "Sequence Listing XML" was filed and a notice was sent to the applicant including the errors according to STIC.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.21.26 Sequence IDs not present in specification

Specific deficiency - Sequences appearing in the specification are not identified by sequence identifiers (i.e., "SEQ ID NO:X" or the like) in accordance with 37 CFR 1.831(c).

Required response – Applicant must provide: A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 inserting the required sequence identifiers, consisting of:

• A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

• A copy of the amended specification without markings (clean version); and

• A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where a sequence is embedded in the text of the description or claims of the patent application and reference has not also been 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.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.22.26 Sequence IDs not present in drawings

Specific deficiency - Sequences appearing in the drawings are not identified by sequence identifiers in accordance with 37 CFR 1.831(c). Sequence identifiers for sequences (i.e., "SEQ ID NO:X" or the like) must appear either in the drawings or in the Brief Description of the Drawings.

Required response – Applicant must provide:

Amended drawings in accordance with 37 CFR 1.121(d) inserting the required sequence identifiers;

AND/OR

A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 inserting the required sequence identifiers (i.e., "SEQ ID NO:X" or the like) into the Brief Description of the Drawings, consisting of:

• A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

• A copy of the amended specification without markings (clean version); and

• A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where a sequence is embedded in the text of the drawings and reference has not also been made to the sequence by use of the sequence identifier, preceded by "SEQ ID NO:X" or the like, in either the text of the drawings or the Brief Description or the Drawings.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.23.26 Sequence in specification, drawings, or claims that is not in XML

This application contains sequence disclosures in accordance with the definitions for nucleotide and/or amino acid sequences set forth in 37 CFR 1.831(a) and 1.831(b). However, this application fails to comply with the requirements of 37 CFR 1.831-1.834. The examiner has noted that **[1]**. Applicant must provide:

• A replacement "Sequence Listing XML" part of the disclosure, as described above in item 1. or 2., as well as

• A statement that identifies the location of all additions, deletions, or replacements of sequence information in the "Sequence Listing XML" as required by 37 CFR 1.835(b)(3);

• A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.835(b)(4);

• A statement that the "Sequence Listing XML" includes no new matter in accordance with 37 CFR 1.835(b)(5); and

• A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 inserting the required incorporation by reference paragraph as required by 37 CFR 1.835(b)(2), consisting of:

o A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

o A copy of the amended specification without markings (clean version); and

o A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. Examiner needs to identify the sequences missing from the "Sequence Listing XML" that were found in the claims, drawing or specification in <1>.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.24.26 Missing, Defective, or Incomplete Incorporation by Reference Paragraph

Specific deficiency - The incorporation by reference paragraph required by 37 CFR 1.834(c)(1), 37 CFR 1.835(a)(2), or 1.835(b)(2) is missing, defective or incomplete.

Required response - Applicant must:

• Provide a substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 inserting the required incorporation by reference paragraph, consisting of:

• A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

• A copy of the amended specification without markings (clean version); and

• A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where incorporation by reference paragraph is missing, defective or incomplete. The examiner should clearly specify how any defective incorporation by reference paragraphs is non-compliant, if applicable.

4. The form paragraph maybe used when the incorporation by reference paragraphs provides erroneous information on the size of the sequence listing XML file, recites the size in kilobytes (KB) instead of bytes, and/or the file name for the sequence listing XML contains an error.

5. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.25.26 Amendment Missing Statement of No New Matter

Specific deficiency - The "Sequence Listing XML" has not been entered into the application because the required statement of no new matter, in accordance with 37 CFR 1.835(a)(4) or 37 CFR 1.835(b)(5), is missing.

Required response - Applicant must submit a statement that the "Sequence Listing XML," identified by the date the "Sequence Listing XML" was filed, includes no new matter.

Examiner Note:

1. This form paragraph should only be used where a compliant "Sequence Listing XML" is not required for examination of the application.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where the statement of no new matter is missing.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.26.26 Amendment Missing Statement of Support

Specific deficiency - The "Sequence Listing XML" has not been entered into the application because the required statement of support for the "Sequence Listing XML" in the application as filed, in accordance with 37 CFR 1.835(a)(3) or 1.835(a)(4), is missing.

Required response - Applicant must submit a proper statement that indicates the basis for the "Sequence Listing XML," with specific references to particular parts of the application as originally filed (specification, claims, drawings) for added, deleted and/or modified sequence data.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where the statement that indicates the basis for the "Sequence Listing XML" is missing.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.27.26 "Sequence Listing XML" contains foreign language text

Specific deficiency - This application fails to comply with the requirements of 37 CFR 1.831-1.835 because the language-dependent free text elements within the "Sequence Listing XML" are not in the English language, as required by 37 CFR 1.833(b)(3).

Required response - As required by 37 CFR 1.835(d)(2), applicant must provide a translated "Sequence Listing XML" part of the disclosure in compliance with 37 CFR 1.831-1.834.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where the "Sequence Listing XML" contains foreign text in the language-dependent free text elements.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.28.26 "Sequence Listing XML" bibliographic information does not match application

Specific deficiency - The "Sequence Listing XML" submitted in the present application does not appear to be the correct sequence listing file. One or more of the following identifying information: invention title, applicant file reference, applicant name, inventor name, or earliest priority application do not match those in the official application.

Required response - Applicant must provide:

• A replacement "Sequence Listing XML" part of the disclosure, as described above in item 1. or 2., as well as

o A statement that identifies the location of all additions, deletions, or replacements of sequence information in the replacement "Sequence Listing XML" as required by 37 CFR 1.835(b)(3);

o A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.835(b)(4),

o A statement that the replacement "Sequence Listing XML" includes no new matter in accordance with 37 CFR 1.835(b)(5), and

o A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 inserting the required incorporation by reference paragraph as required by 37 CFR 1.835(b)(2), consisting of:

A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

A copy of the amended specification without markings (clean version); and

A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where there is a discrepancy between the "Sequence Listing XML" bibliographic information and the application information in an application filed under 35 U.S.C. 111(a), if the application is a national stage, then the examiner should contact IPLA.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.29.26 Amendment Missing Statement of Location of Additions, Deletions or Replacements of Sequence Information

Specific deficiency - The "Sequence Listing XML" has not been entered into the application because the required statement that identifies the location of all additions, deletions or

replacements of the sequence information relative to the replaced "Sequence Listing XML" is missing. See 37 CFR 1.835(b)(3).

Required response - Applicant must provide a proper statement identifying all changes in the replacement "Sequence Listing XML" relative to the replaced sequence data.

Examiner Note:

1. This form paragraph should only be used where compliance with 37 CFR 1.831-1.835 is not required for examination of the application, or where a lack of compliance with 37 CFR 1.831-1.835 arose during prosecution.

2. This form paragraph must be preceded by form paragraph 24.17.26.

3. This form paragraph should be used for an application where the replacement "Sequence Listing XML" is not accompanied by a statement that indicates the location of all additions, deletions or replacements in the sequence information.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

2417 Helpful Hints for Sequence Rules Compliance under WIPO Standard ST.26 [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

The Office has the following non-exhaustive tips regarding sequence rules in compliance with WIPO Standard ST.26, as incorporated by reference in 37 CFR 1.839 and implemented by 37 CFR 1.831-1.835.

—Compliance is not required for establishing a filing date.

A sequence listing submitted with a new application in compliance with 37 CFR 1.831-1.834 is considered part of the specification and thus contributes to disclosure of the invention for the purpose of 35 U.S.C. 112.

—Compliance is not per se a new matter issue. See MPEP § 2414.02 for adding initial sequence listings under 37 CFR 1.835(a)(2)-(4) and MPEP § 2414.03 for amendments adding to, deleting from, or replacing sequence information under 37 CFR 1.835(b)(2)-(4).

Amino acid sequences required to be included in a "Sequence Listing XML" include those comprising D-amino acids, modified amino acids and linear portions of branched sequences.

Nucleotide sequences required to be included in a "Sequence Listing XML" include those comprising modified nucleotides, nucleotide analogs and linear portions of branched sequences.

Sequences of <10 nucleotides and <4 amino acids are prohibited in a "Sequence Listing XML."

Uracil is represented by "t" in RNA and DNA, and is considered to be a modified nucleotide that requires additional annotation when present in a DNA sequence.

Thymine is represented by "t" in RNA and DNA, and is considered to be a modified nucleotide that requires additional annotation when present in an RNA sequence.

Amino acid representations are each one letter.

Each sequence disclosed in the "Sequence Listing XML" must have its sequence identification number (identifier) associated with "SEQ ID NO:" or the like designator next to each occurrence within the disclosure, i.e. SEQ ID NO: X.

Variations of "SEQ ID NO:" format (e.g. "Seq Id No.") are permitted. Parentheses are NOT required; the sequence identifier, however, must be clearly separated from the sequence shown in the specification.

Sequences in figures can have "SEQ ID NO:" or the like with the sequence identifier in the figure itself OR in the Brief Description of the Drawings, but it does not have to be in both places.

Each "Sequence Listing XML" filed in a 111(a) application must be accompanied by an incorporation by reference statement that includes the name of the

file, the date of creation, and the size of the file in bytes.

A "Sequence Listing XML" filed along with a new application need not be accompanied by any statement about new matter.

An added "Sequence Listing XML" filed after the filing date of the application must be accompanied by a statement: • indicating that no new matter has been introduced; AND • identifying the support for each new sequence and sequence-associated data within the application as filed.

An amended "Sequence Listing XML," including one submitted to address a defective sequence listing (validated as CRFD), must be accompanied by a statement: • indicating that no new matter has been introduced; • identifying the locations of additions, deletions or replacements of sequence information relative to the replaced listing; AND • identifying the support for the additions, deletions or replacements of sequence information within the application as filed.

The international filing date (PCT filing date) is controlling for submission of a sequence listing in compliance with WIPO Standard ST.26 in 35 U.S.C. 371 (national stage) applications.

A sequence listing as an XML file in compliance with WIPO Standard ST.26 that is part of the description of an international (PCT) application and transmitted by the International Bureau of WIPO to the USPTO upon entry into the national phase does not require amendment of the specification of the application to add a statement about incorporation by reference, and no statement about new matter is needed.

A sequence listing as an XML file in compliance with WIPO Standard ST.26 that is part of the description of an international (PCT) application and transmitted by the International Bureau of WIPO to the USPTO upon entry into the national phase does not need to be amended to update bibliographic data, such as the 371 application number.

In cases where no sequence is present in a "Sequence Listing XML" for a sequence identification number, i.e. an intentionally skipped sequence, "000" is entered as the sequence.

Sequences that are prohibited from inclusion in a "Sequence Listing XML" may be presented within the description, but should not be associated with a sequence identification number.

2418 WIPO Sequence Tool [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

To comply with rules based on WIPO Standard ST.26 (i.e., 37 CFR 1.831 - 1.835), patent applicants can generate a compliant "Sequence Listing XML" using WIPO Sequence, a desktop application developed by WIPO and adopted by WIPO member states, including the US. WIPO Sequence has two functions: An authoring function and a validation function. Patent applicants are able to author and validate their "Sequence Listing XML" using WIPO Sequence to comply with the requirements of WIPO Standard ST.26. A "Sequence Listing XML" that passes validation by WIPO Sequence will be accepted by all the IPOs of the WIPO member states, although translation may be needed. Thus, the burden of generating a WIPO Standard ST.26 compliant sequence listing in XML file format 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**.

The first step to generate a new sequence listing using WIPO Sequence is to create a project. The user saves (authors) patent application data and biological sequence data in the project, validates the project to ensure all required sequence-associated information is present, and generates a "Sequence Listing XML" in compliance with 37 CFR 1.831 - 1.834 (complying with WIPO Standard ST.26 in 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, WIPO Standard ST.25 sequence listing text files, raw data files, multi-sequence format files, and FASTA format files. The entry of sequence-associated data, such as feature keys, qualifiers, and organism names, are facilitated by their selection from drop-down lists, simplifying the creation of a sequence listing. Applicant and inventor names, as well as custom organism names, can be stored in WIPO Sequence for easy access from other projects. 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 XML." This human readable version must not be submitted in a patent application.

WIPO Sequence includes an integrated validation function that will alert users to most errors in a project (i.e., "Sequence Listing XML" 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 XML." While the validation function will alert a user to most errors in a project or "Sequence Listing XML," 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 XML." The WIPO Sequence validation function can also be used to validate a "Sequence Listing XML" that was previously generated. As for validation of a project, the validation of a "Sequence Listing XML" generates a report that clearly lists every detected error, the location of the error, and the detected value of the error in the "Sequence Listing XML." It is highly recommended that a "Sequence Listing XML" produced by WIPO Sequence be validated following its generation.

To ensure that IPOs can validate applicant-submitted sequence listings generated with WIPO Sequence, WIPO developed a Standard ST.26 sequence listing validation tool, WIPO Sequence Validator. WIPO Sequence Validator is for use by IPOs to verify that filed sequence listings comply with WIPO Standard ST.26 . WIPO Sequence Validator is synchronized with the validation function in the WIPO Sequence tool such that WIPO Sequence Validator applies the same validation rules as WIPO Sequence. The USPTO has integrated WIPO Sequence Validator into its internal IT systems (i.e., CRFValidator). Therefore, filers will have a greater level of confidence that a "Sequence Listing XML" 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.

A "Sequence Listing" in WIPO Standard ST.25 format cannot automatically be converted into WIPO Standard ST.26 format primarily because certain data elements required for a "Sequence Listing XML" compliant with WIPO Standard ST.26 are not present in WIPO Standard ST.25 sequence listing. Therefore, conversion of a "Sequence Listing" in WIPO Standard ST.25 format to WIPO Standard ST.26 format necessarily requires additional input from the applicant. WIPO Sequence, supplemented by significant guidance in Annex VI and Annex VII of WIPO Standard ST.26, will help applicants to accomplish this task without added or deleted subject matter. Users can import a WIPO 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" in WIPO Standard ST.25 format into a WIPO Standard ST.26 "Sequence Listing XML" format.

A sample sequence listing (XML file) specimen can be found in WIPO Standard ST.26, Annex III and at **www.wipo.int/standards/en/xml_material/st26/st26-annex-iii-sequence-listing-specimen.xml**.

2419 Publishing of Patent Grants and Patent Application Publications with a "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

At the USPTO, the Sequence Listing Information Control (SLIC) system is responsible for processing of a "Sequence Listing XML" and exporting the sequence data for publication and grant. Because a "Sequence Listing XML" exists only in .xml file format, the USPTO uses a style sheet to transform a "Sequence Listing XML" file of less than 1 GB in size into an ASCII text file which will be used to present the sequence information in a format that is more easily read.

2419.01 Patent Grants and Patent Application Publications Containing a Non-Lengthy "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

Upon transformation of the .xml file using the style sheet to an ASCII text file, any ASCII text file produced by the USPTO that is under 600KB will be published as part of the specification.

2419.02 Patent Grants and Patent Application Publications Containing a Lengthy "Sequence Listing XML" [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

Upon transformation of the .xml file using the style sheet to an ASCII text file, any ASCII text file produced by the USPTO that is 600KB or more, but under 1GB will be processed for separate publication on Publication Site for Issued and Published Sequences (PSIPS). In place of such lengthy "Sequence Listing", the patent or patent application publication specification will show a page-wide SEQUENCE LISTING statement similar to the example shown below:

SEQUENCE LISTING

The patent contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form from the USPTO website (**https://seqdata.uspto.gov**). An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

2419.03 Patent Grants and Patent Application Publications Containing a "Sequence Listing XML" of 1GB or Larger [R-01.2024]

[Editor Note: This section is applicable to all applications with a filing date, or, for national phase applications, an international filing date, on or after July 1, 2022, having disclosure of one or more nucleotide and/or amino acid sequences as defined in 37 CFR 1.831(b).]

Where a "Sequence Listing XML" is 1GB or larger, exception processing for the listing will be required. That is, the "Sequence Listing XML" cannot be published as part of the specification, nor can it be published on PSIPS, instead it will be published on the USPTO Bulk Data Storage System (BDSS). No transformation of the .xml file into an ASCII text file will occur. In place of the 1GB or larger "Sequence Listing XML," the patent or patent application publication specification will show the page-wide statement similar to the one shown below:

SEQUENCE LISTING

The patent contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form

from the USPTO website **https://bulkdata.uspto.gov**. An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

2420 The Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosures under World Intellectual Property Office Standard ST.25 (WIPO ST.25) - the Sequence Rules [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412- 2419 for applications filed on or after July 1, 2022.]

The sequence rules (37 CFR 1.821 -1.825) set forth the requirements pertaining to patent applications containing nucleotide sequence and/or amino acid sequence disclosures filed under WIPO ST.25. See also MPEP § 608.05 and § 2422.03.

See PCT Rule 5 and Rule 13ter, and MPEP § 1823.02 and § 2422, for the requirements under the PCT for international applications subject to WIPO ST.25 requirements and that disclose nucleic acid or amino acid sequences.

2421 Overview of the Sequence Rules under WIPO ST.25 [R-07.2022]

2421.01 Definition of "Sequence Listing" and Computer Readable Form (CRF) [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412- 2419 for applications filed on or after July 1, 2022.]

The sequence rules (37 CFR 1.821 -1.825) require the use of standard symbols and a standard format for submitting sequence data in most patent applications that disclose nucleic acid or amino acid sequences. For purposes of the sequence rules and the discussion in MPEP Chapter 2400, the phrase "disclose(d) (or disclosure(s) of) nucleic acid or amino acid sequences" is intended to refer to those nucleic acid or amino acid sequences that are described in the patent application by enumeration of their residues and that meet the length thresholds of 37 CFR 1.821(a).

37 CFR 1.821(c) requires that applications containing disclosures of nucleotide and/or amino acid sequences that fall within the definitions of 37 CFR 1.821(a) contain, as a separate part, a disclosure of the nucleotide and/or amino acid sequences, and associated information, using the format and symbols that are set forth in 37 CFR 1.822 and 37 CFR 1.823. This separate part of the disclosure is referred to as the "Sequence Listing". See MPEP § 2422.03 for additional information.

37 CFR 1.821(e) requires a separate computer readable form (CRF) of the "Sequence Listing" under certain circumstances. The CRF is an ASCII plain text file of the sequence information relating to the disclosure of the nucleotide and/or amino acid sequences using the symbols, format, content and file requirements of 37 CFR 1.822. See 37 CFR 1.824. The CRF may be submitted on read-only optical disc or may be submitted as an ASCII plain text file via the Office patent electronic filing system, pursuant to 37 CFR 1.821. The requirements regarding the CRF are discussed in detail in MPEP § 2422.04.

2421.02 Summary of the Requirements of the Sequence Rules [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

The sequence rules define a set of symbols and procedures that are both mandatory and the only way that an applicant is permitted to describe information in the "Sequence Listing" about a nucleotide or amino acid sequence that falls within the definitions used in the rules. Thus, 37 CFR 1.821 defines a

"sequence" and a "Sequence Listing" for the purpose of the rules, the requirements for specific symbols, and formats for the "Sequence Listing", the requirement for a computer readable form (CRF) of the "Sequence Listing" and the deadlines for complying with the requirements. 37 CFR 1.822 to 37 CFR 1.824 set forth detailed descriptions of the requirements that are mandatory for the presentation of sequence data, and 37 CFR 1.825 sets forth procedures that are available to an applicant in the event that amendments to add or replace a "Sequence Listing", including any necessary CRF of the "Sequence Listing" become necessary.

The sequence rules embrace all unbranched nucleotide sequences with ten or more nucleotide bases and all unbranched, non-D amino acid sequences with four or more amino acids, provided that there are at least 10 "specifically defined" nucleotides or 4 "specifically defined" nucleotides or amino acids. The rules apply to all sequences in a given application, whether claimed or not. All such sequences are relevant for the purposes of building a comprehensive database and properly assessing prior art. It is therefore essential that all sequences, whether only disclosed or also claimed, be included in the database.

2421.03 Notification of a Failure to Comply [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

With respect to the Office's determination of compliance with the sequence rules and the opportunities afforded applicants to satisfy the requirements of the rules, applicants will be notified of easily detectable deficiencies early in the application process. Applicants whose computer readable forms (which include a "Sequence Listing" submitted as an ASCII plain text file under 37 CFR 1.821(c)(1) and separate CRFs submitted under 37 CFR 1.821(e)) are not readable, or are missing mandatory elements, will be notified shortly after receipt of the application by the Office. See 37 CFR 1.52(e)(9), 1.821(g), and 1.821(h).

The Office of Patent Application Processing (OPAP) will mail a notice to applicant listing the requirements that have not been met and setting a two month time period within which to comply with the sequence rules, 37 CFR 1.821 - 1.825. Failure to comply with these requirements will result in abandonment of the application under 37 CFR 1.821(g). Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136. Applications will not be subject to examination on the merits until a compliant "Sequence Listing" has been submitted, as jurisdiction remains with Office of Patent Application Processing until compliance with the Notice to Comply With Requirements For Patent Applications Containing Nucleotide Sequence And/Or Amino Acid Sequence Disclosures.

Patent applications filed under 35 U.S.C. 111 on or after December 18, 2013, and international patent applications in which the national stage commenced under 35 U.S.C. 371 on or after December 18, 2013, may be subject to reductions in patent terms adjustment pursuant to 37 CFR 1.704(c)(13) if they are not in condition for examination within eight months from the filing date or date of commencement, respectively. "In condition for examination" includes compliance with 37 CFR 1.821 -1.825 (see 37 CFR 1.704(f)). Deficiencies of a more sophisticated nature will likely only be detected by the examiner to whom the application is assigned. Applicant will be notified of any errors or inconsistencies detected by the examiner in the next Office action.

