

§ 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 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 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 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 (1998), including Tables 1 through 6 in Appendix 2, herein incorporated by reference. (Hereinafter "WIPO Standard ST.25 (1998)"). This incorporation by reference was approved by the Director of the Federal Register in accordance with 5 U.S.C. 552(a) and 1 CFR part 51. Copies of WIPO Standard ST.25 (1998) may be obtained from the World Intellectual Property Organization; 34 chemin des Colombettes; 1211 Geneva 20 Switzerland. Copies may also be inspected at the National Archives and Records Administration (NARA). For information on the availability of this material at NARA, call 202-741-6030, or go to: http://www.archives.gov/federal_register/code_of_federal_regulations/ibr_locations.html. 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 WIPO Standard ST.25 (1998), Appendix 2, Table 1. Modifications, e.g., methylated bases, may be described as set forth in WIPO Standard ST.25 (1998), Appendix 2, Table 2, 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 WIPO Standard ST.25 (1998),

Appendix 2, Table 3. 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 WIPO Standard ST.25 (1998), Appendix 2, Table 3 with the modified positions; e.g., hydroxylations or glycosylations, being described as set forth in WIPO Standard ST.25 (1998), Appendix 2, Table 4, 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 WIPO Standard ST.25 (1998), Appendix 2, Table 3 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.

(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 which contain disclosures of nucleotide and/or amino acid sequences must contain, as a separate part of the disclosure, a paper or compact disc copy (see § 1.52(e)) disclosing the 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. This paper or compact disc copy is referred to elsewhere in this subpart as the "Sequence Listing." Each sequence disclosed must appear separately in the "Sequence Listing." Each sequence set forth in the "Sequence Listing" must be assigned a separate sequence identifier. The sequence identifiers must begin with 1 and increase sequentially by integers. If no sequence is present for a sequence identifier, the code "000" must be used in place of the sequence. The response for the numeric identifier <160> must include the total number of SEQ ID

§ 1.822

37 CFR Ch. I (7–1–12 Edition)

NOs, whether followed by a sequence or by the code “000.”

(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, 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.

(e) A copy of the “Sequence Listing” referred to in paragraph (c) of this section must also be submitted in computer readable form (CRF) in accordance with the requirements of § 1.824. The computer readable form must be a copy of the “Sequence Listing” and may not be retained as a part of the patent application file. If the computer readable form of a new application is to be identical with the computer readable form of another application of the applicant on file in the Office, reference may be made to the other application and computer readable form in lieu of filing a duplicate computer readable form in the new application if the computer readable form in the other application was compliant with all of the requirements of this subpart. The new application must be accompanied by a letter making such reference to the other application and computer readable form, both of which shall be completely identified. In the new application, applicant must also request the use of the compliant computer readable “Sequence Listing” that is already on file for the other application and must state that the paper or compact disc copy of the “Sequence Listing” in the new application is identical to the computer readable copy filed for the other application.

(f) In addition to the paper or compact disc copy required by paragraph (c) of this section and the computer readable form required by paragraph (e) of this section, a statement that the “Sequence Listing” content of the paper or compact disc copy and the computer readable copy are the same must be submitted with the computer readable form, e.g., a statement that “the sequence listing information recorded in computer readable form is

identical to the written (on paper or compact disc) sequence listing.”

(g) If any of the requirements of paragraphs (b) through (f) 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, 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 submission in reply to a requirement under this paragraph must be accompanied by a statement that the submission includes no new matter.

(h) If any of the requirements of paragraphs (b) through (f) of this section are not satisfied at the time of filing an international application under the Patent Cooperation Treaty (PCT), which application is to be searched by the United States International Searching Authority or examined by the United States International Preliminary Examining Authority, applicant will be sent a notice necessitating compliance with the requirements within a prescribed time period. Any submission in reply to a requirement under this paragraph must be accompanied by a statement that the submission does not include matter which goes beyond the disclosure in the international application as filed. If applicant fails to timely provide the required computer readable form, the United States International Searching Authority shall search only to the extent that a meaningful search can be performed without the computer readable form and the United States International Preliminary Examining Authority shall examine only to the extent that a meaningful examination can be performed without the computer readable form.

[63 FR 29634, June 1, 1998, as amended at 65 FR 54680, Sept. 8, 2000; 69 FR 18803, Apr. 9, 2004; 70 FR 10489, Mar. 4, 2005]

§ 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.