Senior Project - Forge

A Virtual Chemical Library Design Workbench
Senior Project: 2001-2002

Almir Halilcevic, Grant Macklem, Arjune Mirchandani, Sean Murphy and Alan Schmidt
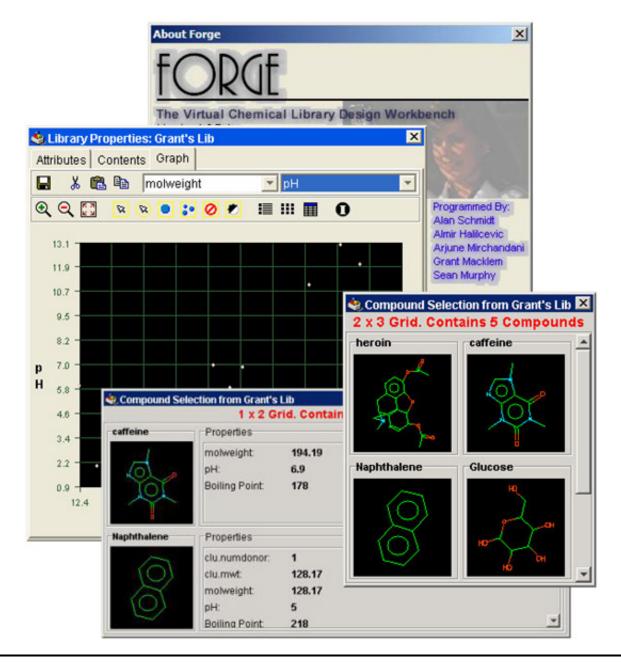
Array BioPharma
Boulder, CO

Array BioPharma is a drug discovery company creating new drugs through innovations in chemistry and structural biology. Array's world-class scientific team integrates chemistry and structural biology with an information-based technology platform to create higher quality drug candidates. Array provides drug discovery expertise to collaborators and for its own proprietary research.

To create new libraries of chemical compounds is an involved and time-consuming process. Though there are a virtually unlimited number of potential compounds, Array BioPharma can only manufacture about 100,000 to 200,000 compounds per year. Choosing which compounds should be manufactured in a library involves selecting compounds with desirable (predicted) physical properties and maximal inter-compound dissimilarity, while minimizing the cost of production.

Because this initial process is time consuming and exploratory, ranging anywhere from one to three months, the project was to develop a computer application that allows users to examine compound properties and store compounds into new libraries. Within newly created libraries, the compounds' properties can be computed (as derived from the compounds' structures), viewed and compared with other libraries, and then saved for later referencing.

Users access the system via a Forge client on their local machine. The software utilizes a networked Oracle database in which user data and profiles, as well as chemical compounds, can be saved and accessed. Most of the chemistry-related computation is done by Daylight, a package of libraries which contain query and manipulation routines for chemical structures. The project was implemented in Java and Perl.