A notification of a failure to comply with the sequence rules will be accompanied by a report of errors associated with any submitted computer readable form (the 37 CFR 1.821(c)(1), the 1.821(e)(1) or 1.821(e)(2) submission). Any inquiries regarding a specific computer readable form that has been processed by the Office should be directed to the Sequence Systems Service Center (S3C) of the Scientific and Technical Information Center. The S3C manages the Sequence Help Desk, and responds to all questions or concerns, from both inside and outside the USPTO, related to computer readable

format sequence listing submissions. The S3C may be contacted at 571-272-2510 or via email at STIC-SSSCHelpdesk@uspto.gov.

2421.04 [Reserved]

2422 Nucleotide and/or Amino Acid Sequence Disclosures in Patent Applications Subject to WIPO ST.25 [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.821 Nucleotide and/or amino acid sequence disclosures in patent applications.

(a) Nucleotide and/or amino acid sequences, as used in §§ 1.821 through 1.825, are interpreted to mean an unbranched sequence of 4 or more amino acids or an unbranched sequence of 10 or more nucleotides. Branched sequences are specifically excluded from this definition. Sequences with fewer than four specifically defined nucleotides or amino acids are specifically excluded from this section. "Specifically defined" means those amino acids other than "Xaa" and those nucleotide bases other than "n," defined in accordance with Appendices A through F to this subpart. Nucleotides and amino acids are further defined as follows:

(1) *Nucleotides:* Nucleotides are intended to embrace only those nucleotides that can be represented using the symbols set forth in Appendix A to this subpart. Modifications (*e.g.*, methylated bases) may be described as set forth in Appendix B to this subpart but shall not be shown explicitly in the nucleotide sequence.

(2) *Amino acids:* Amino acids are those L-amino acids commonly found in naturally occurring proteins and are listed in appendix C to this subpart. Those amino acid sequences containing D-amino acids are not intended to be embraced by this definition. Any amino acid sequence that contains post-translationally modified amino acids may be described as the amino acid sequence that is initially translated using the symbols shown in appendix C to this subpart, with the modified positions (*e.g.*, hydroxylations or glycosylations) being described as set forth in appendix D to this subpart, but these modifications shall not be shown explicitly in the amino acid sequence. Any peptide or protein that can be expressed as a sequence using the symbols in appendix C to this subpart, in conjunction with a description in the Feature section, to describe, for example, modified linkages, cross links and end caps, non-peptidyl bonds, etc., is embraced by this definition.

Note 1 to paragraph (a): Appendices A through F to this subpart contain Tables 1– 6 of the World Intellectual Property Organization (WIPO) Handbook on Industrial Property Information and Documentation, Standard ST.25: Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings in Patent Applications (2009).

(b) Patent applications which contain disclosures of nucleotide and/or amino acid sequences, in accordance with the definition in paragraph (a) of this section, shall, with regard to the manner in which the nucleotide and/or amino acid sequences are presented and described, conform exclusively to the requirements of §§ 1.821 through 1.825.

(c) Patent applications that contain disclosures of nucleotide and/or amino acid sequences, as defined in paragraph (a) of this section, must contain a "Sequence Listing," which is a separate part of the specification containing each of those nucleotide and/or amino acid sequences and associated information using the symbols and format in accordance with the requirements of §§ 1.822 and 1.823. The "Sequence Listing" must be submitted as follows, except for a national stage entry under § 1.495(b)(1), where the "Sequence Listing" has been previously communicated by the International Bureau or originally filed in the United States Patent and Trademark Office and complies with Patent Cooperation Treaty (PCT) Rule 5.2:

(1) As an ASCII plain text file, in compliance with § 1.824, submitted via the USPTO patent electronic filing system or on a read-only optical disc under § 1.52(e), accompanied by an incorporation by reference statement of the ASCII plain text file, in a separate paragraph of the specification, in accordance with § 1.77(b)(5);

(2) As a PDF file via the USPTO patent electronic filing system; or

(3) On physical sheets of paper.

(d) Where the description or claims of a patent application discuss a sequence that is set forth in the "Sequence Listing," in accordance with paragraph (c) of this section, reference must be made to the sequence by use of the sequence identifier (§ 1.823(a)(5)), 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 (§ 1.823(a)(5)), 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 (§ 1.823(a)(5)) in the Brief Description is clear.

(e)(1) If the "Sequence Listing" under paragraph (c) of this section is submitted in an application filed under 35 U.S.C. 111(a) as a PDF file (§ 1.821(c)(2)) via the USPTO patent electronic filing system or on physical sheets of paper (§ 1.821(c)(3)), then the following must be submitted:

(i) A CRF of the "Sequence Listing," in accordance with the requirements of § 1.824; and

(ii) A statement that the sequence information contained in the CRF submitted under paragraph (e)(1)(i) of this section is identical to the sequence information contained in the "Sequence Listing" under paragraph (c) of this section.

(2) If the "Sequence Listing" under paragraph (c) of this section in an application submitted under 35 U.S.C. 371 is a PDF file (paragraph (c)(2) of this section) or on physical sheets of paper (paragraph (c)(3) of this section), and not also as an

ASCII plain text file, in compliance with § 1.824 (paragraph (c)(1) of this section), then the following must be submitted:

(i) A CRF of the "Sequence Listing," in accordance with the requirements of § 1.824; and

(ii) A statement that the sequence information contained in the CRF submitted under paragraph (e)(2)(i) of this section is identical to the sequence information contained in the "Sequence Listing" under paragraph (c)(2) or (3) of this section.

(3) If a "Sequence Listing" in ASCII plain text format, in compliance with § 1.824, has not been submitted for an international application under the PCT, and that application contains disclosures of nucleotide and/or amino acid sequences, as defined in paragraph (a) of this section, and is to be searched by the United States International Searching Authority or examined by the United States International Preliminary Examining Authority, then the following must be submitted:

(i) A CRF of the "Sequence Listing," in accordance with the requirements of § 1.824;

(ii) The late furnishing fee for providing a "Sequence Listing" in response to an invitation, as set forth in § 1.445(a)(5); and

(iii) A statement that the sequence information contained in the CRF, submitted under paragraph (e)(3)(i) of this section, does not go beyond the disclosure in the international application as filed, or a statement that the information recorded in the ASCII plain text file, submitted under paragraph (e)(3)(i) of this section, is identical to the sequence listing contained in the international application as filed, as applicable.

(4) The CRF may not be retained as a part of the patent application file.

(f) [reserved]

(g) If any of the requirements of paragraphs (b) through (e) of this section are not satisfied at the time of filing under 35 U.S.C. 111(a) or at the time of entering the national stage 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. Any amendment to add or replace a "Sequence Listing" and CRF copy thereof in reply to a requirement under this paragraph must be submitted in accordance with the requirements of § 1.825.

(h) If any of the requirements of paragraph (e)(3) of this section are not satisfied at the time of filing an international application under the PCT, and 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. Where a "Sequence Listing" under PCT Rule 13 *ter* is provided in reply to a requirement under this paragraph, it must be accompanied by a statement that the information recorded in the ASCII plain text file under paragraph (e)(3)(i) of this section is identical to the sequence listing contained in the international application as filed, or does not go beyond the disclosure in the international application as filed, as applicable. It must also be accompanied by the late furnishing fee, as set forth in § 1.445(a)(5). If the applicant fails to timely provide the required CRF, the United States International Searching Authority shall search only to the extent that a meaningful search can be performed without the CRF, and the United States International Preliminary Examining Authority shall examine only to the extent that a meaningful examination can be performed without the CRF.

I. APPENDICES A-F REFERENCED IN 37 CFR 1.821 AND 1.822

37 CFR 1.821 and 37 CFR 1.822 reference Appendices A-F, which contain Tables 1–6 of the World Intellectual Property Organization (WIPO) Handbook on Industrial Property Information and Documentation, Standard ST.25: Standard for Nucleotide and Amino Acid Sequence Listings in Patent Applications (2009) (hereinafter WIPO Standard ST.25 (2009)). Appendices A-F are reproduced below. The current version of WIPO Standard ST.25 is available online at **www.wipo.int /export/sites/ www/standards/en/pdf/03-25-01.pdf**.

Appendix A provides that the bases of a nucleotide sequence should be represented using the following one-letter symbol for nucleotide sequence characters:

Appendix A to Subpart G of Part 1 - List of Nucleotides

Symbol	Meaning	Origin of designation
a	a	*a*denine.
g	g	*g*uanine.
c	c	*c*ytosine.
t	t	thymine.
u	u	*u*racil.
r	g or a	purine.
y	t/u or c	p *y*rimidine.
m	a or c	a *m*ino.
k	g or t/u	*k*eto.
s	g or c	strong interactions 3H-bonds.
w	a or t/u	*w*eak interactions 2H-bonds.
b	g or c or t/u	not a.
d	a or g or t/u	not c.
h	a or c or t/u	not g.
v	a or g or c	not t, not u.
n	a or g or c or t/u, unknown, or other	a *n*y.

Appendix B provides that modified bases may be represented as the corresponding unmodified bases in the sequence itself, if the modification is further described in numeric identifier <223> of the Feature section of the "Sequence Listing". The symbols from the list below may be used in the description (i.e., the specification and drawing, or in the Feature section of the "Sequence Listing") but these symbols may not be used in the sequence itself. Modifications not listed in Appendix B may also be represented as the corresponding unmodified base in the sequence itself, and the modification should be described using its full chemical name in the Feature section of the "Sequence Listing".

Appendix B to Subpart G of Part 1 - List of Modified Nucleotides

Symbol	Meaning
ac4c	4-acetylcytidine.
chm5u	5-(carboxyhydroxymethyl)uridine.
cm	2'-O-methylcytidine.
cmnm5s2u	5-carboxymethylaminomethyl-2-thiouridine.
cmnm5u	5-carboxymethylaminomethyluridine.
d	dihydrouridine.
fm	2'-O-methylpseudouridine.
gal q	beta, D-galactosylqueuosine.
gm	2'-O-methylguanosine.
i	inosine.
i6a	N6-isopentenyladenosine.
m1a	1-methyladenosine.
m1f	1-methylpseudouridine.
m1g	1-methylguanosine.
m1i	1-methylinosine.
m22g	2,2-dimethylguanosine.
m2a	2-methyladenosine.
m2g	2-methylguanosine.
m3c	3-methylcytidine.
m5c	5-methylcytidine.
m6a	N6-methyladenosine.
m7g	7-methylguanosine.
mam5u	5-methylaminomethyluridine.
mam5s2u	5-methoxyaminomethyl-2-thiouridine.
man q	beta, D-mannosylqueuosine.
mcm5s2u	5-methoxycarbonylmethyl-2-thiouridine.
mcm5u	5-methoxycarbonylmethyluridine.
mo5u	5-methoxyuridine.
ms2i6a	2-methylthio-N6-isopentenyladenosine.
ms2t6a	N-((9-beta-D-ribofuranosyl-2-methylthiopurine-6-yl)carbamoyl)threonine.
mt6a	N-((9-beta-D-ribofuranosylpurine-6-yl)N-methylcarbamoyl)threonine.
mv	uridine-5-oxyacetic acid-methylester.
o5u	uridine-5-oxyacetic acid.
osyw	wybutoxosine.
p	pseudouridine.
q	queuosine.
s2c	2-thiocytidine.
s2t	5-methyl-2-thiouridine.
s2u	2-thiouridine.
s4u	4-thiouridine.
t	5-methyluridine.
t6a	N-((9-beta-D-ribofuranosylpurine-6-yl)-carbamoyl)threonine.
tm	2'-O-methyl-5-methyluridine.
um	2'-O-methyluridine.
yw	wybutosine.
x	3-(3-amino-3-carboxy-propyl)uridine, (acp3)u.

Appendix C provides that the amino acids should be represented using the following three-letter symbols with the first letter as a capital.

Appendix C to Subpart G of Part 1 - List of Amino Acids

Symbol	Meaning
Ala	Alanine.
Cys	Cysteine.
Asp	Aspartic Acid.
Glu	Glutamic Acid.
Phe	Phenylalanine.
Gly	Glycine.
His	Histidine.
Ile	Isoleucine.

Symbol	Meaning
Lys	Lysine.
Leu	Leucine.
Met	Methionine.
Asn	Asparagine.
Pro	Proline.
Gln	Glutamine.
Arg	Arginine.
Ser	Serine.
Thr	Threonine.
Val	Valine.
Trp	Tryptophan.
Tyr	Tyrosine.
Asx	Asp or Asn.
Glx	Glu or Gln.
Xaa	unknown or other.

Appendix D provides that modified and unusual amino acids may be represented as the corresponding unmodified amino acids in the sequence itself if the modification is further described in numeric identifier <223> of the Feature section of the "Sequence Listing". The symbols from the list below may be used in the description (i.e., the specification and drawings, or in the Feature section of the "Sequence Listing") but these symbols may not be used in the sequence itself. Modifications not listed in Appendix D may also be represented as the corresponding unmodified amino acid in the sequence itself, and the modification should be described using its full chemical name in the Feature section of the "Sequence Listing".

Appendix D to Subpart G of Part 1 - List of Modified and Unusual Amino Acids

Symbol	Meaning
Aad	2-Aminoadipic acid.
bAad	3-Aminoadipic acid.
bAla	beta-Alanine, beta-Aminopropionic acid.
Abu	2-Aminobutyric acid.
4Abu	4-Aminobutyric acid, piperidinic acid.
Acp	6-Aminocaproic acid.
Ahe	2-Aminoheptanoic acid.
Aib	2-Aminoisobutyric acid.
bAib	3-Aminoisobutyric acid.
Apm	2-Aminopimelic acid.
Dbu	2,4 Diaminobutyric acid.
Des	Desmosine.
Dpm	2,2'-Diaminopimelic acid.
Dpr	2,3-Diaminopropionic acid.
EtGly	N-Ethylglycine.
EtAsn	N-Ethylasparagine.
Hyl	Hydroxylysine.
aHyl	allo-Hydroxylysine.
3Hyp	3-Hydroxyproline.
4Hyp	4-Hydroxyproline.
Ide	Isodesmosine.
aIle	allo-Isoleucine.
MeGly	N-Methylglycine, sarcosine.
MeIle	N-Methylisoleucine.
MeLys	6-N-Methyllysine.
MeVal	N-Methylvaline.
Nva	Norvaline.
Nle	Norleucine.
Orn	Ornithine.

Appendix E provides for feature keys related to nucleotide sequences.

Appendix E to Subpart G of Part 1 - List of Feature Keys Related to Nucleotide Sequences

Key	Description
allele	a related individual or strain contains stable, alternative forms of the same gene, which differs from the presented sequence at this location (and perhaps others).
attenuator	(1) region of DNA at which regulation of termination of transcription occurs, which controls the expression of some bacterial operons; (2) sequence segment located

Key	Description
	between the promoter and the first structural gene that causes partial termination of transcription.
C_region	constant region of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains; includes one or more exons depending on the particular chain.
CAAT_signal	CAAT box; part of a conserved sequence located about 75 bp upstream of the start point of eukaryotic transcription units which may be involved in RNA polymerase binding; consensus=GG (C or T) CAATCT.
CDS	coding sequence; sequence of nucleotides that corresponds with the sequence of amino acids in a protein (location includes stop codon); feature includes amino acid conceptual translation.
conflict	independent determinations of the "same" sequence differ at this site or region.
D-loop	displacement loop; a region within mitochondrial DNA in which a short stretch of RNA is paired with one strand of DNA, displacing the original partner DNA strand in this region; also used to describe the displacement of a region of one strand of duplex DNA by a single stranded invader in the reaction catalyzed by RecA protein.
D-segment	diversity segment of immunoglobulin heavy chain, and T-cell receptor beta chain.
enhancer	a cis-acting sequence that increases the utilization of (some) eukaryotic promoters, and can function in either orientation and in any location (upstream or downstream) relative to the promoter.
exon	region of genome that codes for portion of spliced mRNA; may contain 5'UTR, all CDSs, and 3'UTR.
GC_signal	GC box; a conserved GC-rich region located upstream of the start point of eukaryotic transcription units which may occur in multiple copies or in either orientation; consensus=GGGCGG.
gene	region of biological interest identified as a gene and for which a name has been assigned.
iDNA	intervening DNA; DNA which is eliminated through any of several kinds of recombination.
intron	a segment of DNA that is transcribed, but removed from within the transcript by splicing together the sequences (exons) on either side of it.
J_segment	joining segment of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains.
LTR	long terminal repeat, a sequence directly repeated at both ends of a defined sequence, of the sort typically found in retroviruses.

Key	Description
mat_peptide	mature peptide or protein coding sequence; coding sequence for the mature or final peptide or protein product following post- translational modification; the location does not include the stop codon (unlike the corresponding CDS).
misc_binding	site in nucleic acid which covalently or non-covalently binds another moiety that cannot be described by any other Binding key (primer_bind or protein_bind).
misc_difference	feature sequence is different from that presented in the entry and cannot be described by any other Difference key (conflict, unsure, old_sequence, mutation, variation, allele, or modified_base).
misc_feature	region of biological interest which cannot be described by any other feature key; a new or rare feature.
misc_recomb	site of any generalized, site-specific or replicative recombination event where there is a breakage and reunion of duplex DNA that cannot be described by other recombination keys (iDNA and virion) or qualifiers of source key (/insertion_seq, /transposon, /proviral).
misc_RNA	any transcript or RNA product that cannot be defined by other RNA keys (prim_transcript, precursor_RNA, mRNA, 5'clip, 3'clip, 5'UTR, 3'UTR, exon, CDS, sig_peptide, transit_peptide, mat_peptide, intron, polyA_site, rRNA, tRNA, scRNA, and snRNA).
misc_signal	any region containing a signal controlling or altering gene function or expression that cannot be described by other Signal keys (promoter, CAAT_signal, TATA_signal, –35_signal, –10_signal, GC_signal, RBS, polyA_signal, enhancer, attenuator, terminator, and rep_origin).
misc_structure	any secondary or tertiary structure or conformation that cannot be described by other Structure keys (stem_loop and D-loop).
modified_base	the indicated nucleotide is a modified nucleotide and should be substituted for by the indicated molecule (given in the mod_base qualifier value).
mRNA	messenger RNA; includes 5' untranslated region (5'UTR), coding sequences (CDS, exon) and 3' untranslated region (3'UTR).
mutation	a related strain has an abrupt, inheritable change in the sequence at this location.
N_region	extra nucleotides inserted between rearranged immunoglobulin segments.
old_sequence	the presented sequence revises a previous version of the sequence at this location.
polyA_signal	recognition region necessary for endonuclease cleavage of an RNA transcript that is followed by polyadenylation; consensus=AATAAA.

Key	Description
polyA_site	site on an RNA transcript to which will be added adenine residues by post-transcriptional polyadenylation.
precursor_RNA	any RNA species that is not yet the mature RNA product; may include 5' clipped region (5'clip), 5' untranslated region (5'UTR), coding sequences (CDS, exon), intervening sequences (intron), 3' untranslated region (3'UTR), and 3' clipped region (3'clip).
prim_transcript	primary (initial, unprocessed) transcript; includes 5' clipped region (5'clip), 5' untranslated region (5'UTR), coding sequences (CDS, exon), intervening sequences (intron), 3' untranslated region (3'UTR), and 3' clipped region (3'clip).
primer_bind	non-covalent primer binding site for initiation of replication, transcription, or reverse transcription; includes site(s) for synthetic, for example, PCR primer elements.
promoter	region on a DNA molecule involved in RNA polymerase binding to initiate transcription.
protein_bind	non-covalent protein binding site on nucleic acid.
RBS	ribosome binding site.
repeat_region	region of genome containing repeating units.
repeat_unit	single repeat element.
rep_origin	origin of replication; starting site for duplication of nucleic acid to give two identical copies.
rRNA	mature ribosomal RNA; the RNA component of the ribonucleoprotein particle (ribosome) which assembles amino acids into proteins.
S_region	switch region of immunoglobulin heavy chains; involved in the rearrangement of heavy chain DNA leading to the expression of a different immunoglobulin class from the same B-cell.
satellite	many tandem repeats (identical or related) of a short basic repeating unit; many have a base composition or other property different from the genome average that allows them to be separated from the bulk (main band) genomic DNA.
scRNA	small cytoplasmic RNA; any one of several small cytoplasmic RNA molecules present in the cytoplasm and (sometimes) nucleus of a eukaryote.
sig_peptide	signal peptide coding sequence; coding sequence for an N-terminal domain of a secreted protein; this domain is involved in attaching nascent polypeptide to the membrane; leader sequence.
snRNA	small nuclear RNA; any one of many small RNA species confined to the nucleus; several of the snRNAs are involved in splicing or other RNA processing reactions.
source	identifies the biological source of the specified span of the sequence; this key is mandatory; every entry will have,

Key	Description
	as a minimum, a single source key spanning the entire sequence; more than one source key per sequence is permissible.
stem_loop	hairpin; a double-helical region formed by base-pairing between adjacent (inverted) complementary sequences in a single strand of RNA or DNA
STS	Sequence Tagged Site; short, single-copy DNA sequence that characterizes a mapping landmark on the genome and can be detected by PCR; a region of the genome can be mapped by determining the order of a series of STSs.
TATA_signal	TATA box; Goldberg-Hogness box; a conserved AT-rich septamer found about 25 bp before the start point of each eukaryotic RNA polymerase II transcript unit which may be involved in positioning the enzyme for correct initiation; consensus=TATA(A or T)A(A or T).
terminator	sequence of DNA located either at the end of the transcript or adjacent to a promoter region that causes RNA polymerase to terminate transcription; may also be site of binding of repressor protein.
transit_peptide	transit peptide coding sequence; coding sequence for an N-terminal domain of a nuclear-encoded organellar protein; this domain is involved in post-translational import of the protein into the organelle.
tRNA	mature transfer RNA, a small RNA molecule (75-85 bases long) that mediates the translation of a nucleic acid sequence into an amino acid sequence.
unsure	author is unsure of exact sequence in this region.
V_region	variable region of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains; codes for the variable amino terminal portion; can be made up from V_segments, D_segments, N_regions, and J_segments.
V_segment	variable segment of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains; codes for most of the variable region (V_region) and the last few amino acids of the leader peptide.
variation	a related strain contains stable mutations from the same gene (for example, RFLPs, polymorphisms, etc.) which differ from the presented sequence at this location (and possibly others).
3'clip	3'-most region of a precursor transcript that is clipped off during processing.
3'UTR	region at the 3' end of a mature transcript (following the stop codon) that is not translated into a protein.
5'clip	5'-most region of a precursor transcript that is clipped off during processing.

