

SOMA - Environment for Integrating Molecular Modeling Workflows on a Multiplatform Computing Grid

Pekka Lehtovuori, Tapani Kinnunen and Tommi Nyrönen

The Finnish IT Center for Science CSC, P.O. Box 405, FI-02101 Espoo, Finland, <http://www.csc.fi>

Molecular modelers typically utilize combinations of computing tools. The traditional mechanism of creating and sharing workflows that integrate computing programs is a manual process, and as such is technologically inefficient and cumbersome.

To alleviate this challenge, we have developed the SOMA modeling environment at the Finnish IT Center for Science. The SOMA environment unites multiplatform UNIX/LINUX computing resources and third-party chemoinformatics software offering a possibility to construct workflows for calculating molecular structure and properties. The modeling environment consists of XML descriptions for scientific programs, the core workflow management program Grape, toolkit for parsing program input and output and an extranet interface. The program Grape and the developed XML descriptions of programs allow scientist to link molecular modeling software into highly sophisticated computational workflows without in-depth knowledge of programming or UNIX scripting. The SOMA environment collects the calculated data produced by the workflow and stores the computed information in the Chemical Markup Language (CML) format. The extranet interface is used for user authentication, building the program configurations and the workflow. The interface also includes versatile analysis tools for sorting, filtering, and visualizing the results stored in the CML/XML format.

Reference:

SOMA - Workflow for Small Molecule Property Calculations on a Multiplatform Computing Grid, Lehtovuori, P. T.; Nyronen, T. H. J. Chem. Inf. Model; (Article); 2006; ASAP Article; DOI: 10.1021/ci050388n