Cando: A compiled programming language for setting up simulations and designing macromolecules

Christian Schafmeister
Temple University
Chemistry Department
Nov. 4 4:00 PM SERC 703

There is a great need for better software tools for chemistry. The programming languages that could be used to develop these tools are either fast, and difficult to write complex software within (compiled languages like Fortran, C, C++) or expressive, interactive and domain specific but about 100x slower (interpreted languages like Matlab, Python, R, Mathematica, Tcl etc.). I present "Cando", a compiled programming language for Chemistry that combines the speed of compiled languages with the expressiveness and interactivity of interpreted scripting languages. Cando is a superset of an old and highly respected programming language called Common Lisp. Cando is built upon the modern compiler backend LLVM which is the basis of Clang, a new C++/C compiler being developed by Apple, Google and Sony. Cando has a unique feature that allows it to easily interoperate with C++ libraries allowing modern scientific software libraries like openMM and openBabel to be integrated within Cando as facilities that extend the language. Cando contains a rich collection of built-in facilities to represent, build, manipulate and extract information from molecular systems. Cando has built in garbage collection to ease memory management and it extends garbage collection to all chemistry classes. Furthermore, Cando has many facilities that make it excellent for developing chemistry tools. It can automatically assign atom types and force field parameters for the Generalized Amber Force Field, it can identify and assign stereochemistry to atoms, it has SMARTs built in for cheminformatics and pattern recognition of molecular fragments. While there has been a great deal of development of fast code for molecular dynamics and quantum mechanics the software tools to design molecules, and set up large numbers of simulations and to analyze the results is lacking. It is this lack that Cando is designed to meet.

I will demonstrate this software in the context of designing functionalized spiroligomers, synthetic macromolecules that we have been developing within my group for the past 14 years to act as ligands for proteins, catalysts and molecular devices.