



AMERICAN INTELLECTUAL PROPERTY LAW ASSOCIATION

2001 JEFFERSON DAVIS HIGHWAY ■ SUITE 203 ■ ARLINGTON, Virginia 22202

August 22, 2005

Mail Stop Comments - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Attention: Lisa J. Hobbs

Comments Regarding the Proposed Rule: "Acceptance, Processing, Use and Dissemination of Chemical and Three-Dimensional Biological Structural Data in Electronic Format"
70 Federal Register 35573 (June 21, 2005)

Dear Ms. Hobbs:

The American Intellectual Property Law Association (AIPLA) appreciates the opportunity to present its views on the notice of proposed rulemaking entitled "Acceptance, Processing, Use and Dissemination of Chemical and Three-Dimensional Biological Structural Data in Electronic Format" published in the Federal Register on June 21, 2005.

AIPLA is a national bar association whose more than 16,000 members are primarily lawyers in private and corporate practice, in government service, and in the academic community. AIPLA represents a wide and diverse spectrum of individuals, companies, and institutions involved directly or indirectly in the practice of patent, trademark, copyright, and unfair competition law, as well as other fields of law affecting intellectual property. Our members represent both owners and users of intellectual property.

This advance notice of proposed rule making seeks comment on the advisability of the United States Patent and Trademark Office (PTO) requiring submission of chemical and three-dimensional (3-D) biological structural data in electronic format. This advance notice also seeks comment on the format such submissions may take.

This advance notice considers three different scientific disciplines, each of which has its own issues and history:

- A. 3-D Biological Structural Data;
- B. 3-D Small Molecule Structural Data; and
- C. Chemical Structure Data.

As the differences between these three disciplines significantly outweigh their similarities, they should be discussed separately. In considering each of the above disciplines, the PTO must consider: the benefits of requiring submission of electronic data to the examination process; the incremental burden such submission would impose on applicants (and their respective clients); the availability of such data; and the value of such information to the

innovative community as a whole. The comments below are directed for the most part to the issue of requiring submission of 3-D biological structural data.

A. Electronic Submission of 3-D Biological Structural Data

A fundamental goal of the patent system is to encourage the development of technology by making the subject matter of patent applications available to the public. The enablement requirement of 35 U.S.C. §112, first paragraph, requires the specification be written, “in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same...” The current practice regarding applications claiming 3-D biological structures fails to advance this goal.

Under current practice it is most common for such structural data to be included as tables of text. In many instances these tables are not included in the body of the specification, but instead are appended as drawings pursuant to 35 U.S.C. § 113 and 37 C.F.R. § 1.81, *et seq.* Such an application will frequently contain several hundred sheets of drawings of these tables.

The evaluation of such an application by those skilled in the relevant art requires significant and costly work. The most common way of evaluating such a patent application entails the line-by-line transcription of the data in these tables into a spreadsheet which can then be used by the scientists.¹ It must be assumed that examiners from the PTO must also undergo some such labor-intensive process in order to properly examine such an application.

Most scientific journals require the submission of the 3-D biological structure data to an approved database before publication of the manuscript. The journal *Science*, for example, requires that protein structural information be submitted to the Worldwide Protein Data Bank².

The unintended consequence of the PTO not requiring electronic submission, in contrast to a similar requirement of most major journals, is that the filing of a patent application is a less effective means of disclosure than is a scientific publication. Such an application provides significantly less information (than would a submitter of a manuscript) in exchange for possible receiving much more (a period of exclusivity). As the rules are currently applied, a patent application purporting to claim a three-dimensional structure of a protein cannot be effectively examined nor can it be used by the scientific community as a foundation to “promote the Progress of Science and the useful Arts.”³

A desired feature of the patent system is that a person skilled in the art should be able to review all the patent literature and (with some assistance from a patent practitioner) be able to determine if their own invention infringes any granted patent. A consequence of the current rules is that there is no reasonable means by which an academician, or similarly

¹ Optical character recognition (OCR) technology has been attempted by some practitioners, but, to date, the error rate has been greater than manual entry.

² See, H.M. Berman, K. Henrick, H. Nakamura (2003): “Announcing the worldwide Protein Data Bank” *Nature Structural Biology* 10(12): 980 for an overview of the formation of this database. Its self-defined mission is, “to maintain a single Protein Data Bank Archive of macromolecular structural data that is freely and publicly available to the global community.” <http://www wwpdb.org/> (last visited August 11, 2005).

³ U.S. CONST. art. I, §8, cl. 8.

situated scientist, could determine the degree of similarity between their own crystal of a particular protein and that in a granted patent.

