



**Address** 100 Barr Harbor Drive  
PO Box C700  
W. Conshohocken, PA  
19428-2959 | USA

**Phone** 610.832.9500  
**Fax** 610.832.9555  
**e-mail** [service@astm.org](mailto:service@astm.org)  
**Web** [www.astm.org](http://www.astm.org)

---

**Committee E13 on MOLECULAR SPECTROSCOPY AND CHROMATOGRAPHY**

**Minutes for E13.15 Subcommittee Working Group**

10:30 am – 12:30 pm EST

January 19, 2007

Virtual Meeting

- I. Introductions and Welcome:** Gary Kramer, E13.15 Chair called the meeting to order at 10:30 am EST. The ASTM release form for recording the session was displayed on the screen. There being no objections, the meeting was recorded.
- II. Attendees:**
- |                                   |                           |
|-----------------------------------|---------------------------|
| Mark Bean, GSK                    | Peter Linstrom, NIST      |
| Stuart Chalk, Univ. of N. Florida | Dave Martinsen, ACS       |
| Maren Fiege, Waters               | Dale O'Neill, Agilent     |
| Patrick Gleichmann, NIST          | Alex Roth, Fachhochschule |
| Gary Kramer, NIST                 | Wiesbaden                 |
| Robert Lancashire, UWI            |                           |
- III. Next Meeting:** February 9, 2007, 10:30 - 12:30 EST.
- The UV/Vis Technique Group will meet on February 6, 2007 at 10:30 am.
  - The Core Group will meet January 23, 2007 at 10:30 am EST. They generally meet every other Tuesday.
- IV. Minutes:** The minutes were delivered to late for reading. The approval will be delayed until the next meeting.
- V. Core Developers:**
- a. The core group has discussed several approaches to handling hyphenated data in AnIML. They plan to make a decision next meeting. A new idea was circulated by Mark Bean, and Mark agreed to send that document to all on the E13.15 virtual meeting list. However, the core group should make a recommendation prior to discussion by the main group.
- VI. Technique Committees:**
- a. **UV/Vis:** Progress is being made. The spectrum page specification has been completed, and the group is now discussing the time-dependent page. Some

issues have been raised on automated methods and sequences, for discussion by the main group.

## VII. Methods and Sequences:

- a. The group spent some time discussing the meaning of methods and sequences. The consensus was that an automated method is a set of commands to be run on a sample or group of samples. The group of samples can be thought of as a sequence. The method could be thought of consisting of an acquisition step, a separation step, and a quantitation step, or as a chromatography technique, a UV/Vis technique, and an MS technique. Postprocessing steps could also be a part of the method. Each time you iterate over a new sample in the sequence, the same group of methods might be run. At some time after the experiment is completed, the data could be re-analyzed with another set of parameters, or a different algorithm. This wouldn't invalidate the initial dataset, but would simply create a new processed dataset using the same raw data. It was pointed out that AnIML can already support this, through arbitrary concatenation of experiment steps. The question is whether or not there are standard methods which could be applied. For example, could standard peak finding or thresholding methods be created, or is this something that each vendor does in a unique way? Could we define some standard methods, and then let vendors extend those if necessary?
- b. Also noted was a difficulty with the sample container element. It is not part of sample, which could appear in many places. It could be defined as chemical entity, or entities, in a specific location. If you divide a sample and put it into three vials, do you still have one sample, or do you now have three samples? With a fiber optic probe, the sample might be difficult to define. The location of the container relative to the instrument is important. In an autosampler, for instance, the position could be significant. If a rotating autosampler doesn't stop to position the sides of a cuvette orthogonal to a beam, the results could be dependent on autosampler position.
- c. The question of whether there is one AnIML file per sample, or many samples per AnIML file was also considered. It was noted that AnIML can accommodate both, but this doesn't necessarily solve the sample question. One could envision a scenario where the parameter (x,y,z) could account for the sample location, with an integer sequence number to indicate every measurement done over the lifetime of the instrument. The question then is whether or not this should be part of the core, or should be defined in all of our technique definitions. We have already talked about the concept of modules, which would allow this construction to be "included" in techniques. Is there a need for a sequence technique, or an experiment step sequence? Adding more and more techniques will make things more complex for users. The module solution would allow the same sample module, with the location and sequence parameters to be included in all techniques. The alternative, to

create a separate sample introduction/injection technique, would remove the burden for each vendor to create such a technique. The core group needs to decide which approach is best.

## **VIII. Miscellaneous:**

### **a. Float Data Type:**

- i. Stuart Chalk circulated an email recommending that we report the float datatype in scientific notation. This would allow software processing AnIML documents to gather significant figures and decimal place information in addition to the numeric value. It would save us having to change the default float types to contain attributes to include these parameters.
- ii. It was noted that regulatory requirements specify that the original data be stored. To accommodate those needs, we can't reformat into scientific notation - it wouldn't be the original data anymore. If the instrument does some processing before reporting the data, the data that the instrument reports is still what must be recorded for regulatory compliance.
- iii. There was also some discussion about what is meant by significant figures and uncertainty, and whether or not the uncertainty was inherent in the instrument, or is determined experimentally by taking multiple measurements and computing the uncertainty from those measurements.

### **b. Default Values for Missing Legacy Data:**

- i. Another question related to specifying default values for required AnIML elements. Legacy data may not contain all of the AnIML required elements. The discussion here centered on whether or not a default value should be assigned, which might cause confusion as to whether it was a real measured value or not, or whether a "null" or "not available" value should be substituted for a real value. The group leaned toward the latter.
- ii. A suggestion was made that digital signatures, if allowed on all AnIML elements, could be used as references to document any default values used. There was concern that this would make the AnIML file much larger than it is already. Alternately, xpath specifications could be used to point to where default values were used. Another suggestion was that to add an optional attribute *unavailable-legacy="yes"* to solve the issue. Other options discussed were: to create a new data type of unavailable, which would add complexity; put in default values of 0.0 or 0; remove required elements which legacy data are not likely to have, which could lead to the JCAMP issues we are trying to resolve. The consensus was that required parameters should be

minimized, and that if a default value attribute is used, it should only apply to required elements. It was suggested that this would be an appropriate matter for the core group to discuss.

- c. **MD5 or other checksums** for ensuring data accuracy, need to be added to the AnIML specification, but how should this be done. It could be integrated into the XML file when saved. It was noted that digital signatures allow for MD5, but W3C no longer recommends MD5. Rather, they recommend SHA1. If a digital signature is wrapped around the entire file, the issue is addressed. AnIML is not inventing anything, just using existing XML standards.
- IX. AnIML Wiki:** The benefits of a wiki environment for AnIML documentation were discussed, and Mark Bean agreed to create a wiki environment, put Burkhard's existing documentation into the wiki, and then send the url and access information to the group. It was noted that YAML might be a good language to enable in the wiki. This would allow the wiki content to be extracted and turned into documentation.
- X. Adjourn:** 12:30 pm EST.

Minutes prepared by David Martinsen, ACS  
ASTM E13.15 Secretary