FDA U.S. FOOD & DRUG ADMINISTRATION

Disclaimer: The findings and conclusions in this poster reflect the views of the authors and should not be construed to represent FDA's views or policies. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

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

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.

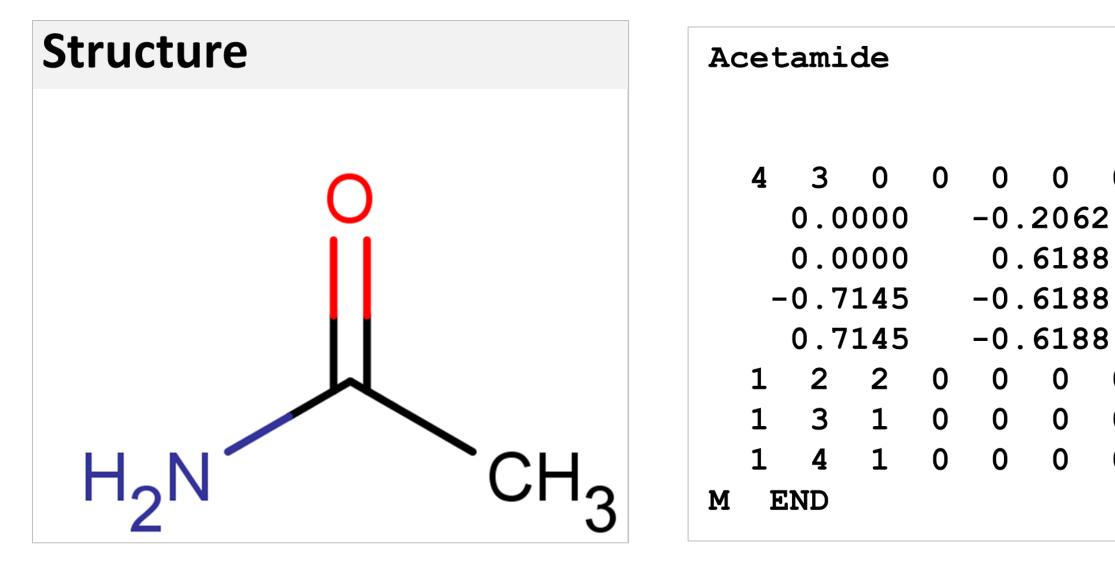


Figure 1: Structure of Acetamide and its corresponding computer-readable SD file format.

Under a collaboration between Office of Pharmaceutical Quality, Office of Translational Science and Office of Commissioner, DMF holders may now provide SD Files through the eCTD, with standardized associated data fields meeting Agency cheminformatics needs. To assist industry in creating the SD File, a Quick Guide has been drafted.² The Division of Lifecycle API (DLAPI) will recommend DMF holders provide impurity structures as a single SD File using the language below:

Please note, the Electronic Submissions Gateway (ESG) has been updated to accept chemical structures as structure-data files (SD File)¹. Please assist an efficient review of your DMF by providing all chemical structures (i.e. drug substance, starting materials, intermediates and impurities) in a single, comprehensive SD File. The SD file should include the structure of the drug substance, the structure of each other chemical, the name or abbreviation of each chemical as it appears in the DMF, DMF Number, and a unique identifier for cross-reference (e.g. Structure 1, Structure 2, etc.). If available, the following data items may also be included: UNII code², CASnumber, role of chemical (e.g., active ingredient, process impurity, degradant, metabolite, starting *material, intermediate).*

² Please see companion poster by Kim et al., titled "Structure-Data File from Sponsors' Submissions Promotes Rapid Chemical **Registration and Computational Analyses**"

Stop Re-drawing Chemicals! Implementation of Computer-Readable Chemical **Structure Format for Drug Impurities** Barbara Scott¹, David Green¹, Marlene Kim², Naomi Kruhlak², Tyler Peryea³, David Skanchy¹ ¹Office of Pharmaceutical Quality/Office of New Drugs/Division of Lifecycle API (DLAPI) ²Office of Translational Science/Office of Clinical Pharmacology/Division of Applied Regulatory Science (DARS) ³Office of the Commissioner/Office of Chief Scientist/Office of Health Informatics

0	0 0 0999		V2000													
2	0.00	00	С	0	0	0	0	0	0	0	0	0	0	0	0	
3	0.00	00	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	0.00	00	N	0	0	0	0	0	0	0	0	0	0	0	0	
3	0.00	00	С	0	0	0	0	0	0	0	0	0	0	0	0	
0																
0																
0																

Please refer to the Quick Guide to Creating a Structure-Data File (SD File) for DMF Submissions³ 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 <u>DMFOGD@fda.hhs.gov</u>. SDfiles (multiple structures and optional data). CTFILE FORMATS BIOVIA Databases 2016. Dassault Systèmes © 2015. http://help.accelrysonline.com/ulm/onelab/1.0/content/ulm_pdfs/direct/reference/ctfileformats2016.pdf ² https://fdasis.nlm.nih.gov/srs/ ³<u>https://www.fda.gov/drugs/forms-submission-requirements/drug-master-files-dmfs</u>

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.