Key	Description
5'UTR	region at the 5' end of a mature transcript (preceding the initiation codon) that is not translated into a protein.
–10_signal	pribnow box; a conserved region about 10 bp upstream of the start point of bacterial transcription units which may be involved in binding RNA polymerase; consensus=TAtAaT.
–35_signal	a conserved hexamer about 35 bp upstream of the start point of bacterial transcription units; consensus=TTGACa [ ] or TGTTGACA [ ].

Appendix F provides for feature keys related to protein sequences.

Appendix F to Subpart G of Part 1-List of Feature Keys Related to Protein Sequences

Key	Description
CONFLICT	different papers report differing sequences.
VARIANT	authors report that sequence variants exist.
VARSPLIC	description of sequence variants produced by alternative splicing.
MUTAGEN	site which has been experimentally altered.
MOD_RES	post-translational modification of a residue.
ACETYLATION	N-terminal or other.
AMIDATION	generally at the C-terminal of a mature active peptide.
BLOCKED	undetermined N- or C-terminal blocking group.
FORMYLATION	of the N-terminal methionine.
GAMMA-CARBOXYGLUTAMIC ACID HYDROXYLATION	of asparagine, aspartic acid, proline, or lysine.
METHYLATION	generally of lysine or arginine.
PHOSPHORYLATION	of serine, threonine, tyrosine, aspartic acid or histidine.
PYRROLIDONE CARBOXYLIC ACID	N-terminal glutamate which has formed an internal cyclic lactam.
SULFATATION	generally of tyrosine.
LIPID	covalent binding of a lipidic moiety.
MYRISTATE	myristate group attached through an amide bond to the N-terminal glycine residue of the mature form of a protein or to an internal lysine residue.
PALMITATE	palmitate group attached through a thioether bond to a cysteine residue or through an ester bond to a serine or threonine residue.
FARNESYL	farnesyl group attached through a thioether bond to a cysteine residue.
GERANYL-GERANYL	geranyl-geranyl group attached through a thioether bond to a cysteine residue.

Key	Description
GPI-ANCHOR	glycosyl-phosphatidylinositol (GPI) group linked to the alpha- carboxyl group of the C-terminal residue of the mature form of a protein.
N-ACYL DIGLYCERIDE	N-terminal cysteine of the mature form of a prokaryotic lipoprotein with an amide- linked fatty acid and a glyceryl group to which two fatty acids are linked by ester linkages.
DISULFID	disulfide bond; the 'FROM' and 'TO' endpoints represent the two residues which are linked by an intra-chain disulfide bond; if the `FROM' and `TO' endpoints are identical, the disulfide bond is an interchain one and the description field indicates the nature of the cross-link.
THIOLEST	thiolester bond; the 'FROM' and 'TO' endpoints represent the two residues which are linked by the thiolester bond.
THIOETH	thioether bond; the 'FROM' and 'TO' endpoints represent the two residues which are linked by the thioether bond.
CARBOHYD	glycosylation site; the nature of the carbohydrate (if known) is given in the description field.
METAL	binding site for a metal ion; the description field indicates the nature of the metal.
BINDING	binding site for any chemical group (co- enzyme, prosthetic group, etc.); the chemical nature of the group is given in the description field.
SIGNAL	extent of a signal sequence (prepeptide).
TRANSIT	extent of a transit peptide (mitochondrial, chloroplastic, or for a microbody).
PROPEP	extent of a propeptide.
CHAIN	extent of a polypeptide chain in the mature protein.
PEPTIDE	extent of a released active peptide.
DOMAIN	extent of a domain of interest on the sequence; the nature of that domain is given in the description field.
CA_BIND	extent of a calcium-binding region.
DNA_BIND	extent of a DNA-binding region.
NP_BIND	extent of a nucleotide phosphate binding region; the nature of the nucleotide phosphate is indicated in the description field.
TRANSMEM	extent of a transmembrane region.
ZN_FING	extent of a zinc finger region.
SIMILAR	extent of a similarity with another protein sequence; precise information, relative to that sequence, is given in the description field.
REPEAT	extent of an internal sequence repetition.
HELIX	secondary structure: Helices, for example, Alpha-helix, 3(10) helix, or Pi- helix.
STRAND	secondary structure: Beta-strand, for example, Hydrogen bonded beta-strand, or Residue in an isolated beta-bridge.

Key	Description
TURN	secondary structure Turns, for example, H-bonded turn (3-turn, 4-turn, or 5-turn).
ACT_SITE	amino acid(s) involved in the activity of an enzyme.
SITE	any other interesting site on the sequence.
INIT_MET	the sequence is known to start with an initiator methionine.
NON_TER	the residue at an extremity of the sequence is not the terminal residue; if applied to position 1, this signifies that the first position is not the N- terminus of the complete molecule; if applied to the last position, it signifies that this position is not the C-terminus of the complete molecule; there is no description field for this key.
NON_CONS	non consecutive residues; indicates that two residues in a sequence are not consecutive and that there are a number of unsequenced residues between them.
UNSURE	uncertainties in the sequence; used to describe region(s) of a sequence for which the authors are unsure about the sequence assignment.

II. INTERNATIONAL AND FOREIGN APPLICATIONS

The requirements of 37 CFR 1.821 through 37 CFR 1.825 are the result of an effort to harmonize the USPTO requirements with international sequence listing requirements to the extent possible. The requirements of 37 CFR 1.821 through 37 CFR 1.825 substantially correspond to the requirements of WIPO Standard ST.25 (2009). However, the requirements of 37 CFR 1.821 through 37 CFR 1.825 are less stringent than the requirements of WIPO Standard ST.25 (2009). Thus, applicants who have filed or wish to file international applications or applications in countries that adhere to WIPO Standard ST.25 (2009) should be aware of the following requirements:

(A) The data in numeric identifier <221> must use selections from Tables 5 and 6 of WIPO Standard ST.25 (2009) to comply with that standard. The terms from these Tables are considered language neutral vocabulary;

(B) WIPO Standard ST.25 (2009), paragraph 24, requires a blank line between numeric identifiers in the sequence listing when the digit in the first or second position of the numeric identifier changes;

(C) Where the sequence listing forming part of the description of the international application contains free text, e.g., free text in numeric identifier <223>, any such free text shall be repeated in the main part of the description in the language thereof (PCT Rule 5.2(b)). It is recommended that the free text in the language of the main part of the description be put in a specific section of the description called “Sequence Listing Free Text”;

(D) A sequence listing filed after the international filing date is generally not considered to be part of the disclosure and usually will not be published as part of the international application publication (see PCT Article 34 and PCT Rules 26 and 91 for exceptions); and

(E) Paragraph 4(v) of WIPO Standard ST.25 (2009) requires an accompanying statement with the specific wording “the information recorded in electronic form furnished under PCT Rule 13ter is identical to the sequence listing **as contained in the international application**”.

With further regard to requirements (A) and (B), is noted that PatentIn Version 3.5.1 software (see MPEP § 2430) generates sequence listings that meet all of the requirements of WIPO Standard ST.25 (2009). Applicants should similarly be aware that filing requirements for sequence listings may differ between a national US application, a foreign application and an international application during international phase. For example, where an international application is filed in paper, the sequence listing part of the international application

must similarly be provided in paper. In addition, a copy of the sequence listing in ASCII plain text, to be used for the purpose of the international search (PCT Rule 13ter) must be filed on read-only optical disc or via the USPTO electronic filing system. Furthermore, in contrast to US national applications, a sequence listing filed with RO/US in ASCII plain text that is 300 MB or more in size is not subject to a size fee during the international phase of an international application.

2422.01 Nucleotide and/or Amino Acids Disclosures Requiring a "Sequence Listing" [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

I. LENGTH THRESHOLDS

37 CFR 1.821(a) presents a definition for "nucleotide and/or amino acid sequences." This definition sets forth limits, in terms of numbers of amino acids and/or numbers of nucleotides, at or above which compliance with the sequence rules is required. Nucleotide and/or amino acid sequences as used in 37 CFR 1.821 through 37 CFR 1.825 are interpreted to mean an unbranched sequence of four or more amino acids or an unbranched sequence of ten or more nucleotides. Branched sequences are specifically excluded from this definition. Sequences with fewer than ten specifically defined nucleotides or four specifically defined amino acids are specifically excluded from 37 CFR 1.821. "Specifically defined" means those amino acids other than "Xaa" and those nucleotide bases other than "n" defined in Appendices A-F to 37 CFR part 1, Subpart G (see MPEP § 2422(I)).

The limit of four or more amino acids was established for consistency with limits in place for industry database collections whereas the limit of ten or more nucleotides, while lower than certain industry database limits, was established to encompass those nucleotide sequences to which the smallest probe will bind in a stable manner.

II. REPRESENTATION OF NUCLEIC ACIDS AND AMINO ACIDS

37 CFR 1.821(a)(1) and 37 CFR 1.821(a)(2) present further definitions for those nucleotide and amino acid sequences that are intended to be embraced by the sequence rules. Situations in which the applicability of the rules is in issue will be resolved on a case-by-case basis.

Nucleotide sequences are further limited to those that can be represented by the symbols set forth in 37 CFR 1.822(b) and Appendices A and B to 37 CFR part 1, Subpart G (see MPEP § 2422(I)). The presence of other than typical 5' to 3' phosphodiester linkages in a nucleotide sequence does not render the rules inapplicable. For example, the Office does not want to exclude linkages of the type commonly found in naturally occurring nucleotides, e.g., eukaryotic end capped sequences.

Amino acid sequences are further limited to those in 37 CFR 1.822(b) and Appendices C and D to 37 CFR part 1, Subpart G (see MPEP § 2422(I)) and those L-amino acids that are commonly found in naturally occurring proteins. The presence of one or more D-amino acids in a sequence will exclude that sequence from the scope of the rules. Voluntary compliance is, however, encouraged in these situations; the symbol "Xaa" can be used to represent D-amino acids. The sequence rules embrace "[a]ny peptide or protein that can be expressed as a sequence using the symbols in Appendix C to 37 CFR part 1, Subpart G (see MPEP § 2422(I)) in conjunction with a description in the Feature section to describe, for example, modified linkages, cross links and end caps, non-peptidyl bonds, etc." 37 CFR 1.821(a)(2).

With regard to amino acid sequences, the use of the terms "peptide or protein" implies that the amino acids in a given sequence are linked by at least three consecutive peptide bonds. Accordingly, an amino acid sequence is not excluded from the scope of the rules merely due to the presence of a single non-peptidyl bond. If an amino acid sequence can be represented by a string of amino acid

abbreviations, modifications in the sequence, if any, set forth in the Features section, the sequence comes within the scope of the rules. However, the rules are not intended to encompass the subject matter that is generally referred to as synthetic resins.

III. SEQUENCES DISCLOSED IN APPLICATION TEXT

The requirement for compliance in 37 CFR 1.821(c) is directed to "disclosures of nucleotide and/or amino acid sequences." (Emphasis added.) All sequences, whether claimed or not, that meet the length thresholds in 37 CFR 1.821(a) are subject to the "Sequence Listing" rules. The goal of the Office is to build a comprehensive database that can be used for, *inter alia*, assessing the prior art. It is therefore essential that all sequences, whether only disclosed or also claimed, be included in the database. In those instances in which prior art sequences are only referred to in a given application by name and a publication or accession reference, they need not be included as part of the "Sequence Listing", unless the referred-to sequence is "essential material" per MPEP § 608.01(p). However, if the applicant presents the sequence as a string of particular nucleotide bases or amino acids, whether by way of symbols, words or chemical structure, it is necessary to include the sequence in the "Sequence Listing" regardless of whether the applicant considers the sequence to be prior art, so long as the sequence meets the criteria of 37 CFR 1.821(a). In general, any sequence that is disclosed and/or claimed as a sequence, i.e., as a string of particular nucleotide bases or amino acids, and that otherwise meets the criteria of 37 CFR 1.821(a), must be set forth in the "Sequence Listing".

IV. VARIANTS OF A PRESENTED SEQUENCE

It is generally acceptable to present a single, primary sequence in the specification and "Sequence Listing" by enumeration of its residues in accordance with the sequence rules ("primary sequence") and to discuss and/or claim variants of that primary sequence without presenting each variant as a separate sequence in the "Sequence Listing". Where the variant sequence meets the length thresholds of 37 CFR 1.821(a) and is disclosed by enumeration of its residues anywhere in an application, it must be presented in a "Sequence Listing" in a manner that complies with the requirements of the sequence rules. However, the primary sequence should be annotated in the "Sequence Listing" to reflect such variants. By way of example only, the following types of sequence disclosures would be treated as noted herein by the Office. With respect to a primary sequence and "conservatively modified variants thereof," the sequences may be described as SEQ ID NO:X (the primary sequence) and "conservatively modified variants thereof," if desired. With respect to a sequence that "may be deleted at the C-terminus by 1, 2, 3, 4, or 5 residues," all of the implied variations do not need to be included in the "Sequence Listing". In this latter example, only the sequence without deletions needs to be included in the "Sequence Listing", though applicant is encouraged to annotate the sequence to indicate that deletions have been made at the C-terminus by 1, 2, 3, 4, or 5 residues.

The Office's database will only contain the unmodified sequence. It is strongly recommended that any sequences appearing in the claims, or sequences that are considered essential to understanding the invention, be included in the "Sequence Listing" as a separate sequence.

V. SEQUENCE IDENTIFIER

37 CFR 1.821(d) and 37 CFR 1.823(a)(5) require that each disclosed nucleic acid and/or amino acid sequence in the application appear separately in the "Sequence Listing", with each sequence further being assigned a sequence identifier, referred to as "SEQ ID NO." or the like. 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 for an amino acid or nucleotide sequence in the specification or claims where it is clear that a sequence from the "Sequence Listing" is shown in the description or claims. The sequence identifiers must begin with 1 and increase sequentially by integers. The requirement for sequence identifiers, at a minimum, requires that each sequence be assigned a different number for purposes of identification. However, where practical and for ease of reference, sequences should be presented in the

"Sequence Listing" in numerical order and in the order in which they are discussed in the application.

37 CFR 1.821(d) further requires that where the description or claims of a patent application discuss a sequence that is set forth in the "Sequence Listing", a reference to the sequence identifier of that sequence is required at all occurrences, even if in the text of the description or claims where the sequence is set forth by enumeration of its residues. This requirement is also intended to permit references elsewhere in the application (e.g., specification, claims, or drawings) to sequences set forth in the "Sequence Listing" by the use of assigned sequence identifiers without repeating the sequence. Sequence identifiers can also be used to discuss and/or claim parts or fragments of a properly presented sequence. For example, language such as "residues 14 to 243 of SEQ ID NO:23" is permissible and the fragment need not be separately presented in the "Sequence Listing". Where a sequence that meets the length thresholds of 37 CFR 1.821(a) is disclosed by enumeration of its residues anywhere in an application, it must be presented in a "Sequence Listing" in a manner that complies with the requirements of the sequence rules.

The rules do not alter, in any way, the requirements of 35 U.S.C. 112. The implementation of the rules has had no effect on disclosure and/or claiming requirements. The rules, in general, or the use of sequence identifiers throughout the specification and claims, specifically, should not raise any issues under 35 U.S.C. 112(a) or 35 U.S.C. 112(b). The use of sequence identifiers (SEQ ID NO:X or the like) only provides a shorthand way for applicants to discuss and claim their inventions. These identifiers do not in any way restrict the manner in which an invention can be claimed.

2422.02 The Requirement for Exclusive Conformance; Sequences Presented in Drawing Figures [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

For all applications that disclose nucleic acid and/or amino acid sequences that fall within the definition set forth in 37 CFR 1.821(a), 37 CFR 1.821(b)requires exclusive conformance to the requirements of 37 CFR 1.821 through 37 CFR 1.825 with regard to the manner in which the disclosed nucleotide and/or amino acid sequences are presented and described. This requirement is necessary to minimize any confusion that could result if more than one format for representing sequence data was employed in a given application.

Pursuant to 37 CFR 1.83(a), sequences that are included in sequence listings should not be duplicated in the drawings. However many significant sequence characteristics may only be demonstrated by a figure. This is especially true in view of the fact that the representation of double stranded nucleotides is not permitted in the "Sequence Listing" and many significant nucleotide features, such as "sticky ends" and the like, may only be shown effectively by reference to a drawing figure. Further, the similarity or homology between/among sequences may only be depicted in an effective manner in a drawing figure. Similarly, drawing figures are recommended for use with amino acid sequences to depict structural features of the corresponding protein, such as epitopes and interaction domains. The situations discussed herein are given by way of example only and there may be many other reasons for including a sequence in a drawing. However, when a sequence is presented in a drawing, the sequence must still be included in the "Sequence Listing" if the sequence falls within the definition set forth in 37 CFR 1.821(a), and a sequence identifier ("SEQ ID NO:X" or the like) must be used, either in the drawing itself or in the Brief Description of the Drawings.

2422.03 Sequence Listing Submission [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26

requirements for applications filed on or after July 1, 2022.]

37 CFR 1.821(c) requires that applications containing disclosures of nucleotide and/or amino acid sequences that fall within the definitions of 37 CFR 1.821(a) contain, as a separate part of the specification, a disclosure of each of the nucleotide and/or amino acid sequences, and associated information, using the format and symbols that are set forth in 37 CFR 1.822 and 37 CFR 1.823. This separate part of the specification is referred to as the "Sequence Listing". Except for submission of a national phase application under 35 U.S.C. 371 of an international application (PCT) that is compliant with PCT Rule 5.2(a), the "Sequence Listing" required pursuant to 37 CFR 1.821(c) must be submitted (1) as an ASCII plain text file via the USPTO patent electronic filing system or on a read-only optical disc under 37 CFR 1.52(e), accompanied by an incorporation by reference statement of the ASCII plain text file, in a separate paragraph of the specification, in accordance with 37 CFR 1.77(b)(5) (see 37 CFR 1.821(c)(1)); (2) as a PDF image file submitted via the USPTO patent electronic filing system (see 37 CFR 1.821(c)(2)); or (3) on physical sheets of paper (see 37 CFR 1.821(c)(3)). If the "Sequence Listing" required by 37 CFR 1.821(c) is submitted on physical sheets of paper or as a PDF image file, the "Sequence Listing" is a separate part of the specification which must begin on a new page within the specification. Also, in an application filed under 35 U.S.C. 111(a) in which sequence information is submitted in an ASCII plain text file in compliance with 37 CFR 1.824 and as a PDF image file or on physical sheets of paper, the PDF image file or the physical sheets of paper will comply with the listing requirement under 37 CFR 1.821(c) and the ASCII plain text file will comply with the CRF requirement under 37 CFR 1.821(e)(1)(i).

When the "Sequence Listing" is submitted as a PDF image file under 37 CFR 1.821(c)(2) via USPTO patent electronic filing system or on physical sheets of paper under 37 CFR 1.821(c)(3), 37 CFR 1.821(e)(1) requires that the copy of the 1.821(c) "Sequence Listing" must also be submitted in a separate computer readable form (CRF) in accordance with the requirements of 37 CFR 1.824. Similarly, in the case of a national stage application submitted under 35 U.S.C. 371, when the "Sequence Listing" is submitted as a PDF image file via the USPTO patent electronic filing system under 37 CFR 1.821(c)(2) or on physical sheets of paper under 37 CFR 1.821(c)(3), 37 CFR 1.821(e)(2) requires that the copy of the "Sequence Listing" referred to in 37 CFR 1.821(c) must also be submitted in a separate CRF in accordance with the requirements of 37 CFR 1.824. 37 CFR 1.821(e)(3) requires that, in text format in an international application under the PCT that is to be searched by the United States International Searching Authority or examined by the United States International Preliminary Examining Authority, a CRF in accordance with the requirements of 37 CFR 1.824 must be submitted if a "Sequence Listing" in ASCII plain text format in compliance with 37 CFR 1.824 has not been submitted and the application contains disclosures of nucleotide and/or amino acid sequences, as defined in 37 CFR 1.821(a).

At entry into the national stage under 35 U.S.C. 371, an international application compliant with PCT Rule 5.2 that contains a "Sequence Listing" in ASCII plain text format as part of the description satisfies the requirements of 37 CFR 1.821(c) and 37 CFR 1.821(e). If the international application was previously communicated by the International Bureau under PCT Article 20 and/or was originally filed in the United States Patent and Trademark Office, then no further submission of a "Sequence Listing" or incorporation by reference into the specification is required. Alternatively, if applicant must provide a copy of the international application as required by 37 CFR 1.495(b)(1), the copy of the international application must include the "Sequence Listing" part of the application, and no incorporation by reference into the specification is required. See also 37 CFR 1.823(b)(2) and 1.825(c).

Whether submitted via the USPTO patent electronic filing system or on read-only optical disc(s), the ASCII plain text file must contain a copy of a single "Sequence Listing" in a single file. One hundred (100) megabytes is the size limit for "Sequence Listing" and CRF text files submitted via the USPTO patent electronic filing system, and "Sequence Listing" and CRF text files cannot be compressed when submitted via the USPTO patent electronic

filing system. If a user wishes to submit an electronic copy of a "Sequence Listing" or CRF text file that exceeds 100 megabytes, the "Sequence Listing" or CRF must be filed on read-only optical disc(s).