When considering whether to require the electronic submission of 3-D structural data, one consideration must be the burden on the applicants. For proteins, it is not possible to develop the 3-D structure without having generated the necessary data. Such applicants will necessarily possess any information that may be required for an electronic submission. Such information would most assuredly be present in electronic form, notwithstanding any requirement from the USPTO.

An additional factor promoting the electronic submission of 3-D data for proteins is the existence of a universally accepted, publicly available format. As is noted in the advance notice of proposed rules, the mmCIF format (with minor, mutually recognizable variations) is the one standard used by all practitioners in this field.

Another consideration is the state of the art as it applies to the patent literature. The claiming of particular three-dimensional structures of proteins is still a developing field. While there are examples of such patents granting in the United States and elsewhere, there have not been an extraordinary number of such patents. A field like this, in its nascency, is one ripe for the introduction of rules that will guide the development of this field of patents.

Specific questions:

A. Questions pertaining to the Creation of 3-D Structural Data Files

1. *What benefits do you foresee for the applicant if electronic filing is adopted? What disadvantages do you foresee?*

A primary advantage for the applicant of a patent claiming a protein by its 3-D characteristics is the assurance of a better examination than is currently received. The current system discourages the examination of a claimed 3-D structure of a protein against the prior art, as the submitted 3-D structure is not in electronic form. A granted patent that has some assurance of a proper examination is less likely to incur costly and disruptive litigation.

A disadvantage of providing 3-D structural information electronically is the creation of an additional bureaucratic step in the procurement process. While some applicants may see the dissemination of their structure in a readily usable format to be a disadvantage, most would appreciate that such openness is a prerequisite to obtaining a patent.

2. *What types of 3-D data would be best submitted electronically?*

Three-dimensional data for proteins claimed by their 3-D structure should be cited for the reasons cited above. At a minimum, the amount of 3-D data required should be that required for publication in scientific journals. This means that not only the coordinate data, but sufficient underlying data to permit the skilled practitioner to evaluate the model proposed by the applicants.

If a small molecule is claimed by its 3-D structure, there is justification in requiring the submission of this 3-D structure electronically. If, however, a small molecule is claimed by means other than its 3-D structure (e.g., solid state nuclear magnetic resonance spectrum, crystalline shape, X-ray diffraction pattern, melting point, etc.), the 3-D structure should not be required.

3. *Should electronic submission of 3-D data be mandatory, optional, or mandatory for some types (e.g., protein crystals) and optional for others (e.g., small organic crystals)?*

Three-dimensional structural information should be mandatory for protein crystals as there can be no meaningful search or examination conducted on them absent the information in a usable format. Such information is almost universally in the hands of the applicants at the time of filing the application and would, therefore, constitute little additional effort.

Three-dimensional structural information should not be mandatory for small molecule patents, as there are readily available means of searching and examining these applications in existence. Few small molecules are claimed by their 3-D structure, but even if that were to change, their smaller size and relative simplicity suggest that proper examination can still be accomplished without the need for imposing new rules. As most organic small molecule crystals are claimed not by their 3-D structure, but instead by a physicochemical property (e.g., X-ray diffraction pattern, solid state NMR spectrum, or melting point), the 3-D data may not be available to the applicant at the time of filing the application.

There would seem to be little benefit in developing a system where such submission would be optional. Those scientists wishing to share their structures with the general scientific community may avail themselves of the appropriate databases. There would seem to be little motivation to complicate a patent application by requiring it to contain information that is not required to assess patentability. There is also a significant question about whether it is fiscally responsible for the PTO to establish a system (i.e., optional submission of electronic data) that would provide only marginal benefits.

4. *If electronic submission is mandatory, should the USPTO require all 3-D information cited in application to be submitted in electronic format, including prior art, or only new data?*

Applicants should not be required to submit electronic data they do not possess. If an applicant details specific 3-D structural data in its application, there are reasonable grounds for requiring its submission in electronic format. If, however, the applicant merely cites a publicly available reference (to another patent, a scientific publication, or a database listing), the applicant should not be required to submit that data in electronic format.

5. *Have tables of 3-D data generally been created for some other purposes before preparation of a patent application, e.g., for publication in a scientific journal or submission to a database?*

Yes; however, the practice has been standardized.

6. *Have most of the 3-D tables been submitted to a database before inclusion in a patent application? If so, which one?*

Practices vary as to the timing of submissions. When submitted, proteins and nucleic acids are generally sent to the worldwide Protein Data Bank (wwPDB) while small molecule data is sent to the Cambridge Structural Database (CSD) which is maintained by Cambridge Crystallographic Data Centre (CCDC).

7. *Have most of the 3-D tables been published before inclusion in a patent application?*

There is no standard practice; however, most of the 3-D tables have not been published prior to inclusion in a patent application.

