# Report on the preparation of a new WIPO standard on the presentation of nucleotide and amino acid sequence listings using eXtensible Markup Language (XML)

*Document prepared by the European Patent Office (EPO)*

## BACKGROUND

1. The Task Force on Sequence Listings was created by the Committee on WIPO Standards (CWS), at its first session (October 25 to 29, 2010), to deal with Task No. 44 (see paragraph 29 of document CWS/1/10):

“Prepare a recommendation on the presentation of nucleotide and amino acid sequence listings based on eXtensible Markup Language (XML) for adoption as a WIPO standard. The proposal of the new WIPO standard should be presented along with a report on the impact of the said standard on the current WIPO Standard ST.25, including the proposed necessary changes to Standard ST.25.”

1. The Task Force was also requested:

“To liaise with the appropriate PCT body with regard to the possible impact of such standard on Annex C to the Administrative Instructions under the PCT.”

1. The European Patent Office (EPO) was assigned the role of Task Force Leader and has since then held six rounds of discussions on WIPO's Wiki and submitted a final draft for public consultation. The principle of differentiating the technical aspects of ST.25 from Annex C (PCT Administrative Instructions) was agreed upon at the eighteenth session Meeting of International Authorities in February 2011 (see paragraphs 88 to 92 of document  PCT/MIA/18/16) and at the fourth session of the PCT Working Group in June 2011 (see paragraphs 180 to 188 of document PCT/WG/4/17).
2. On the basis of the comments received from the Task Force members, a final round of discussions was held in order to achieve a common agreement on the Standard requirements.

## PROGRESS REPORT

1. The Task Force started operating in February 2011 on the basis of drafts prepared by the EPO. Many Offices participated in the process and posted useful comments on WIPO's related Wiki.
2. In March 2012, the Task Force finalized a draft of the standard that could be used by the offices to consult their respective public. Several significant issues were raised by public comments and they have been tackled in cooperation with the database providers DDBJ, EBI, NCBI.
3. The sixth round of discussions was finished on September 2013 and the draft incorporating the improvements resulting from the public consultation and further discussions among the Task Force members and Database providers, was posted on WIPO’s wiki for a final review.
4. On the basis of the comments received from the Task Force members, a final round of discussions took place aimed at achieving a common agreement on the Standard requirements. On a provisional basis, the Task Force named this Standard ST.26. The main body and annexes proposed by the Task Force for consideration and approval of the CWS introduce the following improvements when compared to the current ST.25:
   1. All (PCT) procedural issues are transferred to the PCT Administrative Instructions: the new standard is focused on technical aspects, thus to enable an optimal presentation of the sequence listings (the biotech-related part) and the appropriate format of the submission (namely XML);
   2. The biotech-related part has been considerably improved to reflect modern industry standards, for example:
   * inclusion of modified nucleotides and amino acids not previously provided for (e.g. D-amino acids, PNA, morpholinos etc.) which have gained importance in industry and need to be electronically searchable;
   * clear instructions for gapped sequences and sequence variants;
   * clarification with regard to features and annotations;
   * consistency with latest public biological sequence repositories consortia requirements (INSDC and UniProt); and
   * the XML definition is self-contained and independent of ST.36 or ST.96.
   1. The syntax provided by the Document Type Definition (DTD) used in the ST.26 increases the data accuracy and enables automatic data quality control.
5. The Task Force will continue to work on the transition aspects in 2014 and 2015 with the objective to submit for the consideration and approval of the CWS, at its fifth session, the recommendations for the transition from ST.25 to ST.26.

## ROADMAP

1. A new round of discussions will continue after the CWS/4 session, focused on the preparation of the recommendations for the transition phase to be presented at the CWS session in 2015.

[Annex II follows]