I. ASCII PLAIN TEXT FILE SUBMITTED VIA USPTO PATENT ELECTRONIC FILING SYSTEM

The Office strongly suggests filing the "Sequence Listing" required by 37 CFR 1.821(c) as an ASCII plain text file via the USPTO patent electronic filing system. See 37 CFR 1.821(c)(1). If sequence information is submitted in an application filed under 35 U.S.C. 111(a) as an ASCII plain text file or as a national stage application submitted under 35 U.S.C. 371 as an ASCII plain text file in compliance with 37 CFR 1.824 via the USPTO patent electronic filing system and applicant has not filed a "Sequence Listing" as a PDF image file or on physical sheets of paper, the ASCII plain text file will serve as both the "Sequence Listing" under 37 CFR 1.821(c) and the CRF of the "Sequence Listing" under 37 CFR 1.821(e). See 37 CFR 1.821(e)(1). See 37 CFR 1.821(e)(2). Note that for applications filed under 35 U.S.C. 111(a), the specification must contain a statement in a separate paragraph (see 37 CFR 1.77(b)(5)) that incorporates by reference the material in the "Sequence Listing" ASCII plain text file identifying the name of the ASCII plain text file, the date of creation, and the size of the ASCII plain text file in bytes (see 37 CFR 1.823(b)(1)). However, an incorporation by reference statement is not required in an international application and is not required in an application file under 35 U.S.C. 371 where the "Sequence Listing" has been previously communicated to the International Bureau or originally filed in the USPTO and complies with Patent Cooperation Treaty Rule 5.2. See 37 CFR 1.821(c), 1.823(b)(2), and 1.825(c). See also MPEP § 2422.03(a) for additional information pertaining to USPTO patent electronic filing system submissions of a "Sequence Listing".

II. ASCII PLAIN TEXT FILE ON READ-ONLY OPTICAL DISC

If the "Sequence Listing" as required by 37 CFR 1.821(c) is submitted on read-only optical disc(s) in accordance with 37 CFR 1.52(e), the specification must contain an incorporation by reference of the material on the read-only optical disc in a separate paragraph (see 37 CFR 1.77(b)(5)) identifying the name of the file, the date of creation, and the size of the file in bytes (37 CFR 1.823(b)(1)). However, an incorporation by reference statement is not required in an international application and is not required in an application file under 35 U.S.C. 371 where the "Sequence Listing" has been previously communicated to the International Bureau or originally filed in the USPTO and complies with Patent Cooperation Treaty Rule 5.2. See 37 CFR 1.821(c), 1.823(b)(2), and 1.825(c). It is noted that a "Sequence Listing" may be compressed using WinZip®, 7-Zip, or Unix®/Linux® Zip (37 CFR 1.824(b)(2)(ii)). If a compressed ASCII plain text file does not fit on a single read-only optical disc due to storage limitations of the read-only optical disc, the compressed ASCII plain text file may be split into multiple file parts and placed on multiple read-only optical discs which are labeled in compliance with 37 CFR 1.52(e)(5)(vi) and (37 CFR 1.824(b)(2)(iv)).

The read-only optical disc used to submit the "Sequence Listing" may also contain "Large Tables" if the table has more than 50 pages of text. See 1.52(e)(1)(iii) and 37 CFR 1.58(c), (f) and (i). The read-only optical disc and duplicate copy must be labeled "Copy 1" and "Copy 2," respectively, and a statement stating that the copies are identical must be included. If the two read-only optical discs are not identical, the Office will use the disc labeled "Copy 1" for further processing (37 CFR 1.58(i)). See also MPEP § 608.05.

If the "Sequence Listing" is submitted in an application filed under 35 U.S.C. 111(a) as an ASCII plain text file in compliance with 37 CFR 1.824 on read-only optical disc(s) in accordance with 37 CFR 1.52(e) and applicant has not filed a "Sequence Listing" as a PDF image file or on physical sheets of paper, the ASCII plain text file will serve as both the "Sequence Listing" under 37 CFR 1.821(c) and the CRF of the "Sequence Listing" under 37 CFR 1.821(e). See 37 CFR 1.821(e)(1). Similarly, if the "Sequence Listing" is filed in a national stage application submitted under 35 U.S.C. 371 as an ASCII plain text file in compliance with 37 CFR 1.824 on read-only optical disc(s) in accordance with 37 CFR 1.52(e) and applicant has not filed a

"Sequence Listing" as a PDF image file or on physical sheets of paper, the ASCII plain text file will serve as both the "Sequence Listing" under 37 CFR 1.821(c) and the CRF of the "Sequence Listing" under 37 CFR 1.821(e).

A. ASCII Plain Text Files Up to 300 MB

When the "Sequence Listing" or CRF is submitted via read-only optical disc(s), the text file may either be compressed or not compressed. If the ASCII plain text file is not compressed, the ASCII plain text file must be contained on a single read-only optical disc. However, if the file does not fit on a single read-only optical disc even after compression, a compressed ASCII plain text file may be split into multiple file parts, in accordance with the target read-only optical disc size, and labeled in compliance with 37 CFR 1.52(e)(5)(vi). See 37 CFR 1.824(b).

B. ASCII Plain Text Files 300 MB or Over

Any "Sequence Listing" ASCII plain text file of 300 MB or more is subject to a fee under 37 CFR 1.21(o) to manage handling of the oversized submission (37 CFR 1.52(f)(3)). Pricing for this fee is divided into two tiers with Tier 1 for file sizes 300 MB to 800 MB and Tier 2 for file sizes greater than 800 MB. The level of effort associated with the handling of mega-"Sequence Listing" is significant, because the Office's systems require extra storage and special handling for files beyond 300 MB. The fee should encourage applicants to draft their specifications such that sequence data that is not essential material is not required to be included in a "Sequence Listing". A reduced number of mega-"Sequence Listings" will benefit the Office and the public by reducing the strain on Office resources, thus facilitating the effective administration of the patent system.

The fee under 37 CFR 1.21(o) is due upon the first submission of a "Sequence Listing" that exceeds 800 MB, or the first submission of a "Sequence Listing" of at least 300MB, whichever applicable fee is higher. As an example, if an application was filed prior to January 16, 2018 (with or without a text file "Sequence Listing"), and thereafter a mega-"Sequence Listing" that is between 300 and 800 MB is filed, the fee under 37 CFR 1.21(o)(1) is due. If an applicant thereafter files a corrected "Sequence Listing" that is also between 300 and 800 MB, no additional fee is due. If a further corrected "Sequence Listing" is filed and the file size exceeds 800 MB, then the total fee owed under 37 CFR 1.21(o) is the fee set forth in 37 CFR 1.21(o)(2). The fee, which is difference between the current fee and the prior paid fee, is due upon submission of the mega-"Sequence Listing". Subsequent deletion or reduction in size of a "Sequence Listing" does not change the requirement to pay the mega-"Sequence Listing" submission fee.

The fee under 37 CFR 1.21(o) does not apply to international applications in the international stage, but does apply to the submission of mega-"Sequence Listings" received in national stage applications under 35 U.S.C. 371, including mega-"Sequence Listings" received by the Office pursuant to PCT Article 20. See MPEP § 2422.03(a), subsection IV, for additional information.

2422.03(a) "Sequence Listing" Submitted as ASCII Plain Text Files [R-01.2024]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

The Legal Framework for Patent Electronic System (**www.uspto.gov/PatentLegalFramework**) and MPEP § 502.05 provide detailed information pertaining to filing applications and other documents via the USPTO patent electronic filing system. The information below is specific to "Sequence Listing" submissions via the USPTO patent electronic filing system.

Pursuant to the Legal Framework for Patent Electronic System, applicants may submit a "Sequence Listing" under 37 CFR 1.821 as an ASCII plain text file via the USPTO patent electronic filing system or on read-only optical disc(s), provided the specification contains a statement in a separate paragraph (preferably on the first page) that incorporates by reference the material in the ASCII

plain text file identifying the name of the ASCII plain text file, the date of creation, and the size of the ASCII plain text file in bytes. See 37 CFR 1.77(b)(5) and 1.823(b)(1). An exception is that an incorporation by reference statement is not required in an international application and is not required in an application file under 35 U.S.C. 371 where the "Sequence Listing" has been previously communicated to the International Bureau or originally filed in the USPTO and complies with Patent Cooperation Treaty Rule 5.2. See 37 CFR 1.821(c), 1.823(b)(2), and 1.825(c). The requirements of 37 CFR 1.52(e) for documents submitted on read-only optical disc(s) are not applicable to a "Sequence Listing" submitted as ASCII plain text files via the USPTO patent electronic filing system. However, each text file must be an ASCII plain text file and have a file name with a ".txt" extension. See 37 CFR 1.824.

I. ASCII PLAIN TEXT FILES SERVE AS BOTH THE "SEQUENCE LISTING" AND THE CRF

It is recommended that a "Sequence Listing" be submitted as an ASCII plain text file via the USPTO patent electronic filing system rather than as a PDF image file. See subsection IV, below, for information regarding filing an international application (PCT) with a sequence listing ASCII plain text file via the USPTO patent electronic filing system.

If the "Sequence Listing" is submitted in an application filed under 35 U.S.C. 111(a) as an ASCII plain text file in compliance with 37 CFR 1.824 via the USPTO patent electronic filing system and applicant has not filed a "Sequence Listing" as a PDF image file or on physical sheets of paper, the ASCII plain text file will serve as both the "Sequence Listing" under 37 CFR 1.821(c) and the CRF of the "Sequence Listing" under 37 CFR 1.821(e). See 37 CFR 1.821(e)(1). Similarly, if the "Sequence Listing" is filed in a national stage application submitted under 35 U.S.C. 371 as an ASCII plain text file in compliance with 37 CFR 1.824 via the USPTO patent electronic filing system, the ASCII plain text file will serve as both the "Sequence Listing" under 37 CFR 1.821(c) and the CRF of the "Sequence Listing" under 37 CFR 1.821(e). Thus, the following are not required and should not be submitted: (1) a second copy of the "Sequence Listing" as a PDF image file or on physical sheets of paper; and (2) a statement under 37 CFR 1.821(e)(1)(ii) or (2)(ii) (indicating that the sequence information contained in the "Sequence Listing" under 37 CFR 1.821(c) and CRF copy of the "Sequence Listing" under 37 CFR 1.821(e)(1)(ii) or (2)(ii) are identical). Also, the practice of CRF transfers has been eliminated. Checker software that may be used to check a "Sequence Listing" for compliance with the requirements of 37 CFR 1.824 is available on the USPTO website at **www.uspto.gov/Checker4**. The User Notes on the Checker website should be consulted for an explanation of the scope of errors and content that is able to be verified by the Checker software.

If a user adds a "Sequence Listing" (under PCT Rule 13ter) as an ASCII plain text file via the USPTO patent electronic filing system in response to a requirement under 37 CFR 1.821(h), the "Sequence Listing" ASCII plain text file must be accompanied by a statement that the ASCII plain text file does not go beyond the disclosure in the international application as filed and the late furnishing fee as set forth in 37 CFR 1.445(a)(5). However, if a user submits an amendment to add or replace a "Sequence Listing" (under 37 CFR 1.821(c)) as an ASCII plain text file via the USPTO patent electronic filing system in response to a requirement under 37 CFR 1.821(g), the submission must comply with 37 CFR 1.825. See MPEP § 2426.

In applications filed under 35 U.S.C. 111(a) and 371, submission of the "Sequence Listing" as a PDF image file or on physical sheets of paper is not recommended. Where the "Sequence Listing" is submitted as a PDF image or on physical sheets of paper, applicant must provide the CRF required by 37 CFR 1.821(e)(1)(i). Note that the "Sequence Listing" in the PDF image file or on physical sheets of paper will not be excluded when determining the application size fee. The USPTO prefers the submission of sequence information in an ASCII plain text file via the USPTO patent electronic filing system because as stated above, if in an application filed under 35 U.S.C. 111(a) or 35 U.S.C. 371 applicant has not filed a second copy of the "Sequence Listing" as a PDF image file or on physical sheets of paper (see 37 CFR 1.821(e)(1)), the ASCII plain text file will serve as both the

"Sequence Listing" required by 37 CFR 1.821(c) and the CRF required by 37 CFR 1.821(e).

II. APPLICATION SIZE FEE

Any "Sequence Listing" or CRF of a "Sequence Listing" submitted as an ASCII plain text file via the USPTO patent electronic filing system that is in compliance with 37 CFR 1.821(c) or (e) will be excluded when determining the application size fee required by 37 CFR 1.16(s) or 1.492(j) as per 37 CFR 1.52(f)(2). A "Sequence Listing" submitted as a PDF image file via the USPTO patent electronic filing system or on read-only optical disc will not be excluded when determining the application size fee.

See subsection IV, below, for additional information regarding application size fees in an international application (PCT).

III. SIZE RESTRICTIONS FOR ASCII PLAIN TEXT FILES

One hundred (100) megabytes is the size limit for "Sequence Listing" ASCII plain text files submitted via the USPTO patent electronic filing system, and compression is not allowed for a "Sequence Listing" submitted via the USPTO patent electronic filing system. If a user wishes to submit a "Sequence Listing" ASCII plain text file that exceeds 100 megabytes, it is recommended that the user file the application without the "Sequence Listing" using the USPTO patent electronic filing system to obtain the application number and confirmation number, and then file the "Sequence Listing" on read-only optical disc(s) in accordance with 37 CFR 1.52(e) and 1.824 on the same day by using Priority Mail Express® from the USPS in accordance with 37 CFR 1.10, or hand delivery, in order to secure the same filing date for all parts of the application. Note: a submission of a "Sequence Listing" in electronic form of 300 MB or more in size is subject to an oversized submission fee set forth in 37 CFR 1.21(o). See 37 CFR 1.52(f)(3). Alternatively, a user may submit the application on physical sheets of paper and include the "Sequence Listing" ASCII plain text file on read-only optical disc(s) in accordance with 37 CFR 1.52(e). "Sequence Listing" ASCII plain text files may not be divided into multiple files so as to not exceed the 100 MB size limit for filing via the USPTO patent electronic filing system, and any "Sequence Listing" greater than 100 MB must be submitted on read-only optical disc(s). If the "Sequence Listing" is filed on a read-only optical disc, the "Sequence Listing" must be a single file, and any not compressed file must be contained on a single read-only optical disc. However, a compressed 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 must have a label permanently affixed thereto on which the following information has been hand-printed or typed: (i) First-named inventor (if known); (ii) Title of the invention; (iii) Attorney docket or file reference number (if applicable); (iv) Application number and filing date (if known); (v) Date on which the data were recorded on the read-only optical disc; and (vi) Disc order (e.g., "1 of X"), if multiple read-only optical discs are submitted. See 37 CFR 1.52(e)(5) and 1.824(b).

See subsection IV.B, below, for information regarding submission of a sequence listing text file that exceeds 100 megabytes in an international application (PCT) filed via the USPTO patent electronic filing system.

The current size limit on the USPTO patent electronic filing system for ASCII plain text file submissions of "Large Tables" and a "Computer Program Listing Appendix" is 25 megabytes per submission, and compression is not allowed for "Large Tables" and a "Computer Program Listing Appendix" submitted as ASCII plain text files via the USPTO patent electronic filing system. Files above the 25 MB limit for "Large Tables" and a "Computer Program Listing Appendix" may be either (1) broken up into multiple files that are no larger than 25 MB each and those smaller files may be submitted via the USPTO patent electronic filing system or (2) submitted on read-only optical disc(s) (see 37 CFR 1.52(e)). If the user chooses to break up a "Large Table" or "Computer Program Listing Appendix" file so that it may be submitted via the USPTO patent electronic filing system, the file names must indicate their order (e.g., "1 of X", "2 of X"). If a user wishes to file an application with a "Large Table" or "Computer Program Listing Appendix" ASCII plain text file that is larger than

25 megabytes, it is recommended that the user file the application without the "Large Table" or "Computer Program Listing Appendix" using the USPTO patent electronic filing system to obtain the application number and confirmation number, and then file the "Large Table" or "Computer Program Listing Appendix" on read-only optical disc(s) in accordance with 37 CFR 1.52(e) and 1.58(c) or 1.96(c) on the same day by using Priority Mail Express® from the USPS in accordance with 37 CFR 1.10, or hand delivery, in order to secure the same filing date for all parts of the application. Alternatively, a user may submit the application on physical sheets of paper and include the "Large Table" or "Computer Program Listing Appendix" ASCII plain text file on read-only optical disc(s) in accordance with 37 CFR 1.52(e) in order to secure the same filing date for all parts of the application. See 37 CFR 1.58(f) and 37 CFR 1.96(c)(4).

IV. FILING A SEQUENCE LISTING IN INTERNATIONAL APPLICATIONS (PCT) VIA THE USPTO PATENT ELECTRONIC FILING SYSTEM

A. Sequence Listing Must Be Presented as a Separate Part of the Description

Under PCT Rule 5.2(a), "where the international application contains disclosure of one or more nucleotide and/or amino acid sequences, the description shall contain a sequence listing complying with the standard provided for in the Administrative Instructions and presented as a separate part of the description in accordance with that standard". The standard is set forth in the PCT Administrative Instructions Annex C, entitled Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings in International Patent Applications Under the PCT. When filing an international application (PCT) using the USPTO patent electronic filing system, it is highly recommended to submit a sequence listing as a single ASCII plain text file with a ".txt" extension on the international filing date. Note that 100 megabytes is the size limit for submitting a sequence listing ASCII plain text file via the USPTO patent electronic filing system. See subsection IV.B, below. Although it is not recommended, applicant can also submit the sequence listing in a PCT application as a PDF image file as part of the application on the international filing date.

If the sequence listing is submitted as an ASCII plain text file when filing a new international application (PCT), applicant need not and should not submit any additional copies, including PDF image files. The single ASCII plain text file is preferred because the ASCII plain text file will serve both as the sequence listing part of the description required under PCT Rule 5.2 and the electronic form under PCT Rule 13 *ter*.1(a) in the absence of a PDF sequence listing file. The check list (Box No. IX) of the PCT Request (form PCT/RO/101) provided via the USPTO patent electronic filing system together with the international application (PCT) must indicate that the sequence listing submitted as an ASCII plain text file forms part of the international application. Furthermore, the statement as set forth in paragraph 4(v) of the AI Annex C (Administrative Instructions under the PCT, Annex C), that "the information recorded in electronic form furnished under PCT Rule 13 *ter* is identical to the sequence listing as contained in the international application," is not necessary.

Submission of the sequence listing part of the description as a PDF image file in a new international application (PCT) is not recommended because, where the application does not contain a sequence listing as an ASCII plain text file, the International Searching Authority or the International Preliminary Authority may invite applicant to furnish a copy of the sequence listing in ASCII plain text format for the purposes of international search and/or preliminary examination. Any sequence listing submitted in response to this invitation will not form part of the application.

When a sequence listing is filed via the USPTO patent electronic filing system in a new PCT international application as both a PDF image file and an ASCII plain text file, but the Request form Box No. IX does not indicate which one forms part of the international application, the PDF image copy of the sequence listing will be considered to form part of the application and the ASCII plain text file will be considered to be an accompanying item for search purposes under PCT Rule 13ter.1(a) only.

The international filing fee for an international application (PCT) that includes a sequence listing, filed via the USPTO patent electronic filing system, is calculated based on the type of sequence listing file that is part of the description of the international application (PCT). A sequence listing filed as an ASCII plain text file will not be included in the sheet count of the international application (PCT). A sequence listing filed as a PDF image file will be included in the sheet count of the international application (PCT), when it is part of the description. When a new PCT international application is filed via the USPTO patent electronic filing system and contains the sequence listing part of the description as a PDF image file and a sequence listing ASCII plain text to be used only for search purposes under PCT Rule 13ter.1(a), the sheets of the PDF image file will count towards excess sheet fees, if any.

B. File Size and Quantity Limits

One hundred (100) megabytes is the size limit for sequence listing ASCII plain text files submitted via the USPTO patent electronic filing system. See 37 CFR 1.824(b)(1). Sequence listing ASCII plain text files must not be partitioned into multiple files for filing via the USPTO patent electronic filing system. The sequence listing must be in a single ASCII plain text file.

Applicant may use the USPTO patent electronic filing system to file part of the international application (PCT) and to obtain the international application (PCT) number and the confirmation number, and then file the remainder of the international application (PCT) on the same day as one or more follow-on submissions using the USPTO patent electronic filing system, in order to secure the same filing date for all parts of the international application (PCT). However, applicant is not permitted to file part of the international application (PCT) electronically via the USPTO patent electronic filing system, and then file the remainder of the international application (PCT) on paper to secure a filing date of all parts of the international application (PCT).

In the situation where applicant needs to file a sequence listing that is over one hundred (100) megabytes, applicant may use the USPTO patent electronic filing system to file the international application (PCT) without the sequence listing to obtain the international application (PCT) number and the confirmation number, and then file the sequence listing on read-only optical discs on the same day by using Priority Mail Express® from the USPS in accordance with 37 CFR 1.10, or by hand delivery, in order to secure the same filing date for all parts of the international application (PCT). However, the read-only optical discs must not contain PDF image files and must fully comply with the guidelines for filing a sequence listing on electronic media. The check list of the PCT Request provided via the USPTO patent electronic filing system together with the international application (PCT) must indicate that the sequence listing will be filed separately on physical data carrier(s), on the same day and in the form of an Annex C/ST.25 text file. The sequence listing must be a single file, but the file may be split for submission on multiple physical media using software designed to divide a file into multiple files for subsequent concatenation. If the user breaks up a sequence listing for submission on multiple read-only optical discs, the read-only optical discs must be labeled to indicate their order (e.g., “1 of X”, “2 of X”).

