

Introduction

The Agency receives thousands of impurity structures per year. For example, Drug Master File (DMF) holders provide a hazard assessment for impurity classification which includes the use of (quantitative) structure-activity relationship ((Q)SAR) models.¹ Impurities consist of starting materials, intermediates, by-products, and degradants, and can reach numbers >15 for a single DMF. Currently, structures associated with these impurities are submitted through the eCTD as unmanipulable portable data files (PDF). Submission of structures in PDF format requires end users at the FDA to manually redraw each structure which is tedious, prone to errors, and an inefficient use of resources. In addition, further review of the impurities using cheminformatics tools—application of (Q)SAR models and registration in the FDA Global Substance Registration System (GSRS)—or use in the CMC review document, requires structures to be converted to a computer-readable format. This poster introduces the benefits that will result when industry submits structures in a single SD File that can then be used by multiple groups within the Agency for their various work purposes.

¹ICH M7: Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk

Please refer to the *Quick Guide to Creating a Structure-Data File (SD File) for DMF Submissions*³ for column definitions. The SD File should be located in Module 3 section 3.2.S.3.2 of the eCTD. {If you have a complex substance that requires additional data elements to describe you can obtain a UNII for the substance by contacting FDA-SRS@fda.hhs.gov} If you have any questions on how to format an SD File or trouble submitting it in the eCTD, please contact DMFOGD@fda.hhs.gov.

¹ SDfiles (multiple structures and optional data). CTFE FORMATS BIOVIA Databases 2016. Dassault Systèmes © 2015.

http://help.accelrys.com/ulm/one/1.0/content/ulm_pdfs/direct/reference/ctfileformats2016.pdf

