

## Sharing chemical structures with peer-reviewed publications. Are we there yet?

Antony J. Williams

In the domain of chemistry one of the greatest benefits to publishing research is that data can be shared. Unfortunately, the vast majority of chemical structure data associated with scientific publications remain locked up in document form, primarily in PDF files or trapped on webpages. Despite the explosive growth of online chemical databases and the overall maturity of cheminformatics platforms, many barriers stifle the exchange of chemical structures via publications. These challenges include incomplete support by accepted standards (especially InChI) for advanced stereochemistry, organometallic compounds and generic “Markush” representations, the difference between human-readable and computer-readable forms of data, and challenges with the computer representation of chemical structures. To address these obstacles to chemical structure sharing, US EPA National Center for Computational Toxicology scientists are using a combination of cheminformatics applications and online repositories to distribute chemical structure data associated with their publications. This presentation will describe how EPA-NCCT chemical structure data that is amenable to indexing and distribution are shared and highlight the benefit of open data sharing for modeling, data integration, and increasing research impact. *This abstract does not reflect U.S. EPA policy.*