Submissions of very lengthy sequence listings (300 MB or over a.k.a. mega-sequence listings) in international applications are not subject to the mega-sequence listing submission fees set forth in 37 CFR 1.21(o). However, for mega-sequence listing submissions on or after January 16, 2018, the fee under 37 CFR 1.21(o) does apply to the submission of mega-sequence listings received in national stage applications under 35 U.S.C. 371, including mega-sequence listings received by the Office pursuant to PCT Article 20. Similarly, if an international application is filed at RO/US with a mega-sequence listing, and thereafter a bypass continuing application is filed under 35 U.S.C. 111(a), the fee under 37 CFR 1.21(o) will be due in the continuing application for mega-sequence listing submissions on or after January 16, 2018. See 37 CFR 1.52(f)(3).

2422.04 The Requirement for a Computer Readable Copy of the "Sequence Listing" [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.821(e) requires the submission of a copy of the "Sequence Listing" in computer readable form (CRF) in an application filed under 35 U.S.C. 111(a) or in a national stage application submitted under 35 U.S.C. 371. A separate computer readable form must be submitted via the USPTO patent electronic filing system or on read-only optical disc(s), as permitted by 37 CFR 1.824(b), when the "Sequence Listing" required by 37 CFR 1.821(c) is submitted as a PDF image file or on physical sheets of paper in a U.S. application filed under 35 U.S.C. 111(a) (see 37 CFR 1.821(e)(1)) or when the "Sequence Listing" required by 37 CFR 1.821(c) is submitted as a PDF image file or on physical sheets of paper and not also submitted as an ASCII plain text file in a national stage application (see 37 CFR 1.821(e)(2)). However, the Office prefers submission of sequence information as an ASCII plain text file via the USPTO patent electronic filing system or on read-only optical disc(s) without a copy of the "Sequence Listing" as a PDF image file or on physical sheets of paper in all applications because such an ASCII submission will serve as both the "Sequence Listing" required by 37 CFR 1.821(c) and the CRF of the "Sequence Listing" required by 37 CFR 1.821(e) and the "Sequence Listing" submitted as an ASCII plain text file will not be included in the application size fee determination under 37 CFR 1.52(f)(1) or (2). See MPEP § 2422.03(a)(I).

The information on the computer readable form will be entered into the Office's database for searching and printing nucleotide and amino acid sequences. This electronic database will also enable the Office to provide published sequence data, in electronic form, to the National Center for Biotechnology Information (NCBI) for publication in GenBank, and enable NCBI to exchange data with the DNA Data Bank of Japan (DDBJ) and the European Bioinformatics Institute (EBI). It should be noted that the Office's database complies with the confidentiality requirement imposed by 35 U.S.C. 122. Unpublished pending application sequences are maintained in the database separately from published or patented sequences. That is, the Office will not exchange or make public any information on any sequence until the patent application containing that information is published or matures into a patent, or as otherwise allowed by 35 U.S.C. 122.

The Office may permit correction of the "Sequence Listing" submitted pursuant to 37 CFR 1.821(c), whether on physical sheets of paper or as a PDF image file, at the least, during the pendency of a given application by reference to the computer readable form thereof submitted pursuant to37 CFR 1.821(e) if both the "Sequence Listing" and computer readable form were submitted at the time of filing of the application and the totality of the circumstances otherwise substantiate the proposed correction. A mere discrepancy between the "Sequence Listing" and the computer readable form may not, in and of itself, be sufficient to justify a proposed correction. In this regard, the Office will assume that the computer readable form has been incorporated by reference into the application when the "Sequence Listing" and computer readable form were submitted at the time of filing of the application. The Office will attempt to accommodate or address all correction issues, but it must be kept in mind that the real burden rests with the applicant to ensure that any discrepancies between the "Sequence Listing" and the CRF copy are eliminated or minimized. Applicants should be aware that there will be instances where the applicant may have to suffer the consequences of any discrepancies between the two. It is noted that in an application filed under 35 U.S.C. 111(a) in which applicant has not filed a second copy of the "Sequence Listing" as a PDF image file or on physical sheets of paper (see 37 CFR 1.821(e)(1)), an ASCII plain text file will serve as both the "Sequence Listing" required by 37 CFR 1.821(c) and the CRF required by 37 CFR 1.821(e), eliminating any chance for discrepancies. Filing the "Sequence Listing" as an ASCII plain text file submitted via the USPTO patent electronic filing system that complies with both 37

CFR 1.821(c) and (e) is the Office's preferred method of receiving a "Sequence Listing".

The Office does not desire to be bound by a requirement to permanently preserve computer readable forms submitted on read-only optical disc(s) for support, priority or correction purposes. Thus, once use of the CRF by the Office for processing has ended, i.e., once the Office has entered the data contained on the computer readable form into the appropriate database, the Office does not intend to further preserve the CRF submitted by the applicant, and applicant should not expect to have the read-only optical disc(s) returned. See 37 CFR 1.52(e)(6).

2422.05 [Reserved]

2422.06 Requirement for Statement Regarding Information Contained in the "Sequence Listing" and Separate Computer Readable Form [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

When a separate computer readable form (CRF) of a "Sequence Listing" is submitted in an application filed under 35 U.S.C. 111(a) because the "Sequence Listing" is filed as a PDF image file (37 CFR 1.821(c)(2)) or on physical sheets of paper (37 CFR 1.821(c)(3)) (see 37 CFR 1.821(e)(1)(i)) or filed in a national stage application submitted under 35 U.S.C. 371 because the "Sequence Listing" is filed as a PDF image file (37 CFR 1.821(c)(2)) or on physical sheets of paper (37 CFR 1.821(c)(3)) and not also as an ASCII plain text file in compliance with 37 CFR 1.824 (see 37 CFR 1.821(e)(2)(i)), 37 CFR 1.821(e)(1)(ii) and (2)(ii) require, a statement that the information contained in the "Sequence Listing" and the separate CRF are identical. When a CRF is submitted in the international stage of an international application under the PCT in response to an notice requesting a ASCII plain text formatted sequence listing by the United States International Searching Authority or by the United States International Preliminary Examining Authority because a sequence listing in ASCII plain text format in compliance with 37 CFR 1.824 has not been filed (see 37 CFR 1.821(e)(3)(i)), 37 CFR 1.821(e)(3)(iii) requires a statement that the information contained in the CRF does not go beyond the disclosure in the international application as filed or a statement that the information recorded in the ASCII plain text file of the CRF is identical to the sequence listing contained in the international application as filed, as applicable. Such a statement may be made by a registered practitioner, the applicant, an inventor, or the person who actually compares the sequence data on behalf of the aforementioned. See MPEP § 2428 for further information and Sample Statements.

Note that, in an application filed under 35 U.S.C. 111(a), if a "Sequence Listing" is filed as an ASCII plain text file via the USPTO patent electronic filing system or on a read-only optical disc under 37 CFR 1.52(e), and applicant has not filed a "Sequence Listing" as a PDF image file or on physical sheets of paper, the ASCII plain text file will serve as both the "Sequence Listing" required by 37 CFR 1.821(c) and the computer readable form (CRF) required by 37 CFR 1.821(e) (37 CFR 1.821(e)(1)). Also, in a national stage application submitted under 35 U.S.C. 371, if a "Sequence Listing" is filed as an ASCII plain text file via the USPTO patent electronic filing system or on a read-only optical disc under 37 CFR 1.52(e), the ASCII plain text file will serve as both the "Sequence Listing" required by 37 CFR 1.821(c) and the computer readable form (CRF) required by 37 CFR 1.821(e). See MPEP § 2422.03(a), subsection I, for additional information. See also MPEP § 2422.03(a) subsection IV, for additional information regarding international stage applications. Thus, for applications filed under 35 U.S.C. 111(a) and 35 U.S.C. 371, the following are not required and should not be submitted: (1) a second copy of the "Sequence Listing" as a PDF image file or on physical sheets of paper; and (2) a statement under 37 CFR 1.821(e)(1)(ii) or (2)(ii) (indicating that the sequence information contained in the "Sequence Listing" under 37 CFR 1.821(c) and CRF copy of the "Sequence Listing" under 37 CFR 1.821(e)(1)(ii) or (2)(ii) are identical).

2422.07 Requirements for Compliance and Consequences of Non-Compliance [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.821(g) requires compliance with the requirements of 37 CFR 1.821(b) through (e), as discussed above, if they are not satisfied at the time of filing under 35 U.S.C. 111(a) or at the time of entering the national stage of an international application under 35 U.S.C. 371, within the period of time set in a notice requiring compliance. When applicant files an amendment to comply with the requirements of 37 CFR 1.821(g) and that amendment adds or replaces a "Sequence Listing" and CRF copy thereof, the amendment must be submitted in accordance with the requirements of 37 CFR 1.825. Failure to provide a proper reply in compliance with 37 CFR 1.825 will result in the abandonment of the application. See MPEP § 2426. Extensions of time in which to reply to a requirement under this paragraph are available pursuant to 37 CFR 1.136. Note, however, that patent applications filed under 35 U.S.C. 111 on or after December 18, 2013, and international patent applications in which the national stage commenced under 35 U.S.C. 371 on or after December 18, 2013, may be subject to reductions in patent terms adjustment pursuant to 37 CFR 1.704(c)(13) if they are not in condition for examination within eight months from the filing date or date of commencement, respectively. "In condition for examination" includes compliance with 37 CFR 1.821 through 1.825 (see 37 CFR 1.704(f)).

Provisional applications filed under 35 U.S.C. 111(b) need not comply with 37 CFR 1.821 through 1.825, however, applicants are encouraged to file a "Sequence Listing" as defined in 37 CFR 1.821(c) for ease of identification of the sequence information contained in the provisional application.

If any of the requirements of 37 CFR 1.821(e)(3) are not satisfied at the time of filing an international application under the Patent Cooperation Treaty (PCT), and 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. Where a sequence listing under PCT Rule 13ter is provided in reply to a under 37 CFR 1.821(h), the sequence listing must be accompanied by a statement that the information recorded in the ASCII plain text file under 37 CFR 1.821(e)(3)(i) is identical to the sequence listing contained in the international application as filed, or does not go beyond the disclosure in the international application as filed, as applicable. Such a statement may be made by a registered practitioner, the applicant, an inventor, or the person who actually compares the sequence data on behalf of the aforementioned. Also, the ASCII plain text file under 37 CFR 1.821(e)(3)(i) must be accompanied by the late furnishing fee, as set forth in 37 CFR 1.445(a)(5). International applications that fail to comply with any of the requirements of 37 CFR 1.821(e)(3) will be searched and/or examined only to the extent possible without the benefit of the information in computer readable form. See PCT Administrative Instructions Section 513(c).

The requirement to submit a statement that a submission in reply to the requirement under 37 CFR 1.821(h) does not go beyond the disclosure in the application as filed or that the information recorded in the ASCII plain text file under 37 CFR 1.821(e)(3)(i) is identical to the sequence listing contained in the international application as filed is not the first instance in which the applicant has been required to ensure that there is not new matter upon amendment. The requirement is analogous to that found in 37 CFR 1.125 regarding substitute specifications. When a substitute specification is required because the number or nature of amendments would make it difficult to examine the application, the applicant must include a statement that the substitute specification includes no new matter. The necessity of requiring sequence information as an ASCII plain text file is similar to the necessity of requiring a substitute specification and, likewise, the burden is on the applicant to ensure that no new matter is added. Applicants have a duty to comply with the statutory prohibition (35 U.S.C.

132 and 35 U.S.C. 251) against the introduction of new matter.

The correction of errors in sequencing or any other errors that are made in describing an invention are subject to the statutory prohibition (35 U.S.C. 132 and 35 U.S.C. 251) against the introduction of new matter.

2422.08 Presumptions Regarding Compliance [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

Neither the presence nor absence of information which is not required under the sequence rules will create a presumption that such information is necessary to satisfy any of the requirements of 35 U.S.C. 112. Further, the grant of a patent on an application that is subject to 37 CFR 1.821 through 37 CFR 1.825 constitutes a presumption that the granted patent complies with the requirements of these rules.

2422.09 [Reserved]

2423 Symbols and Format To Be Used for Nucleotide and/or Amino Acid Sequence Data for WIPO ST.25 [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.822 Symbols and format to be used for nucleotide and/or amino acid sequence data.

(a) The symbols and format to be used for nucleotide and/or amino acid sequence data shall conform to the requirements of paragraphs (b) through (e) of this section.

(b) The code for representing the nucleotide and/or amino acid sequence characters shall conform to the code set forth in appendices A and C to this subpart. No code other than that specified in these sections shall be used in nucleotide and amino acid sequences. A modified base or modified or unusual amino acid may be presented in a given sequence as the corresponding unmodified base or amino acid if the modified base or modified or unusual amino acid is one of those listed in appendices B and D to this subpart, and the modification is also set forth in the Feature section. Otherwise, each occurrence of a base or amino acid not appearing in appendices A and C, shall be listed in a given sequence as "n" or "Xaa," respectively, with further information, as appropriate, given in the Feature section, by including one or more feature keys listed in appendices E and F to this subpart.

Note 1 to paragraph (b): Appendices A through F to this subpart contain Tables 1– 6 of the World Intellectual Property Organization (WIPO) Handbook on Industrial Property Information and Documentation, Standard ST.25: Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings in Patent Applications (2009).

(c) *Format representation of nucleotides.*

(1) A nucleotide sequence shall be listed using the lowercase letter for representing the one-letter code for the nucleotide bases set forth in appendix A to this subpart.

(2) The bases in a nucleotide sequence (including introns) shall be listed in groups of 10 bases except in the coding parts of the sequence. Leftover bases, fewer than 10 in number, at the end of noncoding parts of a sequence shall be grouped together and separated from adjacent groups of 10 or 3 bases by a space.

(3) The bases in the coding parts of a nucleotide sequence shall be listed as triplets (codons). The amino acids corresponding to the codons in the coding parts of a nucleotide sequence shall be listed immediately below the corresponding codons. Where a codon spans an intron, the amino acid symbol shall be listed below the portion of the codon containing two nucleotides.

(4) A nucleotide sequence shall be listed with a maximum of 16 codons or 60 bases per line, with a space provided between each codon or group of 10 bases.

(5) A nucleotide sequence shall be represented, only by a single strand, in the 5 to 3 direction, from left to right.

(6) The enumeration of nucleotide bases shall start at the first base of the sequence with number 1. The enumeration shall be continuous through the whole sequence in the direction 5 to 3. The enumeration shall appear in the right margin, next to the line containing the one-letter codes for the bases and giving the number of the last base of that line.

(7) For those nucleotide sequences that are circular in configuration, the enumeration method set forth in paragraph (c)(6) of this section remains applicable with the exception that the designation of the first base of the nucleotide sequence may be made at the option of the applicant.

Note 2 to paragraph (c): Appendices A through F to this subpart contain Tables 1– 6 of the World Intellectual Property Organization (WIPO) Handbook on Industrial Property Information and Documentation, Standard ST.25: Standard for

the Presentation of Nucleotide and Amino Acid Sequence Listings in Patent Applications (2009).

(d) *Representation of amino acids.*

(1) The amino acids in a protein or peptide sequence shall be listed using the three-letter abbreviation, with the first letter as an upper case character, as in Appendix C to this subpart.

(2) A protein or peptide sequence shall be listed with a maximum of 16 amino acids per line, with a space provided between each amino acid.

(3) An amino acid sequence shall be represented in the amino to carboxy direction, from left to right, and the amino and carboxy groups shall not be represented in the sequence.

(4) The enumeration of amino acids may start at the first amino acid of the first mature protein, with the number 1. When represented, the amino acids preceding the mature protein, (*e.g.*, pre-sequences, pro-sequences, pre-pro-sequences, and signal sequences) shall have negative numbers, counting backwards starting with the amino acid next to number 1. Otherwise, the enumeration of amino acids shall start at the first amino acid at the amino terminal as number 1, and shall appear below every five amino acids of the sequence. The enumeration method for amino acid sequences that is set forth in this section remains applicable for amino acid sequences that are circular in configuration, with the exception that the designation of the first amino acid of the sequence may be made at the option of the applicant.

(5) An amino acid sequence that contains internal terminator symbols (*e.g.*, "Ter," "*," or "." etc.) may not be represented as a single amino acid sequence but shall be represented as separate amino acid sequences.

Note 3 to paragraph (d): Appendices A through F to this subpart contain Tables 1– 6 of the World Intellectual Property Organization (WIPO) Handbook on Industrial Property Information and Documentation, Standard ST.25: Standard for the Presentation of Nucleotide and Amino Acid Sequence Listings in Patent Applications (2009).

(e) A sequence with a gap or gaps shall be represented as a plurality of separate sequences, with separate sequence identifiers (§ 1.823(a)(5)), with the number of separate sequences being equal in number to the number of continuous strings of sequence data. A sequence composed of one or more noncontiguous segments of a larger sequence or segments from different sequences shall be presented as a separate sequence.

Appendices A through F referenced in 37 CFR 1.822 are reproduced in MPEP § 2422(I).

2423.01 Format and Symbols To Be Used in a "Sequence Listing" [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.822 sets forth the format and symbols to be used for listing nucleotide and/or amino acid sequence data. The symbols for representing the nucleotide and/or amino acid characters in the sequences are set forth in Appendices A and C to Subpart G of Part 1 of the CFR. See MPEP § 2422(I). No other symbols shall be used in nucleotide and amino acid sequences. The "modified base" and "modified and unusual amino acid" symbols appearing in Appendices B and D to Subpart G of Part 1 of the CFR (see 37 CFR 1.822 and MPEP § 2422(I)) are not to be set forth in the sequences recited in the "Sequence Listing". However, "modified base" or "modified and unusual amino acid" symbols may be used in the written description and/or drawing portions of the specification. To properly enter notations for modified bases or amino acids in the "Sequence Listing", the Feature section of the "Sequence Listing" should be used. That is, a modified base or amino acid may be presented in a given sequence as the corresponding unmodified base or amino acid if the modified base or amino acid is one of those listed in Appendices B and D to Subpart G of Part 1 of the CFR and the modification is also set forth in the Feature section of the "Sequence Listing". Otherwise, all nucleotide bases or amino acids not appearing in Appendices A and C to Subpart G of Part 1 of the CFR must be listed in a given sequence as "n" or "Xaa," respectively, with further information given in the Feature section of the "Sequence Listing" by including one or more feature keys listed in Appendices E and F to Subpart G of Part 1 of the CFR. See 37 CFR 1.822(b).

In 37 CFR 1.822(b) and 37 CFR 1.822(d), the use of three-letter symbols for amino acids is required in the "Sequence Listing". The three-letter symbols must be presented using the upper case for the first character and lower case for the remaining two characters. Applicants are encouraged to use the three-letter symbols for amino acids throughout the disclosure, instead of the one-letter symbols, for easier reading of the application and any patent issuing therefrom.

37 CFR 1.822(c) through (e) set forth the format for presenting sequence data. These paragraphs set forth the manner in which the characters in sequences are to be grouped, spaced, presented and numbered.

2423.02 Depiction of Coding Regions [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

If applicant chooses to depict coding regions, 37 CFR 1.822(c)(3) requires the amino acids corresponding to the codons in the coding parts of a nucleotide sequence to be listed immediately below the corresponding codons. Further, in 37 CFR 1.822(c)(3), the situation in which a codon spans an intron has been addressed. In those situations, the "amino acid symbol shall be listed below the portion of the codon containing two nucleotides." This requirement clarifies the representation of an amino acid that corresponds to a codon that spans an intron.

It should be noted that the sequence rules do not, in any way, require the depiction of coding regions or the amino acids corresponding to the codons in those coding regions. 37 CFR 1.822(d) only requires that where amino acids corresponding to the codons in the coding parts of a nucleotide sequence are depicted, they must be depicted below the corresponding codons. There is absolutely no requirement in the rules to depict coding regions. However, when the coding parts of a nucleotide sequence and their corresponding amino acids have been enumerated by their residues, those amino acids must also be set forth as a separate sequence if the amino acid sequence meets the length thresholds in 37 CFR 1.821(a).

2423.03 Presentation and Numbering of Sequences [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.822(c)(5) provides that nucleotide sequences shall only be represented by a single strand, in the 5 to 3 direction, from left to right. That is, double stranded nucleotides shall not be represented in the "Sequence Listing". A double stranded nucleotide may be represented as two single stranded nucleotides, and any relationship between the two may be shown in the drawings.

The procedures for presenting and numbering amino acid sequences are set forth in 37 CFR 1.822(d). Two alternatives are presented for numbering amino acid sequences. Amino acid sequences may be numbered with respect to the identification of the first amino acid of the first mature protein or with respect to the first amino acid appearing at the amino terminal. The numbering procedure for nucleotides is set forth in 37 CFR 1.822(c)(6). Sequences that are circular in configuration are intended to be encompassed by these rules, and the numbering procedures described above remain applicable with the exception that the designation of the first nucleotide base or amino acid of the sequence may be made at the option of the applicant. See 37 CFR 1.822(c)(7) and (d)(4).

In 37 CFR 1.822(e) the procedures for presenting and numbering hybrid and gapped sequences are set forth. A sequence with a gap or gaps shall be presented as a plurality of separate sequences, each having separate sequence identifiers, with the number of separate sequences being equal in number to the number of continuous strings of sequence data. The term "gap" is not intended to embrace a gap or gaps that is/are introduced into the presentation of otherwise continuous sequence information in, e.g., a drawing figure, to show alignments or similarities with other sequences. The "gaps" referred to in this section are gaps representing unknown or undisclosed regions in a sequence between regions that are known or disclosed. On the other hand, a sequence that contains one or more regions of contiguous "n" or "Xaa" residues, wherein the exact number of "n" or "Xaa" residues in each region is

disclosed, must be included in the "Sequence Listing" as a single sequence with a single sequence identifier. A sequence disclosed by enumeration of its residues that is constructed as a single continuous sequence from one or more non-contiguous segments of a larger sequence or segments from different sequences must be included in the "Sequence Listing" as a single sequence with a single sequence identifier. A fragment of a larger sequence need not be enumerated by its residues, and may be referred to in the specification, claims or drawings as, e.g., "residues 2 through 33 of SEQ ID NO:12," assuming that SEQ ID NO:12 has been properly included in the "Sequence Listing".