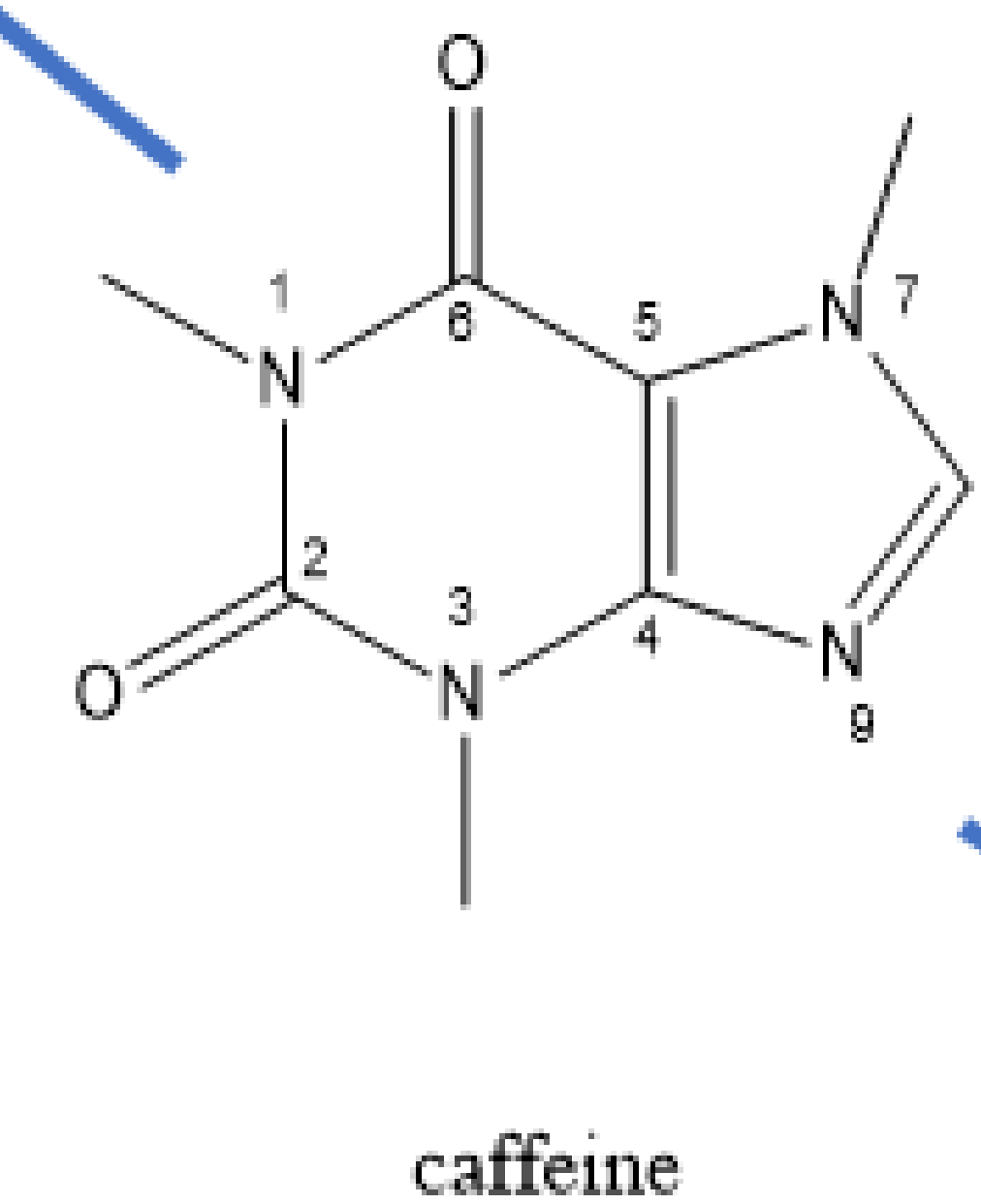
² <https://fdasis.nlm.nih.gov/srs/>

³ <https://www.fda.gov/drugs/forms-submission-requirements/drug-master-files-dmfs>

Who will use the SD File and what do they do with it?

FDA Global Substance Registration System (GSRS)

- Automatic notification that structures exist
- Find existing UNIs for structure matches
- Bulk load new structures for assigning new UNIs
- Flag structures that are invalid / ambiguous for review
- Link application records to the GSRS records for traceability
- Prioritize records for deeper curation adding:
 - ✓ Synonyms
 - ✓ CAS RN numbers
 - ✓ Relationships to other substances
 - ✓ References



Structure from SD File Submitted in eCTD by Industry

Division of Lifecycle API Assessors

- Extracts structure from SD File into chemdraw format
- Uses it in DMF review document
- Uses it to submit (Q)SAR consult via Panorama
- Uses it to submit safety consult

FDA Computational Toxicology Consult Service

- Standardize and register structures
- Modify SD File, if needed
- Run SD File in (Q)SAR software

What is an SD File: The 'Ask'

An SD File is an extensible, portable text file encoding one or more computer-readable chemical structures linked to associated data fields. An example SD File shown in Figure 1.

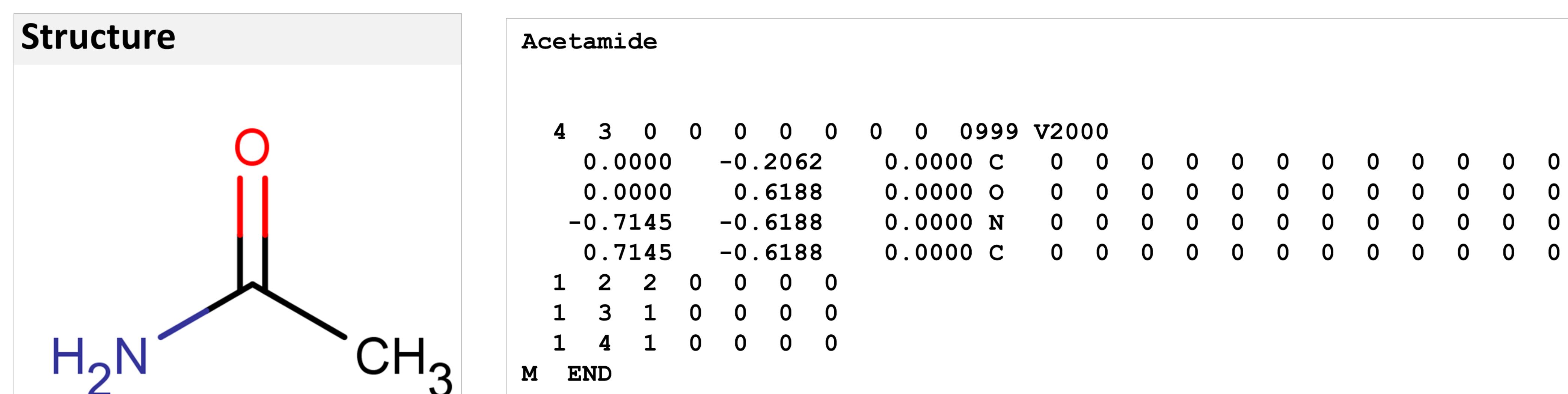


Figure 1: Structure of Acetamide and its corresponding computer-readable SD file format.

Benefits of the SD File for CMC Review

Benefits:

- The SD File can be opened from the eCTD and the machine-readable data interpreted as the corresponding structure(s), as shown in Figure 1. Any associated data fields will be automatically linked to the structures. The structures can then be used by the quality assessor for inclusion into the DMF review or for submission of a (Q)SAR or safety consult.
- Importantly, the responsibility for submission of the correct structure is placed on industry.
- Submission of the SD File will result in fewer typographical errors on the part of the Agency that could be propagated as multiple groups re-draw the structure for their own review and databasing purposes.
- Review efficiency will increase as our valuable assessor resources will not be used to re-draw structures.

Conclusions

The Agency receives thousands of impurity structures per year. Use of the SD File format eliminates redrawing of structures, increases review efficiency, and is compatible with internal cheminformatics tools.

Acknowledgements

Many thanks to Norman Schuff for his help in editing the SD File Quick Guide.