8. *Database providers require certain annotation data. Would any of the annotation data currently required by 3-D database providers be unknown or proprietary at the time of filing a patent application (e.g., method used for crystal creation)?*

On a case-by-case basis there may be certain fields that are not necessary to satisfy the requirements of 35 U.S.C. § 112 that may be proprietary. It is difficult to determine in the abstract what those fields might be.

Many of the annotated fields, however, should be completed in order to satisfy the requirements of section 112. The method of creating the crystal, for example, may be required to meet the enablement and/or best mode requirements.

As a series of coordinates is only an interpretation of the underlying data, it is necessary to understand how the applicants derived the interpretation. This can be done either through submission of the underlying data or by completion of certain annotation fields.

For example, if the applicants relied on correlation with a known structure of a related protein, it would be essential for the annotations to provide the methods used by the applicants in determining their molecular model. Other information that might be seen as essential to reproduce the experiments described in the application include: how the crystal was grown; the presence of co-factors; the particular gene employed; and the presence of non-naturally occurring amino acids.

9. *Database providers often establish a controlled vocabulary for annotation or feature description information. Would there be any problems created during patent application prosecution if the electronic file relied on dynamic controlled dictionaries or vocabularies, controlled and maintained by database providers, not the USPTO, for the description of features, etc. What would be the pros and cons if the USPTO were to incorporate by reference a public database controlled vocabulary into any adopted standard?*

The greatest concern about a dynamic controlled dictionary or vocabulary is its changing nature over time. The dynamic nature of such a dictionary increases the concern of a definition altering its meaning over the life of the patent. This could

lead to greater uncertainty as to the metes and bounds of the invention, as well as how it was enabled and described. The result of this could be increased litigation exposure for the scientific community as a whole.

An ideal system would be one in which the same mmCIF file would be acceptable for submission to both the scientific database as a prerequisite for publication and the USPTO as part of the application. Some of the particular additional fields necessary for USPTO submission (see answer to Question 10), however, may make that type of universality impractical.

10. *Is there annotation data specific to a patent application that does not appear in public database files but that would be desirable to provide for an electronic submission in a patent application (e.g., continuing application data, attorney's docket number)?*

Yes, an electronic submission of 3-D structural data should include sufficient identifying information to permit the attorney, the client, the PTO, and the scientific community to pair the electronic submission with the relevant patent application(s). Similarly, as a patent application is more likely than a scientific publication to contain more than one 3-D structure, there should be tags linking one data set to the relevant place in the application.

An excellent model for the type of identifying data necessary is the sequence listing rules found at 37 C.F.R. § 1.821, *et seq.* Sequence listings provide the necessary bibliographic information, as well as sequence identification numbers that tie a specific electronic dataset to a specific sequence in the specification.

B. Questions Pertaining to the USPTO Receipt of 3-D Files

1. *In general, 3-D structure data tables submitted as part of a patent application are quite lengthy. Should the USPTO require that all 3-D files greater than a certain size be submitted in electronic media only?*

There is minimal value to the scientific community as a whole in lengthy paper copies of structural data. It would seem to be prudent to require lengthy 3-D files be submitted in electronic form only.

One concern is the consequence of a resubmission of the electronic medium due to an error or problem with its reading. Such errors do arise with sequence listings, such as an unreadable disc or corrupt file, or a minor error like a transcriptional error in the identifying data (e.g., mistyped serial number). If the only means by which the 3-D data was sent is a corrupted file, the absence of the paper copy could give rise to an inference that the application has failed to meet the enablement and written description requirements of 35 U.S.C. § 112, and therefore, is not entitled to its filing date. Any proposed rules must provide a mechanism by which the applicants can repair such an error without loss of the filing date (*cf.* 37 C.F.R. § 1.825(d)).

2. *Should the USPTO require submission in electronic format at the time of filing, or, if a paper copy is filed, permit the electronic submission to be filed later (with a*

statement indicating that the electronic version is the same as the version originally filed)?

The PTO should permit applicants a limited period of time for the submission of the electronic format of the data, so long as the data was provided in paper format at the time of filing. The timing of such a filing can parallel the current practice done with sequence listings.

- 3. Should any statement that comes with an electronic file outline the authoring tool and certify the use of a validation tool?*

The statement accompanying the electronic file should provide sufficient guidance to ensure the file is read properly. It should include the authoring tool and should certify the use of a validation tool.

- 4. Should the rules be revised to specify that 3-D biological structural data, if a paper copy is provided, is to appear in a special section, e.g., between the specification and the Sequence Listing?*

There should be consistency in where the paper copy of the 3-D structural data is to be found in the application. The location between the specification and the sequence listing is reasonable.

Sincerely,



Michael K. Kirk
Executive Director
AIPLA