2424 Requirements for Content of a "Sequence Listing" Part of the Specification under WIPO ST.25. [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.823 Requirements for content of a "Sequence Listing" part of the specification.

(a) The "Sequence Listing" must comply with the following:

(1) The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement in appendix G to this subpart. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional.

(2) Each item of information shall begin on a new line, with the numeric identifier enclosed in angle brackets, as shown in appendix G to this subpart.

(3) Set forth numeric identifiers <110> through <170> at the beginning of the "Sequence Listing."

(4) Include each disclosed nucleotide and/or amino acid sequence, as defined in § 1.821(a).

(5) Assign a separate sequence identifier to each sequence, beginning with 1 and increasing sequentially by integers, and include the sequence identifier in numeric identifier <210>.

(6) Use the code "000" in place of the sequence where no sequence is present for a sequence identifier.

(7) Include the total number of SEQ ID NOs in numeric identifier <160> , as defined in appendix G to this subpart, whether followed by a sequence or by the code "000."

(8) Must not contain more than 74 characters per line.

(b)(1) Unless paragraph (b)(2) of this section applies, if the "Sequence Listing" required by § 1.821(c) is submitted as an ASCII plain text file 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 ASCII plain text file identifying:

(i) The name of the file;

(ii) The date of creation; and

(iii) The size of the file in bytes.

(2) If the "Sequence Listing" required by § 1.821(c) is submitted as an ASCII plain text file 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 incorporation by reference of the material in the ASCII plain text file is not required.

(3) A "Sequence Listing" required by § 1.821(c) that is submitted as a PDF file (§ 1.821(c)(2)) via the USPTO patent electronic filing system or on physical sheets of paper (§ 1.821(c)(3)), setting forth the nucleotide and/or amino acid sequence and associated information in accordance with paragraph (a) of this section:

(i) Must begin on a new page;

(ii) Must be titled "Sequence Listing";

(iii) Must not include material other than the "Sequence Listing" itself;

(iv) Must have sheets containing no more than 66 lines, with each line containing no more than 74 characters;

(v) Should have sheets numbered independently of the numbering of the remainder of the application; and

(vi) Should use a fixed-width font exclusively throughout.

Appendix G to Subpart G of Part 1 - Numeric Identifiers

Numeric Identifier	Definition	Comments and format	Mandatory (M) or optional (O)
<110>	Applicant	If Applicant is inventor, then preferably max. of 10 names; one	M.

Numeric Identifier	Definition	Comments and format	Mandatory (M) or optional (O)
		name per line; preferable format: Surname, Other Names and/or Initials.	
<120>	Title of Invention		M.
<130>	File Reference	Personal file reference	M when filed prior to assignment or appl. number.
<140>	Current Application Number	Specify as: US 09/999,999 or PCT/US09/99999	M, if available.
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available.
<150>	Prior Application Number	Specify as: US 09/999,999 or PCT/US09/99999	M, if applicable include priority documents under 35 U.S.C. 119 and 120.
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable.
<160>	Number of SEQ ID NOs	Count includes total number of SEQ ID NOs	M.
<170>	Software	Name of software used to create the "Sequence Listing"	O.
<210>	SEQ ID NO:#:	Response shall be an integer representing the SEQ ID NO shown	M.
<211>	Length	Respond with an integer expressing the number of bases or amino acid residues	M.
<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/ RNA molecule shall be further described in the <220> to <223> feature section	M.
<213>	Organism	Scientific name, *i.e.*, Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section	M.
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence	M, under the following conditions: If "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/ RNA.
<221>	Name/Key	Provide appropriate identifier for feature, from WIPO Standard ST.25 (2009), Appendices E and F to this subpart	M, under the following conditions: If "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence.

Numeric Identifier	Definition	Comments and format	Mandatory (M) or optional (O)
<222>	Location	Specify location within sequence; where appropriate, state number of first and last bases/amino acids in feature	M, under the following conditions: If “n,” “Xaa,” or a modified or unusual L-amino acid or modified base was used in a sequence.
<223>	Other Information.	Other relevant information; four lines maximum.	M, under the following conditions: If "n,” “Xaa,” or a modified or unusual L- amino acid or modified base was used in a sequence; if ORGANISM is “Artificial Sequence” or “Unknown”; if molecule is combined DNA/ RNA.
<300>	Publication Information	Leave blank after <30>.	O.
<301>	Authors	Preferably max. of 10 named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials.	O.
<302>	Title		O.
<303>	Journal		O.
<304>	Volume		O.
<305>	Issue		O.
<306>	Pages		O.
<307>	Date	Journal date on which data published; specify as yyyy-mm-dd, MMM-yyyy or Season-yyyy	O.
<308>	Database Accession Number.	Accession number assigned by database, including database name	O.
<309>	Database Entry Date.	Date of entry in database; specify as yyyy-mm-dd or MMM-yyyy	O.
<310>	Patent Document Number.	Document number; for patent-type citations only. Specify as, for example, US 09/ 999,999	O.
<311>	Patent Filing Date.	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	O.
<312>	Publication Date.	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	O.
<313>	Relevant Residues.	FROM (position) TO (position)	O.
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual sequence	M.

2424.01 Informational Requirements for the "Sequence Listing" [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.823 sets forth the informational requirements for the content of a "Sequence Listing" part of the specification that must be submitted in accordance with 37 CFR 1.821(c) as part of the application. See MPEP § 2422.03 for a discussion of 37 CFR 1.821(c).

37 CFR 1.823(a) sets forth the content requirements for a "Sequence Listing". Such requirements include, but are not limited to, sequence identifiers, the order and presentation of items of information, mandatory and optional information, the format as to line spacing, and the use of numeric identifiers.

37 CFR 1.823(b)(1) sets forth a requirement for applications, other than an international application in the international stage, to contain, in the specification of the patent application, an express incorporation by reference of the material submitted as an ASCII plain text file via the USPTO patent electronic filing system or on a read-only optical disc(s) identifying the name of the file, the date of creation, and the size of the file in bytes.

37 CFR 1.823(b)(2) specifically exempts international applications during the international stage from the incorporation by reference requirement in 37 CFR 1.823(b)(1).

37 CFR 1.823(b)(3) sets forth the format and content for a "Sequence Listing" that is submitted either as a PDF image file via the USPTO patent electronic filing system or on physical sheets of paper, in accordance with 37 CFR 1.823(b)(1)(i)-(vi).

2424.02 "Sequence Listing" Numeric Identifiers [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.823(a) sets forth the order and presentation of the items of information in the "Sequence Listing". Each item of information in the "Sequence Listing" must include the appropriate numeric identifier and its accompanying information as in Appendix G to Subpart G of Part 1. Each item of information must begin on a new line with the numeric identifier enclosed in angle brackets. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> must be set forth at the beginning of the "Sequence Listing".

See MPEP § 2431 for a sample "Sequence Listing".

2424.03 Additional Miscellaneous Requirements [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

Throughout 37 CFR 1.823(a), the items of information relating to patent applications and patent publications should be provided keeping in mind the appropriate standards that have been established by the World Intellectual Property Organization (WIPO). In general, an application should be identified by a country code, a number and a filing date, while a published patent document should be identified by a country code, a number and kind code. Proper citation of priority patent applications

is covered in MPEP § 214.04. For published patent documents, the country code, number and kind code will appear on the front page of the document. Unpublished PCT applications are identified by the letters PCT, the country code of the Receiving Office, and the four digit year of filing and a six digit number, e.g., PCT/SE2011/000123. A published PCT application is identified by the letters WO, the four digit year of publication and a six digit number, e.g. WO/2015/009167. Country codes from WIPO Standard ST.3 Annex A and kind codes from WIPO Standard ST.16 are accessible on WIPO's website (**www.wipo.int/standards/en/part_03_standards.html**). See also MPEP § 1851. Questions on proper citation of patent documents should be directed to staff in the Office of International Patent Legal Administration.

See Appendix G to Subpart G of Part 1 of the CFR (reproduced in MPEP § 2424) for information regarding numeric identifiers.

2425 Form and Format for a Nucleotide and/or Amino Acid Sequence Submission as an ASCII Plain Text File under WIPO ST.25 [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.824 Form and format for a nucleotide and/or amino acid sequence submissions as an ASCII plain text file.

(a) A "Sequence Listing" under § 1.821(c)(1) and the CRF required by § 1.821(e) submitted as an ASCII plain text file may be created by any means, such as text editors, nucleotide/amino acid sequence editors, or other custom computer programs; however, the ASCII plain text file must conform to the following requirements:

(1) Must 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) Must be in ASCII plain text, where:

(i) All printable characters (including the space character) are permitted; and

(ii) No nonprintable (ASCII control) characters are permitted, except ASCII CRLF or LF as line terminators.

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

(4) Must contain no more than 74 printable characters in each line.

(5) Pagination is not permitted; the ASCII plain text file must be one continuous file, with no "hard page break" codes and no page numbering.

(b) The ASCII plain text file must contain a copy of a single "Sequence Listing" in a single file and be submitted either:

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

(2) On a 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 ASCII plain text 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).

37 CFR 1.824 sets forth the requirements for sequence submissions as an ASCII plain text file. Any computer editing program may be utilized to produce the ASCII plain text file; however, the resultant file must have the characteristics specified in 37 CFR 1.824.

If the "Sequence Listing" required by 37 CFR 1.821(c) cannot be submitted via USPTO patent electronic filing system because it is larger than 100 megabytes (see MPEP § 2422.03(a)(III)), the "Sequence Listing" may be submitted on read-only optical disc(s) in compliance with 37 CFR 1.52(e). See 37 CFR 1.824(b)(2).

2426 Amendments to Add or Replace a "Sequence Listing" and CRF Copy Thereof Subject to WIPO ST. 25 [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

37 CFR 1.825 Amendments to add or replace a "Sequence Listing" and CRF copy thereof.

(a) Any amendment adding a "Sequence Listing" (§ 1.821(c)) after the application filing date must include:

(1) A "Sequence Listing," in accordance with the requirements of §§ 1.821 through 1.824, submitted as:

(i) An ASCII plain text file, under § 1.821(c)(1), via the USPTO patent electronic filing system or on a read-only optical disc, in compliance with § 1.52(e);

(ii) A PDF file via the USPTO patent electronic filing system; or

(iii) Physical sheets of paper;

(2) A request that the amendment be made:

(i) By incorporation by reference of the material in the ASCII plain text file, in a separate paragraph of the specification, identifying the name of the file, the date of creation, and the size of the file in bytes (*see* § 1.77(b)(5)), for a "Sequence Listing" submitted under § 1.821(c)(1), except when submitted to the United States International Preliminary Examining Authority for an international application; or

(ii) By inserting, after the abstract of the disclosure, a "Sequence Listing" submitted as a PDF file under § 1.821(c)(2) or submitted on physical sheets of paper under § 1.821(c)(3), 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 (specification, claims, drawings) for all sequence data in the "Sequence Listing" in the application as originally filed;

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

(5) A new or substitute CRF under § 1.821(e), if:

(i) (The added "Sequence Listing" is submitted as a PDF file, under § 1.821(c)(2), or on physical sheets of paper, under § 1.821(c)(3); and

(ii) A CRF, under § 1.821(e), was not submitted, not compliant with § 1.824, or not the same as the "Sequence Listing"; and

(6) A statement that the sequence information contained in the CRF is the same as the sequence information contained in the added "Sequence Listing," if submitted as a PDF file, under § 1.821(c)(2), or on physical sheets of paper, under § 1.821(c)(3).

(b) Any amendment to a "Sequence Listing" (§ 1.821(c)) must include:

(1) A replacement "Sequence Listing," in accordance with the requirements of §§ 1.821 through 1.824, submitted as:

(i) An ASCII plain text file, under § 1.821(c)(1), via the USPTO patent electronic filing system, or 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);

(ii) A PDF file via the USPTO patent electronic filing system; or

(iii) Physical sheets of paper;

(2) A request that the amendment be made:

(i) By incorporation by reference of the material in the ASCII plain text file, in a separate paragraph of the specification (replacing any prior such paragraph, as applicable) identifying the name of the file, the date of creation, and the size of the file in bytes (see § 1.77(b)(5)) for a "Sequence Listing" under § 1.821(c)(1), except when submitted to the United States International Preliminary Examining Authority for an international application; or

(ii) By placing, after the abstract of the disclosure, a "Sequence Listing" submitted as a PDF file, under § 1.821(c)(2), or on physical sheets of paper, under § 1.821(c)(3) (replacing any prior "Sequence Listing," as applicable), except when submitted to the United States International Preliminary Examining Authority for an international application;

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

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

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

(6) A new or substitute CRF, under § 1.821(e), with the amendment incorporated therein, if:

(i) The replacement "Sequence Listing" is submitted as a PDF file, under § 1.821(c)(2), or on physical sheets of paper, under § 1.821(c)(3); and

(ii) A CRF, under § 1.821(e), was not submitted, not compliant with § 1.824, or not the same as the submitted "Sequence Listing"; and

(7) A statement that the sequence information contained in the CRF is the same as the sequence information contained in the replacement "Sequence Listing" when submitted as a PDF file, under § 1.821(c)(2), or on physical sheets of paper, under § 1.821(c)(3).

(c) The specification of a complete application, filed on the application filing date, with a "Sequence Listing" as an ASCII plain text file, under § 1.821(c)(1), without an incorporation by reference of the material contained in the ASCII plain text file, must be amended to contain a separate paragraph incorporating by reference the material contained in the ASCII plain text file, in accordance with § 1.77(b)(5), except for international applications during the international stage or national stage.

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

37 CFR 1.825 sets forth the procedures for adding or replacing the "Sequence Listing" and the CRF thereof. Under 37 CFR 1.825(a), adding a "Sequence Listing" after the application filing date involves the submission of: (1) a "Sequence Listing" either as a PDF image file, on physical sheets of paper, or as an ASCII plain text file submitted via the USPTO patent electronic filing system or on read-only optical disc(s), (2) a request that the amendment be made by either incorporation by reference in the case of ASCII plain text files or inserting text into the specification in the case of PDF image files or physical sheets of paper, (3) a statement that indicate the basis for the amendment in the application, as originally filed, (4) a statement that the "Sequence Listing" includes no new matter, (5) a new or substitute CRF when the submission is a PDF image file or on physical sheets of paper, and (6) a statement that the sequence information contained in the CRF is the same as the sequence information contained in the added "Sequence Listing" when the submission is a PDF image file or on physical sheets of paper. (See MPEP § 2428 for further information and sample statements.) If the "Sequence Listing" is filed as an ASCII plain text file only, thereby serving both as a "Sequence Listing" under 37 CFR 1.821(c) and a CRF under 37 CFR 1.821(e), then an incorporation by reference statement is required in the specification (replacing any prior such paragraph) (see 37 CFR 1.77(b)(5)). If the "Sequence Listing" is filed as a PDF image file or on physical sheets of paper, then the "Sequence Listing" must be placed after the abstract of the disclosure. Also, if the added "Sequence Listing" is submitted as a PDF image file or on physical sheets of paper and a separate CRF under 37 CFR 1.821(e) was either not submitted, not compliant with 37 CFR 1.824, or not the same as the submitted "Sequence Listing", then a new or substitute separate CRF is required along with a statement that the sequence information contained in the CRF is the same as the sequence information contained in the added "Sequence Listing". It is noted that an incorporation by reference statement should not be submitted for a separate CRF of a "Sequence Listing" submitted under 37 CFR 1.821(e).

Under 37 CFR 1.825(b), amending the "Sequence Listing" involves the submission of a replacement "Sequence Listing" either as a PDF image file, on physical sheets of paper, or as an ASCII plain text file submitted via the USPTO patent electronic filing system or on read-only optical disc(s), in conjunction with statements that identify the location of all deletions, replacements, or additions to the "Sequence Listing", indicate the basis for the amendment in the application, as originally filed, and that the replacement "Sequence Listing" includes no new matter. (See MPEP § 2428 for further information and sample statements.) These statements may be made by a registered practitioner, the applicant, an inventor, or the person who actually compares the sequence data on behalf of the aforementioned. It is noted that any amendment to the information on a read-only optical disc must be by the way of a replacement read-only optical disc. 37 CFR 1.52(e)(7). If the replacement "Sequence Listing" is filed as an ASCII plain text file, then an incorporation by reference statement is required in the specification (replacing any prior such paragraph) (see 37 CFR 1.77(b)(5)). If the replacement "Sequence Listing" is filed as a PDF image file or on physical sheets of paper, then the replacement "Sequence Listing" must be placed after the abstract of the disclosure. Also, if the replacement "Sequence Listing" is submitted as a PDF image file or on physical sheets of paper and a separate CRF under 37 CFR 1.821(e) was either not submitted, not compliant with 37 CFR 1.824, or not the same as the submitted "Sequence Listing", then a new or substitute separate CRF is required along with a statement that the sequence information contained in the CRF is the same as the sequence information contained in the replacement "Sequence Listing". It is noted that an incorporation by reference statement should not be submitted for a separate CRF of a "Sequence Listing" submitted under 37 CFR 1.821(e). If the "Sequence Listing" required by 37 CFR 1.821(c) was filed on a read-only optical disc

that included files in addition to the "Sequence Listing", a replacement read-only optical disc containing the amended "Sequence Listing" also must contain all of the files of the original read-only optical disc that were not amended. This will ensure that the Office, printer, and public can quickly access all of the current files in an application or patent by referencing only the latest version.

The requirement for statements regarding the absence of new matter follows current practice relating to the submission of substitute specifications, as set forth in 37 CFR 1.125. 37 CFR 1.825(c) requires that the specification must be amended to contain an incorporation by reference statement of the material contained in a "Sequence Listing" ASCII plain text file if the "Sequence Listing" ASCII plain text file was filed on the application filing date without an incorporation by reference statement, except if the "Sequence Listing" ASCII plain text file was filed on the application filing date of an international application regardless of whether the international application is currently in the international stage or the national stage. 37 CFR 1.825(d) addresses the situation where amendments to the "Sequence Listing" are made after a patent has been granted, e.g., by a certificate of correction, reissue or reexamination, and states that such amendments must comply with the requirements of 37 CFR 1.825(b).

2427 Form Paragraphs for Applications Subject to WIPO ST.25 [R-01.2024]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

See MPEP § 608.05 for form paragraphs which should be used when notifying applicant that a read-only optical disc submitted in accordance with 37 CFR 1.52(e) (i.e., containing a "Computer Program Listing Appendix", "Sequence Listing", and/or "Large Tables") does not comply with all of the requirements of the 37 CFR 1.52(e). See also MPEP § 608.05(b) for form paragraphs which should be used when a table submitted on read-only optical disc does not comply with 37 CFR 1.52(e).

In order to expedite the processing of applications, minor errors pertaining to compliance with the sequence rules discovered after examination has begun may be handled with the next Office action. Examples of minor errors are: missing any necessary sequence identifiers in the specification/drawings; missing or providing an incomplete incorporation by reference of the "Sequence Listing" in the specification (if required); and missing statement that the sequence information contained in the "Sequence Listing" is identical to the sequence information contained in the computer readable form (if required). Since the application is ready for examination, the examiner may act on the application and include any objections to the application based on minor errors related to the "Sequence Listing" or the CRF of the "Sequence Listing" with the Office action.

¶ 24.01 Heading for Sequence Requirements

REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE AND/OR AMINO ACID SEQUENCE DISCLOSURES

Items 1) and 2) provide general guidance related to requirements for sequence disclosures.

1) 37 CFR 1.821(c) requires that patent applications which contain disclosures of nucleotide and/or amino acid sequences that fall within the definitions of 37 CFR 1.821(a) must contain a "Sequence Listing," *as a separate part of the disclosure* , whichpresents the nucleotide and/or amino acid sequences and associated information using the symbols and format in accordance with the requirements of 37 CFR 1.821 1.825 . This "Sequence Listing" part of the disclosure may be submitted:

a) In accordance with 37 CFR 1.821(c)(1) via the USPTO patent electronic filing system (see Section I.1 of the Legal Framework for Patent Electronic System(**https://www.uspto.gov/PatentLegalFramework**), hereinafter "Legal Framework") as an ASCII text file, together with an incorporation-by-reference of the material in the ASCII text file in a separate paragraph of the specification as required by 37 CFR1.823(b)(1) identifying:

i) the name of the ASCII text file;

ii) the date of creation; and

iii) the size of the ASCII text file in bytes;

b) In accordance with 37 CFR1.821(c)(1) on read-only optical disc(s) as permitted by 37 CFR1.52(e)(1)(ii) , labeled according to 37 CFR1.52(e)(5) , with an incorporation-by-reference of the material in the ASCII text file

according to 37 CFR1.52(e)(8) and 37 CFR1.823(b)(1) in a separate paragraph of the specification identifying:

i) the name of the ASCII text file;

ii) the date of creation; and

iii) the size of the ASCII text file in bytes;

c) In accordance with 37 CFR 1.821(c)(2) via the USPTO patent electronic filing system as a PDF file(not recommended); or

d) In accordance with 37 CFR 1.821(c)(3) on physical sheets of paper (not recommended).