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.

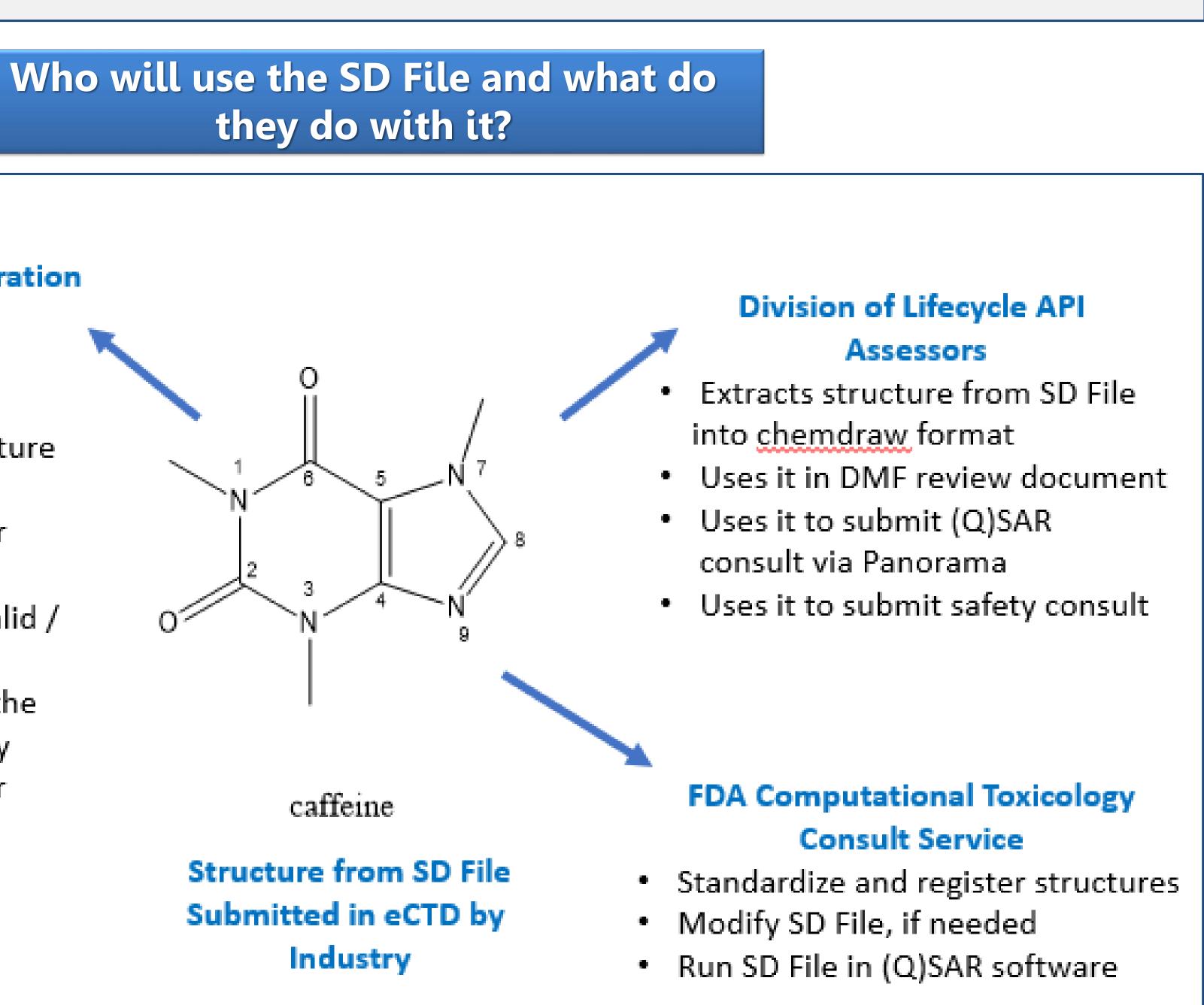
FDA Global Substance Registration System (GSRS)

- Automatic notification that
 - structures exist
- Find existing UNIIs for structure matches
- Bulk load new structures for
- assigning new UNIIs
- 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

Conclusions

Acknowledgements

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



Benefits of the SD File for CMC Review