

[P24] On-Line Reaction Database Cartridge with Built-in Condensed Graph of Reaction-Based Search Engine

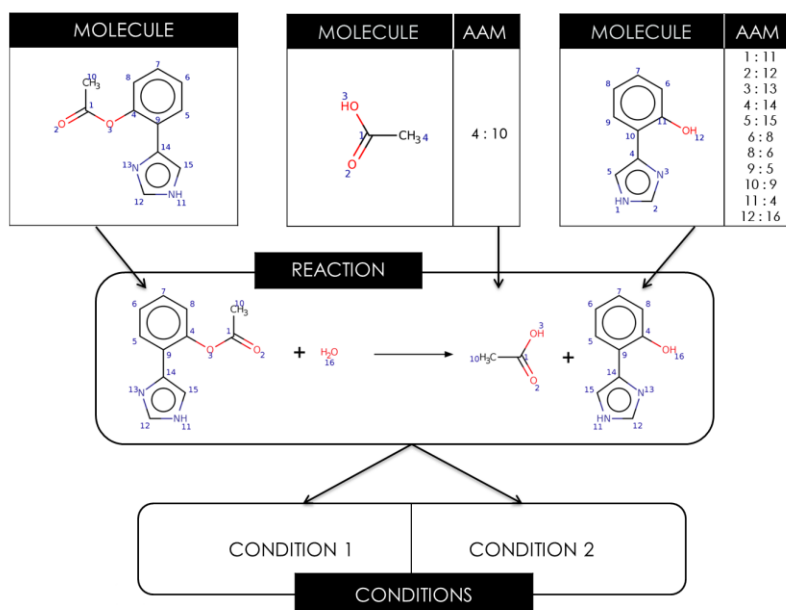
Ramil Nugmanov¹, Adelia Fatykhova¹, Timur Madzhidov¹, Alexandre Varnek^{1,2}

¹ Kazan Federal University, Kremlyovskaya st. 18, 420008, Kazan, Russia;

² University of Strasbourg, rue B. Pascal 1, 67000, Strasbourg, France.

Nowadays, a large volume of data on chemical reactions has accumulated among various research teams around the world. On the other hand, no known commercially available and convenient database cartridge to combine and store this data is designed so far. Thus, our goal was to develop an on-line database of chemical reactions which may provide with free access to information and search, a user-friendly interface for data registration and developers API interface for integration in own projects.

To store reactions, we use the representation of reactions as a combination of molecules with a molecule-to-reaction mapping. Storing links to individual molecules instead of storing each molecule for each reaction saves a huge amount of disk space. This approach allows user modifying structure of particular molecule only once, without any need to rewrite each record in different reactions. It also greatly simplifies the re-standardization. A popular among organic chemists "reaction-by-molecule" search does not require scanning all reactions or applying complex indexes to speed up the search. To simplify the data structure, some important properties of molecules and reactions are stored in separate table in the "json" format, which can be indexed if necessary.



Unique technology of Condensed Graph of Reaction (CGR) allowed us to implement substructure and similarity searches which use the same engine for molecules and reactions and to assess quantitatively reaction similarity. Thanks to special reaction fingerprints, substructure and similarity search for small and/or common fragments has become faster and more efficient.

In the framework of database management, a user-friendly mechanism was implemented to delete / add / change molecules or reactions, requiring minimal user actions and automated updating all dependencies in the database. Developed Application Programmer's Interface (API) on Python language provides an opportunity to integrate CGRdb database cartridge into their project. This may allow user to modify SQL-queries, to manage flexible search/storage settings and to make some other custom settings.

The research was supported by Russian Science Foundation, grant 14-43-00024.