2) When a “Sequence Listing” has been submitted as a PDF file as in 1.c) above (37 CFR 1.821(c)(2)) or on physical sheets of paper as in 1. d) above (37 CFR 1.821(c)(3) , 37 CFR1.821(e)(1)), requires a computer readable form (CRF) of the “Sequence Listing” in accordance with the requirements of 37 CFR 1.824 .

a) If the "Sequence Listing" required by 37 CFR 1.821(c) is filed via the USPTO patent electronic filing system as a PDF,then 37 CFR 1.821(e)(1)(ii) or 1.821(e)(2)(ii) requires submission of a statement that the"Sequence Listing" content of the PDF copy and the CRF copy (the ASCII text file copy) are identical.

b) If the "Sequence Listing" required by 37 CFR 1.821(c) is filed on paper or read-only optical disc,then 37 CFR 1.821(e)(1)(ii) or 1.821(e)(2)(ii) requires submission of a statement that the"Sequence Listing" content of the paper or read-only optical disc copy and the CRF are identical.

Specific deficiencies and the required response to this Office Action are as follows:

Examiner Note:

1. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

2. This form paragraph must be followed by any of form paragraphs 24.0224.16 .

¶ 24.02 No Sequence Listing part of the disclosure and No CRF

Specific deficiency - This application fails to comply with the requirements of 37 CFR 1.821 - 1.825 because it does not contain a "Sequence Listing" as a separate part of the disclosure or a CRF of the “Sequence Listing”.

Required response - Applicant must provide:

• A "Sequence Listing" part of the disclosure; **together with**

• **An amendment specifically directing its entry into the application in accordance with37 CFR 1.825(a)(2) ;**

• A statement that the "Sequence Listing" includes no new matter as required by 37 CFR 1.825(a)(4); and

• A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.825(a)(3).

• If the "Sequence Listing" part of the disclosure is submitted according to item 1) a) or b) above,

Applicant must also provide:

o A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3) and 1.125 inserting the required incorporation-by-reference paragraph, consisting of:

- A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

- A copy of the amended specification without markings (clean version); and

- A statement that the substitute specification contains no new matter.

• If the "Sequence Listing" part of the disclosure is submitted according to item 1) b), c), or d) above, Applicant must also provide:

o A CRF in accordance with 37 CFR 1.821(e)(1) or 1.821(e)(2) as required by 1.825(a)(5); and

o A statement according to item 2) a) or b) above.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application that has no "Sequence Listing" part of the disclosure as required by 37 CFR 1.821(c) and no CRF as required by 37 CFR 1.821(e).

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.03 No Sequence Listing part of the disclosure and Defective CRF

Specific deficiency - This application fails to comply with the requirements of 37 CFR 1.821 - 1.825 because it does not contain a "Sequence Listing" as a separate part of the disclosure and the CRF of the "Sequence Listing" is defective.

Required response - Applicant must provide:

• A "Sequence Listing" part of the disclosure, as described above in item 1); **together with**

• **An amendment specifically directing its entry into the application in accordance with 37 CFR 1.825(a)(2);**

• A statement that the "Sequence Listing" includes no new matter as required by 37 CFR 1.825(a)(4); and

• A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.825(a)(3).

• If the "Sequence Listing" part of the disclosure is submitted according to item 1) a) or b) above, Applicant must also provide:

o A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3) and 1.125 inserting the required incorporation-by-reference paragraph, consisting of:

- A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

- A copy of the amended specification without markings (clean version); and

- A statement that the substitute specification contains no new matter.

• If the "Sequence Listing" part of the disclosure is submitted according to item 1) c) or d) above, Applicant must also provide:

o A CRF in accordance with 37 CFR 1.821(e)(1) or 1.821(e)(2) as required by 37 CFR 1.825(a)(5); and

o A statement according to item 2) a) or b) above.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application that has no "Sequence Listing" part of the disclosure as required by 37 CFR 1.821(c) and the CRF as required by 37 CFR 1.821(e) is defective.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.17 Improper CRF transfer request

Specific deficiency - The present application contains a "Sequence Listing" submitted as either a PDF file pursuant to 37 CFR 1.821(c)(2) or as physical sheets of paper pursuant to 37 CFR 1.821(c)(3). No computer readable form (CRF) of the "Sequence Listing" pursuant to 37 CFR 1.821(e)(1) has been received. In lieu of the CRF, Applicant has filed a request to transfer the CRF from a related or other application of the applicant to the present application to comply with the requirement in 37 CFR 1.821(e)(1). As of November 15, 2021, the practice of transferring a CRF from a previously-filed application of applicant into the present application in order to comply with 37 CFR 1.821(e)(1) has been eliminated.

Required response - Applicant must provide:

• A new CRF of the "Sequence Listing" in accordance with 37 CFR 1.821(e)(1)(i) or 1.821(e)(2)(i); and

• A statement that the content of the CRF is identical to the "Sequence Listing" part of the disclosure submitted as a PDF file (37 CFR 1.821(c)(2)) or on physical sheets of paper (37 CFR 1.821(c)(3)), as required by 37 CFR 1.821(e)(1)(ii) or 1.821(e)(2)(ii).

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.05 The "Sequence Listing" part of the disclosure and the CRF are not the same

Specific deficiency - This application fails to comply with the requirements of 37 CFR 1.821 - 1.825 because the "Sequence Listing" part of the disclosure submitted as a PDF file (37 CFR 1.821(c)(2)) or on physical sheets of paper (37 CFR 1.821(c)(3))is not the same as the CRF of the "Sequence Listing" as required by 37 CFR 1.821(e)(1)(ii) or 1.821(e)(2)(ii).

Required response - Applicant must provide:

• A replacement "Sequence Listing" as described above in items 1) c) or d) in accordance with 37 CFR 1.825(b)(1)(ii) or (iii); **as well as**

• **An amendment specifically directing its entry into the application as required by 37 CFR 1.825(b)(2)(ii);**

• A statement that identified the locations of any deletions, replacements or additions to the "Sequence Listing" as required by 37 CFR 1.825(b)(3);

• A statement that the "Sequence Listing" added by amendment includes no new matter as required by 37 CFR 1.825(b)(5);

• A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.825(b)(4); and

• A statement that the content of the previously-filed CRF is identical to the "Sequence Listing" part of the disclosure added by amendment as required by 37 CFR 1.825(b)(7), where provided under item 1) c) or d) (note that where a "Sequence Listing" part of the disclosure is provided under item 1) a) or b), the text file will also serve as the CRF, and the statement of identity is not required);

OR

• A CRF as required by 37 CFR 1.821(e)(1) or 1.821(e)(2); and

• A statement that the content of the CRF is identical to the "Sequence Listing" part of the disclosure previously submitted as a PDF file (37 CFR 1.821(c)(2)) or on physical sheets of paper (37 CFR 1.821(c)(3)), as required by 37 CFR 1.821(e)(1)(ii) or 1.821(e)(2)(ii).

Examiner Note:

1. This form paragraph must be preceded by 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application in which the "Sequence Listing" part of the disclosure as required by 37 CFR 1.821(c) is not the same as the CRF as required by 37 CFR 1.821(e).

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.06 Missing statement that the "Sequence Listing" (paper or PDF) and the CRF are the same

Specific deficiency - This application fails to comply with the requirements of 37 CFR 1.821 - 1.825 because the application does not contain a statement that the CRF is identical to the "Sequence Listing" part of the disclosure, as described above in item 1), as required by 37 CFR 1.821(e)(1)(ii) or 1.821(e)(2)(ii).

Required response - Applicant must provide such statement.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application that is missing the statement that the CRF is identical to the "Sequence Listing" part of the disclosure.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.07 No Computer Readable Form (CRF) submitted

Specific deficiency - This application contains a "Sequence Listing as a PDF file (37 CFR 1.821(c)(2)) or as physical sheets of paper (37 CFR 1.821(c)(3), but fails to comply with the requirements of 37 CFR 1.821 - 1.825 because a copy of the "Sequence Listing" in computer readable form (CRF) has not been submitted as required by 37 CFR 1.821(e)(1)(i) or 1.821(e)(2)(i) as indicated in item 2) above.

Required response - Applicant must provide:

• A new CRF of the "Sequence Listing" in accordance with 37 37 CFR 1.821(e)(1)(i) or 1.821(e)(2)(i) and

• A statement that the content of the CRF is identical of the "Sequence Listing" part of the disclosure, submitted as a PDF file (37 CFR 1.821(c)(2)) or on physical sheets of paper (37 CFR 1.821(c)(3)), as required by 37 CFR 1.821(e)(1)(ii) or 1.821(e)(2)(ii).

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application that is missing the CRF.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.08 Computer Readable Form (CRF) contains error(s) according to STIC report

Specific deficiency - This application fails to comply with the requirements of 37 CFR 1.821 - 1.825. This application contains a "Sequence Listing" as a PDF file (37 CFR 1.821(c)(2)) or as physical sheets of paper (37 CFR 1.821(c)(3)). A copy of the "Sequence Listing" in computer readable form (CRF) has been submitted; however, the content of the CRF does not comply with one or more of the requirements of 37 CFR 1.822 through 1.824, as indicated in the "Error Report" that indicates the "Sequence Listing" could not be accepted. Refer to attachment or document "Computer Readable Form (CRF) for Sequence Listing – Defective" dated **[1]**.

Required response – Applicant must provide:

• A replacement "Sequence Listing" part of the disclosure, as described above in item 1); **together with**

• **An amendment specifically directing its entry into the application in accordance with 37 CFR 1.825(b)(2)** ;

• A statement that the "Sequence Listing" includes no new matter as required by 37 CFR 1.825(b)(5); and

• A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.825(b)(4).

• If the replacement "Sequence Listing" part of the disclosure is submitted according to item 1) a) or b) above, Applicant must also provide:

o A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3) and 1.125 inserting the required incorporation-by-reference paragraph, consisting of:

- A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

- A copy of the amended specification without markings (clean version); and

- A statement that the substitute specification contains no new matter and

- An amendment to the specification to remove the "Sequence Listing previously submitted as a PDF file (37 CFR 1.821(c)(2)) or as physical sheets of paper (37 CFR 1.821(c)(3))

• If the replacement "Sequence Listing" part of the disclosure is submitted according to item 1) c) or d) above, Applicant must also provide:

o A CRF in accordance with 1.821(e)(1) or 1.821(e)(2) as required by 37 CFR 1.825(b)(6)(ii); and

o Statement according to item 2) a) or b) above.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application where the CRF is defective.

4. In bracket 1, insert the date of the appropriate document.

5. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.09 Computer Readable Form (CRF) damaged or unreadable

Specific deficiency - The ASCII .txt file purported to contain the computer readable form (CRF) copy of the "Sequence Listing" filed with this application in accordance with 37 CFR 1.821(c) has been found to be damaged, unreadable, or otherwise contains an error as indicated on document "Computer Readable Form (CRF) for Sequence Listing - Defective" dated.

Required response – Applicant must provide:

• a replacement "Sequence Listing" in the form of an ASCII plain text file under 37 CFR 1.821(c) as provided for in 37 CFR 1.825(b)(1)(i), together with

• An amendment specifically directing its entry into the application in accordance with 37 CFR 1.825(b)(3);

• A statement that the "Sequence Listing" includes no new matter as required by 37 CFR 1.825(b)(5); and

• A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.825(b)(4).

• A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3), and 1.125 inserting the required incorporation-by-reference paragraph, consisting of:

• A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

• A copy of the amended specification without markings (clean version); and

• A statement that the substitute specification contains no new matter;

OR

• A "Sequence Listing" part of the disclosure, as described above in item 1 c) or 1 d) as provided for in 37 CFR 1.825(b)(1)(ii) or 1.825(b)(1)(iii); **together with**

• An amendment specifically directing its entry into the application in accordance with 37 CFR 1.825(b)(2);

• A statement that the "Sequence Listing" includes no new matter as required by 37 CFR 1.825(b)(5); and

• A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.825(b)(4).

• When the "Sequence Listing" part of the disclosure is submitted according to item 1 c), or 1 d) above, Applicant must also provide:

• A CRF in accordance with 37 CFR 1.821(e)(1) as required by 37 CFR 1.825(b)(6); and

• a statement according to item 2) a) or b) above.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application where the CRF is damaged or unreadable, e.g., SCORE - CRF Problem Report.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.10 Sequence IDs not present in the specification

Specific deficiency - Nucleotide and/or amino acid sequences appearing in the specification are not identified by sequence identifiers in accordance with 37 CFR 1.821(d).

Required response – Applicant must provide:

• a substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3) and 1.125 inserting the required sequence identifiers, consisting of:

o A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

o A copy of the amended specification without markings (clean version); and

o A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application where reference has not been made to the sequence by use of the sequence identifier, preceded by "SEQ ID NO:" 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.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.11 Sequence IDs not present in the drawings

Specific deficiency - Nucleotide and/or amino acid sequences appearing in the drawings are not identified by sequence identifiers in accordance with 37 CFR 1.821(d). Sequence identifiers for nucleotide and/or amino acid sequences must appear either in the drawings or in the Brief Description of the Drawings.

Required response – Applicant must provide:

• Replacement and annotated drawings in accordance with 37 CFR 1.121(d) inserting the required sequence identifiers;

AND/OR

• a substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3) and 1.125 inserting the required sequence identifiers into the Brief Description of the Drawings, consisting of:

o A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

o A copy of the amended specification without markings (clean version); and

o A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application where reference has not been made to the sequence by use of the sequence identifier, preceded by "SEQ ID NO:" in either the text of the drawings or the Brief Description or the Drawings.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.12 Sequences present in the specification or drawings that are not in the CRF or listing

Specific deficiency - This application contains sequence disclosures in accordance with the definitions for nucleotide and/or amino acid sequences set forth in 37 CFR 1.821(a)(1) and (a)(2). However, this application fails to comply with the requirements of 37 CFR 1.821 - 1.825.

The sequence disclosures are located **[1]**.

Required response – Applicant must provide:

• A "Sequence Listing" part of the disclosure, as described above in item 1); **as well as**

• **An amendment specifically directing entry of the "Sequence Listing" part of the disclosure into the application in accordance with 1.825(b)(2);**

• A statement that the "Sequence Listing" includes no new matter in accordance with 1.825(b)(5); and

• A statement that indicates support for the amendment in the application, as filed, as required by 37 CFR 1.825(b)(4).

• If the "Sequence Listing" part of the disclosure is submitted according to item 1) a) or b) above, Applicant must also provide:

o A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3) and 1.125 inserting the required incorporation-by-reference paragraph, consisting of:

- A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

- A copy of the amended specification without markings (clean version); and

- A statement that the substitute specification contains no new matter;

• If the "Sequence Listing" part of the disclosure is submitted according to item 1) b), c), or d) above, Applicant must also provide:

o A replacement CRF in accordance with 1.825(b)(6); and

o Statement according to item 2) a) or b) above.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should be used for an application containing sequence disclosures that are not contained in the Sequence Listing or CRF.

4. In bracket 1, insert the specific location of the sequence disclosures that are not contained in the Sequence Listing or CRF.

5. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.13 Missing or Defective Incorporation by Reference Paragraph

Specific deficiency - The Incorporation by Reference paragraph required by 37 CFR 1.821(c)(1) is missing or incomplete. See item 1) a) or 1) b) above.

Required response – Applicant must provide:

• A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3) and 1.125 inserting the required incorporation-by-reference paragraph, consisting of:

o A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

o A copy of the amended specification without markings (clean version); and

o A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should only be used for a sequence listing under 1) a) or 1) b) in form paragraph 24.01, where the incorporation-by-reference paragraph is missing.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.14 Amendment Missing Instruction to Enter the "Sequence Listing" into the Application

Specific deficiency – The "Sequence Listing" has not been entered into the application because the amendment does not direct entry of either the "Sequence Listing" (as required by 37 CFR 1.825(a)(2) or 1.825(b)(2)) or contain the required Incorporation by Reference paragraph into the application.

Required response – Applicant must provide:

• A substitute specification in compliance with 37 CFR 1.52, 1.121(b)(3) and 1.125 inserting the required incorporation-by-reference paragraph, consisting of:

o A copy of the previously-submitted specification, with deletions shown with strikethrough or brackets and insertions shown with underlining (marked-up version);

o A copy of the amended specification without markings (clean version); and

o A statement that the substitute specification contains no new matter.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should only be used where the instruction to enter the "Sequence Listing" into the application is missing.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.15 Amendment Missing Statement of No New Matter

Specific deficiency – The "Sequence Listing" has not been entered into the application because the required statement of no new matter is missing. See 37 CFR 1.825(a)(4) or 1.825(b)(5).

Required response – Applicant must provide:

• A proper statement of no new matter.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should only be used for an amendment that is missing the statement of no new matter.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

¶ 24.16 Amendment Missing Statement of Support

Specific deficiency – The "Sequence Listing" has not been entered into the application because the required statement of support is missing. See 37 CFR 1.825(a)(3) or 1.825(b)(4).

Required response – Applicant must provide:

• A proper statement of support.

Examiner Note:

1. This form paragraph must be preceded by form paragraph 24.01.

2. This form paragraph should only be used for sequence listing non-compliance where a compliant sequence listing is not required for examination of the application.

3. This form paragraph should only be used for an amendment that is missing the statement of support.

4. This form paragraph may be followed by one or more deficiency form paragraphs.

2428 Sample Statements under WIPO ST.25 [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

Sample language for the statements required to support sequence rule submissions is provided below. These statements are given by way of example only and are not exhaustive; other language may, of course, be used. For the statements that relate to the assertion that the sequence information contained in the PDF image file or the physical sheets of paper and the computer readable form are "identical" or "the same," it is acknowledged that there may be some non-substantive differences between the two, e.g., page numbers and page breaks may be present in the PDF image file or the physical sheets of paper but not in the computer readable form thereof. This requirement for sameness relates to the informational content of the PDF image file or the physical sheets of paper and the CRF relevant to the requirements of the sequence rules.

37 CFR 1.821(e)(1)(ii), (2)(ii), 1.825(a)(6), or (b)(7) - I hereby state that the sequence information contained in the computer

readable form is identical to the sequence information contained in the "Sequence Listing" submitted as a PDF image file or on physical sheets of paper.

37 CFR 1.825(a)(3) and (4) - I hereby state that the amendment adding a "Sequence Listing", made in accordance with 37 CFR 1.825(a) is supported in the application, as originally filed, at [insert the basis for all sequence data in the "Sequence Listing" with specific references to particular parts of the application (specification, claims, drawings) as originally filed]. I hereby state that the "Sequence Listing" does not include new matter.

37 CFR 1.825(b)(3), (4), and (5) - - I hereby state that the previous "Sequence Listing" that is replaced by the currently submitted replacement "Sequence Listing" in accordance with 37 CFR 1.825(b) is amended by [identify the location of all deletions, replacements, or additions to the previous "Sequence Listing"]. I hereby state that the replacement "Sequence Listing", made in accordance with 37 CFR 1.825(b) is supported in the application, as originally filed, at [insert the basis for all sequence data in the "Sequence Listing" with specific references to particular parts of the application (specification, claims, drawings) as originally filed]. I hereby state that the replacement "Sequence Listing", submitted in accordance with 37 CFR 1.825(b), includes no new matter.

2429 Helpful Hints for Sequence Rules Compliance under WIPO ST.25 [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

The Office has the following tips regarding sequence rules in compliance with WIPO ST.25.

—Compliance is not a filing date issue.

—Compliance is not a 35 U.S.C. 112 issue.

—Compliance is not a 35 U.S.C. 119/120 issue.

—Compliance is not *per se* a new matter issue. The standard for resolution of inconsistencies between the "Sequence Listing" (submitted as an ASCII plain text file, on read-only optical disc(s), as a PDF image file, or on physical sheets of paper pursuant to 37 CFR 1.821(c)) and a separate computer readable form thereof pursuant to 37 CFR 1.821(e) (if required) and/or errors in the "Sequence Listing" is based on the new matter standard.

—Compliance can be achieved via amendment. See MPEP § 2426 for additional information regarding amendments to add or replace a "Sequence Listing" and CRF thereof. See 37 CFR 1.825.

—If sequence information is submitted in an application filed under 35 U.S.C. 111(a) or 35 U.S.C. 371 as an ASCII plain text file in compliance with 37 CFR 1.824 via the USPTO patent electronic filing system or on read-only optical disc(s) and applicant has not filed a "Sequence Listing" as a PDF image file or on physical sheets of paper, the ASCII plain text file will serve as both the "Sequence Listing" under 37 CFR 1.821(c) and the CRF of the "Sequence Listing" under 37 CFR 1.821(e). See 37 CFR 1.821(e)(1). Thus, the following are not required and should not be submitted: (1) a second copy of the "Sequence Listing" as a PDF image file or on physical sheets of paper; and (2) a statement under 37 CFR 1.821(e)(1)(ii) or (2)(ii) (indicating that the sequence information contained in the "Sequence Listing" under 37 CFR 1.821(c) and CRF copy of the "Sequence Listing" under 37 CFR 1.821(e)(1)(ii) or 37 CFR 1.21(o) are identical). Any "Sequence Listing" submitted as an ASCII plain text file via the USPTO patent electronic filing system or on read-only optical disc(s) under 37 CFR 1.52(e) and in compliance with 37 CFR 1.821(c) will be excluded when determining the application size fee required by 37 CFR 1.16(s) or 37 CFR 1.492(j) as per 37 CFR 1.52(f)(1) and (2). See MPEP § 2422.03(a) for additional information. See MPEP § 2422.03(a) for additional information.

—The USPTO encourages applicants to file their patent applications via the USPTO patent electronic filing system and imposes a surcharge for non-electronic filing of an original patent application (excluding reissue, design, plant, and provisional applications). Filing a "Sequence Listing" via USPTO patent electronic filing system as a PDF image file or on physical sheets of paper is not recommended. A "Sequence Listing" in PDF format or on physical sheets of paper is treated as the "Sequence Listing" required by 37 CFR 1.821(c) and requires filing of both a separate CRF and a statement that the "Sequence Listing" and the CRF are identical in an application filed under 35 U.S.C. 111(a) regardless of if the sequence information is also filed as an ASCII plain text file, as required by 37 CFR 1.821(e)(1)(ii), and in a national stage application filed under 35 U.S.C. 371 if the sequence information is not also filed as an ASCII plain text file, as required by 37 CFR 1.821(e)(2)(ii). In addition, a "Sequence Listing" submitted in PDF format or on physical sheets of paper as part of the specification is not excluded when determining the application size fee required by 37 CFR 1.16(s) or 1.492(j). See 37 CFR 1.52(f)(1) and (2).

—For international applications (PCT), the check list of the PCT Request filed with the international application must contain an indication that the sequence listing, filed with the PCT application on the international filing date forms part of the international application. See MPEP § 2422.03(a), subsection IV, for information specific to filing sequence listings in international applications (PCT) via the USPTO patent electronic filing system.

—Applicants are reminded that for fee purposes, a table of sequences is not a "Sequence Listing". Such tables are considered part of the specification and are included when determining the application size fee required by 37 CFR 1.16(s) or 1.492(j). See 37 CFR 1.52(f)(1) and (2).

—Applicants are encouraged to draft their specifications such that sequence data that is not essential material is not required to be included in a "Sequence Listing". 37 CFR 1.21(o) and 1.52(f)(3) provide that the submission of an oversized "Sequence Listing" (a mega-"Sequence Listing") of 300 MB or more are subject to additional fees. A mega-"Sequence Listing", in particular, often include sequences that are available in the prior art, are not essential material, and could have been described instead, for example, by name and a publication or accession reference.

—Failure to reply to sequence compliance issues in a timely manner may reduce any patent term adjustment. Patent applications filed under 35 U.S.C. 111(a) on or after December 18, 2013, and international patent applications in which the national stage commenced under 35 U.S.C. 371 on or after December 18, 2013, may be subject to reductions in patent terms adjustment pursuant to 37 CFR 1.704(c)(13) if they are not in condition for examination within eight months from the filing date or date of commencement, respectively. "In condition for examination" includes compliance with 37 CFR 1.821 - 1.825 (see 37 CFR 1.704(f)).

—The copy of the "Sequence Listing" required by 37 CFR 1.821(c) is an integral part of the application. If submitted as a PDF image file or on physical sheets of paper, the "Sequence Listing" must begin on a new page, should appear at the end of the application, and preferably should be numbered independently of the numbering of the remainder of the application. The new page that begins the "Sequence Listing" should be entitled "Sequence Listing." See 37 CFR 1.823(b)(3). If not submitted as such at filing, the "Sequence Listing" must be inserted into the application via amendment, e.g., by preliminary amendment. See 37 CFR 1.825. If submitted as an ASCII plain text file via the USPTO patent electronic filing system or read-only optical disc, the specification must contain an incorporation by reference of the material, except for a national stage entry under 37 CFR 1.495(b)(1), where the "Sequence Listing" has been previously communicated to the International Bureau or originally filed in the USPTO and complies with Patent Cooperation Treaty Rule 5.2.

—A replacement "Sequence Listing" must be used to amend a "Sequence Listing" regardless of whether the "Sequence Listing" was filed as an ASCII plain text file via the USPTO patent electronic filing system, on read-only optical disc(s) (37 CFR 1.821(c)(1), as a PDF file via the USPTO patent electronic filing system (37 CFR 1.821(c)(2), or on

physical sheets of paper (37 CFR 1.821(c)(3). See MPEP § 2426 for additional information regarding amendments to add or replace a "Sequence Listing" and CRF copy thereof.

—The practice of computer readable form transfers from one application to another has been eliminated.

—Angle brackets and numeric identifiers listed in 37 CFR 1.823 and Appendix G to Subpart G of Part 1 of the CFR (reproduced in MPEP § 2424) are very important for our database. Extra punctuation should not be used in a "Sequence Listing".

—A "Sequence Listing" (37 CFR 1.821(c)) or a separate CRF of a "Sequence Listing" (37 CFR 1.821(e)) as an ASCII plain text file cannot contain page numbers. Page numbers should only be placed on PDF image files or on physical sheets of paper of the "Sequence Listing".

—The PatentIn computer program is not the only means by which to comply with the rules. Any word processing program can be used to generate a "Sequence Listing" if it has the capability to convert a file into ASCII plain text. However, use of a word processing program to generate or amend a "Sequence Listing" file is discouraged. Word processing programs often introduce unintended changes to the "Sequence Listing" that render the listing unacceptable. Use of a plain text editor to generate or edit a "Sequence Listing" is recommended.

—If a word processing program is used to generate a "Sequence Listing", hard page break controls should not be used and margins should be adjusted to the smallest setting.

—Word processing files should not be submitted to the Office; the "Sequence Listing" generated by a word processing file should be saved as an ASCII plain text file for submission. Most word processing programs provide this feature.

—Statements in accordance with 37 CFR 1.821(e)(1)(ii), (e)(2)(ii), (e)(3)(iii), and (h) and 37 CFR 1.825(a)(3), (a)(4), (a)(6), (b)(3), (b)(4), (b)(5), and (b)(7) and proper labeling for read-only optical disc(s) in accordance with 37 CFR 1.52(e)(5)(vi) should be noted. Sample statements to support filings and submissions in accordance with 37 CFR 1.821 through 1.825 are provided in MPEP § 2428 Sample Statements.

—Use Box SEQUENCE. See MPEP § 2433.

—On nucleotide sequences, since only single strands may be depicted in the "Sequence Listing", show strands in 5' to 3' direction from left to right in accordance with 37 CFR 1.822(c)(5).

—The single stranded nucleotide depicted in the "Sequence Listing" may represent a strand of a nucleotide sequence that may be single or double stranded which may be, further, linear or circular. An amino acid sequence or peptide may be linear or circular.

—Numeric identifiers "<140>, Current Application Number," "<141>, Current Filing Date," "<150>, Prior Application Number," and "<151>, Prior Application Filing Date," should appear in the "Sequence Listing" in all cases. If the information about the current application is not known or is unavailable at the time of completing the "Sequence Listing", then the lines following numeric identifiers <140> and <141> should be left blank. This would normally be the case when the "Sequence Listing" is included in a newly filed application. Similarly, if information regarding prior applications is inapplicable, or not known at the time of completing the "Sequence Listing" but will be later filed, then the numeric identifiers <150> and <151> should appear with the line following the numeric identifiers left blank.

—The mandatory items of information that must be included in a "Sequence Listing" are identified in the table of numeric identifiers set forth Appendix G to Subpart G of Part 1 of the CFR. See also MPEP § 2424.02.

—Pursuant to 37 CFR 1.83(a), sequences that are included in a "Sequence Listing" should not be duplicated in the drawings. However, significant sequence characteristics that are not readily conveyed by the data in the "Sequence Listing" may be depicted in a drawing figure. However, the sequence information so conveyed must still be

included in a "Sequence Listing" if the sequence falls within the definition set forth in 37 CFR 1.821(a), and the sequence identifier ("SEQ ID NO:X or the like") must be used, either in the drawing or in the "Brief Description of the Drawings." See MPEP § 2422.02 for additional information.

—Inosine may be represented by the use of "I" in the features section, otherwise use "n."

—Stop codons that are represented by an asterisk are not permitted in amino acid sequences.

—Punctuation should not be used in a sequence to indicate unknown nucleotide bases or amino acid residues or to delimit active or functional regions of a sequence. These regions should be noted as Features of the sequence per Appendix G to Subpart G of Part 1 of the CFR (see numeric identifiers <220> - <223>. (Appendix G is reproduced in MPEP § 2424).

—The presence of an unnatural amino acid in a sequence does not have the same effect as the presence of a D-amino acid. The sequence may still be subject to the rules even though one or more of the amino acids is not naturally occurring.

—Cyclic and branched peptides are causing some confusion in the application of the rules. Specific questions should be directed to Sequence Systems Service Center of the Scientific and Technical Information Center at 571-272-2510.

—A cyclic peptide with a tail is regarded as a branched sequence, and thereby exempt from the rules, if all bonds adjacent to the amino acid from which the tail emanates are normal peptide bonds.

—Sequences that have variable-length regions depicted as, for example, Ala Ala Leu Leu (Xaa Xaa)$_n$ Ile Pro where n=0-234 or agccttgggaca(nnnnn)$_m$gtcatt where m=0-354 or Ser Met Ala Xaa Ser where Xaa could be 1, 2, 3, 4 and/or 5 amino acids must still comply with the Sequence Rules. The method to use is to repeat the variable-length region as many times as the maximum length and specify in the Features section that the amino acid (or nucleotide) at a specified position is either absent or present. The variables Xaa and n may stand for only one residue, hence the need to repeat the variable. The correct way to submit the third example is Ser Met Ala Xaa Xaa Xaa Xaa Xaa Ser combined with an explanation in the Features section of the listing that any one or all of amino acids 4-8 can either be present or absent.

—Single letter amino acid abbreviations are not acceptable within the "Sequence Listing" but may appear elsewhere in the application.

—Zero (0) is not used when the numbering of amino acids uses negative numbers to distinguish the mature protein.

—Subscripts or superscripts are not permitted in a "Sequence Listing".

—The exclusive conformance requirement of 37 CFR 1.821(b) requires that any amendment of the sequence information in a "Sequence Listing" be accompanied by an amendment to the corresponding information, if any, embedded in the text of the specification or presented in a drawing figure.

—A mandatory feature is required to cover every "n" or "Xaa" used in a sequence. The feature consists of numeric identifiers <220>, <221>, <222>, and <223>. Numeric identifier <220> should remain blank, numeric Identifier <221> should be selected from Appendices E and F to Subpart G of Part 1 of the CFR (reproduced in MPEP § 2422(I)), numeric identifier <222> should identify the location of the "n" or "Xaa" within the sequence, and numeric identifier should <223> specify what the "n" or "Xaa" can be. When all of the "n" or "Xaa" variables in a sequence are equal to the same thing, a range of the entire sequence can be given for numeric identifier <222> to cover all of the "n" or "Xaa" designators in one feature.

—Remove all non-ASCII characters from the .txt file. For example, an symbol should be spelled out as "alpha."

—Tabs are non-ASCII characters. Do not use tabs in "Sequence Listing" .txt files.

—Make all explanations in a feature section consistent with the molecule type in numeric identifier <212>. For example, if the sequence is type "PRT" do not describe the sequence in a feature section as a "synthetic oligonucleotide."

—A response for numeric identifier <130>, "File Reference," is mandatory if the numeric identifier <140> is not present, e.g., when the "Sequence Listing" is filed before the application number has been assigned. At least one of a numeric identifier <130> with docket number or numeric identifier <140> with current application number must be in the "Sequence Listing". This information is used to ensure that ASCII plain text files are correctly matched to their corresponding applications.

—If a "Sequence Listing" is modified by the addition or deletion of sequences, remember to update the total number of sequences in numeric identifier <160>.

—Numeric identifier <213> can only be one of three choices: Scientific name (i.e. Genus/species), Unknown, or Artificial Sequence. Do not add any extraneous information about the sequence, such as a gene names, in this field. Do not use common names for species. For example, human should be "Homo sapiens" and cow should be "Bos taurus." If a specific genus/species is unknown, use the reply "Unknown" in numeric identifier <213> and add whatever information is known into numeric identifier <223> of the feature section. For example, if only the family "Saccharomycetaceae" is known, numeric identifier <213> should state "Unknown" and numeric identifier <223> could state "fungus of the family Saccharomycetaceae."

—For all sequences using "Unknown" or "Artificial sequence" for numeric identifier <213>, a mandatory feature is required to explain the source of the genetic material. The feature consists of <220>, which remains blank, and <223>, which states the source of the genetic material. To explain the source, if the sequence is put together from several organisms, please list those organisms. If the sequence is made in the laboratory, please indicate that the sequence is synthesized.

—Only use abbreviations that are specifically defined in "WIPO Standard ST.25 (2009)" or that are well known. Do not use abbreviations that are specific to the application at issue and would not be clear to someone who had not read the invention description. When in doubt, use the full name rather than an abbreviation.

—Note that if a "Sequence Listing" provided as an ASCII plain text file or a separate CRF of a "Sequence Listing" is rejected and an error report issued, the errors listed are exemplary and may not be a complete list of all errors in the "Sequence Listing" file. The applicant is required to review the "Sequence Listing" in its entirety and correct all instances of similar errors.

—Any inquiries regarding a specific "Sequence Listing" provided as an ASCII plain text file or a separate CRF of a "Sequence Listing" that has been processed by the Office should be directed to the Sequence Systems Service Center of the Scientific and Technical Information Center at 571-272-2510.

2430 PatentIn Information [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

In those areas of biotechnology in which nucleotide and/or amino acid sequence information is significant, many patent applicants are accustomed to, or familiar with, the submission of such sequence information, in electronic form, to various sequence databases, such as GenBank, which is produced by the National Institutes of Health. In order to facilitate such submissions, or merely for the purpose of researching and developing sequence information, many eventual patent applicants also generate or encode sequence information in computer readable form. In order to further facilitate compliance with the sequence rules, the Office previously made available to the public an input program based on the AuthorIn program produced by GenBank. This input program, called PatentIn version 1.3, was

specifically tailored to the requirements of the sequence rules which were in effect between October 1, 1990 and July 1, 1998. Applications filed in the U.S. or in member countries of WIPO after July 1, 1998, containing sequence listings prepared using PatentIn version 1.3 will not be in compliance with 37 CFR 1.821-1.825 or ST.25.

The current sequence rules, which are embodied in 37 CFR 1.821-1.825 and World Intellectual Property Organization (WIPO) Standard ST.25, became effective July 1, 1998. Several versions of PatentIn have been released since 1998.

PatentIn version 3.5.1, available as of November 2010, and the companion User's Manual, are available on the Office website (**www.uspto.gov/patents/basics/types-patent-applications/utility-patent/checker/patentin**) for free downloading. PatentIn 3.5.1 operates in a Windows environment.

For help related to downloading the PatentIn software program or using PatentIn, contact the Patent Electronic Business Center at 866-217-9197 or via email at ebc@uspto.gov.

While use of the PatentIn program is not required for compliance with the sequence rules, its use is highly recommended, as Office experience has shown that submissions developed with PatentIn are far less likely to include errors than those developed without the program. The many automatic features of the PatentIn program also greatly ease the generation of a "Sequence Listing" when compared to generating them by hand in a word processing environment. This is especially true for a "Sequence Listing" that includes many sequences and/or sequences having great lengths.

Checker software that may be used to check a "Sequence Listing" for compliance with the requirements of 37 CFR 1.824 is available on the USPTO website at **(www.uspto.gov/patents/basics/types-patent-applications/utility-patent/checker/patentin**. The User Notes on the Checker website should be consulted for an explanation of errors that are not indicated, and content that is not verified, by the Checker software.

2431 Sample "Sequence Listing" under WIPO ST.25 [R-07.2022]

[Editor Note: This section is not applicable to 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). See MPEP §§ 2412-2419 for guidance on WIPO ST.26 requirements for applications filed on or after July 1, 2022.]

A sample "Sequence Listing" is included below.

SAMPLE "SEQUENCE LISTING"
(under 37 CFR 1.821-1.825)

<110> Smith, John
Smith, Jane

<120> Example of a Sequence Listing

<130> 01-00001

<140> US 16/999,999
<141> 2018-02-28

<150> EP 91000000
<151> 2017-12-31

<160> 3

<170> PatentIn version 3.5.1

<210> 1
<211> 403
<212> DNA
<213> Paramecium aurelia

<220>
<221> CDS
<222> (341)..(394)

<300>
<301> Doe, Richard
<302> Isolation and Characterization of a Gene Encoding a Protease from
Paramecium sp.
<303> Journal of Fictional Genes
<304> 1
<305> 4
<306> 1-7
<307> 2011-06-20

<400> 1
ctactctact ctactctcat ctactatctt ctttggatct ctgagtctgc ctgagtggta 60

ctcttgagtc ctggagatct ctcctctcac atgtgatcgt cgagactgac cgatagatcg 120

ctgactgact ctgagatagt cgagcccgta cgagacccgt cgagggtgac agagagtggg 180

cgcgtgcgcg cagagcgccg cgccggtgcg cgcgcgagtg cgcggtgggc cgcgcgaggg 240

```
ctttcgcggc agcggcggcg ctttccggcg cgcgcccgtc cgcccctaga cctgagaggt      300

cttctcttcc ctcctcttca ctagagaggt ctatatatac atg gtt tca atg ttc        355
                                            Met Val Ser Met Phe
                                            1               5

agc ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt gtt tgtttgctc          403
Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val
                10                  15


<210>  2
<211>  18
<212>  PRT
<213>  Paramecium aurelia

<220>
<221>  MOD_RES
<222>  (10)..(10)
<223>  METHYLATION

<400>  2

Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu
1               5                   10                  15


Phe Val



<210>  3
<211>  13
<212>  DNA
<213>  Artificial Sequence

<220>
<223>  synthetic primer

<400>  3
tgtttgtttg ctc                                                          13
```

2432 [Reserved]

2433 Box Sequence; Hand Delivery of a "Sequence Listing" and Computer Readable Forms [R-07.2022]

To facilitate administrative processing of all papers and read-only optical discs associated with sequence rule compliance, all computer readable forms, read-only optical discs, fees, and papers accompanying them filed in the Office should be marked "Box SEQUENCE."

Correspondence relating to the sequence rules may also be hand-delivered to the Customer Service Window. In cases of hand delivery to the Customer Service Window, the computer readable form should be placed in a protective mailer labeled with at least the application number, if available. The labeling requirements of 37 CFR 1.52(e)(4) must also be complied with. In no circumstances should additional or complimentary electronic copies be delivered to examiners or other Office personnel.

2434 Examination of Patent Applications Claiming Large Numbers of Nucleotide Sequences [R-07.2022]

Polynucleotide molecules defined by their nucleic acid sequence (hereinafter "nucleotide sequences") that encode different proteins are structurally distinct chemical compounds. These sequences are thus deemed to normally constitute independent and distinct inventions within the meaning of 35 U.S.C. 121. Absent evidence to the contrary, each such nucleotide sequence is presumed to represent an independent and distinct invention, subject to a restriction requirement pursuant to 35 U.S.C. 121 and 37 CFR 1.141 *et seq.*

All pending applications are subject to *Examination of Patent Applications Containing Nucleotide Sequences*, 1316 OG 123 (March 27, 2007). Note, however, that supplemental restriction requirements will not be advanced in applications that have already received an action on their merits in the absence of extenuating circumstances. For national applications filed under 35 U.S.C. 111(a), polynucleotide inventions will be considered for restriction, rejoinder, and examination practice in accordance with the standards set forth in MPEP Chapter 800. Claims to polynucleotide molecules will be considered for independence, relatedness, distinction and burden in the same manner as claims to any other type of molecule.

See MPEP § 1850 for treatment of claims containing nucleotide sequences that lack unity of invention in international applications filed under the Patent Cooperation Treaty (PCT) and national stage applications filed under 35 U.S.C. 371.

2435 Publishing of Patents and Patent Application Publications with a Lengthy "Sequence Listing" [R-07.2022]

Due to the high cost and limited usefulness of the printed paper or composed electronic image versions of nucleotide and/or amino acid sequences, if the "Sequence Listing" portion is lengthy (i.e., at least 600 Kb (about 300 typed pages)), it will no longer be printed with the paper and composed electronic image (page image) versions of patents and patent application publications. The "Sequence Listing" will only be published in electronic form and will be available on the USPTO sequence homepage (**https://seqdata.uspto.gov**) as an ASCII plain text file.

Copies of patents and patent application publications that include sequence listings are available for sale through the Patent and Trademark Copy Fulfillment Branch's Certified Copy Center storefront at **https://certifiedcopycenter.uspto.gov/**. However, these copies will not include a "Sequence Listing" if the "Sequence Listing" is not included in the composed electronic image (page image) version of the patent or patent application publication.

Copies of U.S. patents and U.S. patent application publications are not provided in paper to applicants and are not placed in the application file. See MPEP § 707.05(a). If an applicant desires an electronic copy of a "Sequence Listing", applicant may either download the "Sequence Listing" from the USPTO sequence homepage (**https://seqdata.uspto.gov/**) or specifically request and pay for the electronic copy through the Patent and Trademark Copy Fulfillment Branch of the Public Records Division. Both applicants and members of the general public can obtain an electronic copy of the "Sequence

Listing" through the Patent and Trademark Copy Fulfillment Branch for a separate fee as set forth in 37 CFR 1.19(b)(3). See MPEP § 1730 for contact information for the Patent and Trademark Copy Fulfillment Branch.

If the patent is mailed to applicant, the Office will include a copy of the patent on paper and a copy of the "Sequence Listing" on an electronic medium (e.g., read-only optical disc), if the "Sequence Listing" is not printed in the patent.

If the "Sequence Listing" is not included in the page images of a patent or patent application publication, a standardized statement will appear. Additionally, in the electronic text version of the patent or patent application publication, the statement will include an active hyperlink to a Web page containing the "Sequence Listing". The standardized statement for a patent will read, for example:

"SEQUENCE LISTING"

The patent contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form from the USPTO web site (**https://seqdata.uspto.gov**). An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

The standardized statement for a patent application publication will read, for example:

"SEQUENCE LISTING"

The patent application contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form from the USPTO web site (**https://seqdata.uspto.gov**). An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

Sequence data may also be accessed in a more readily searchable manner from the National Center for Biotechnology Information (NCBI) at **www.ncbi.nlm.nih.gov** or from a commercial vendor. The USPTO forwards a copy of the sequence data to NCBI when a patent including a "Sequence Listing" is granted, and when an application containing a sequence is published pursuant to 35 U.S.C. 122(b). If NCBI elects to include the sequence data in one of its databases, NCBI indexes the sequence data according to patent or patent application publication number. There is currently no fee for the public to use the NCBI site.