

Cyberbridges Project Proposal

“Parallel computations for macromolecular simulation”

Cassian D’Cunha, Ph.D. candidate in Chemistry & Biochemistry

David Chatfield, Assoc. Prof. of Chemistry & Biochemistry (Mr. D’Cunha’s advisor)

Computer simulations of biological macromolecules are widely used in the pharmaceutical industry to accelerate drug discover and in academic research for a host of problems from protein folding to enzyme mechanism. The sizes of the molecular systems and the time scales modeled are memory and compute time intensive. Computer clusters have become popular for such applications, either for running many simultaneous simulations to obtain good sampling statistics or for distributing a single, expensive calculation over many computers.

Our research is focused on an enzyme, chloroperoxidase¹ (CPO), of industrial interest as a potential biological chiral catalyst. CPO has been shown to catalyze a host of stereospecific oxidation reactions with an efficiency much greater than standard organic synthetic routes.² Our research is aimed at elucidating the reaction mechanism of CPO and at predicting useful site-specific mutations for fine-tuning the enzyme’s catalytic activity for particular substrates of interest. Prof. Xiaotang Wang, also at FIU, plans to synthesize and assay mutants forms of CPO we predict to be useful.

The calculations we need to perform are of two types, both of which require substantial computer power: quantum mechanical (QM) calculations for the mechanism work, and molecular mechanical (MM) calculations for the mutation work. Turbomole is a computer program particularly efficient for the type of QM calculations we need to do (density functional theory). We have a computer cluster with thirty dual-processor AMD nodes. We have obtained Turbomole³ and tested it in serial mode. However, it is highly desirable to run Turbomole in parallel mode to speed calculations, which in serial mode can take many hours or days. The length of the calculations limits the accuracy with which we can model the biochemical systems. Turbomole is parallel-capable, but we have not yet been successful in running Turbomole in parallel on the computer cluster.

The computer cluster came with the operating system Scyld (based on RedHat Linux 6.0), which is simple to use but of limited flexibility. For example, we have not been able to run Gaussian, the most popular program for QM calculations involving molecules, on the cluster. This is probably due to an incompatibility with the Scyld operating system. Consequently, we wish to reconfigure the cluster and discard Scyld, both for parallelizing Turbomole and to enable us to run other computational chemistry programs.

Thus the proposed project has two specific aims:

Specific Aim 1: Reconfigure the computer cluster, using Linux, and jettison the Scyld operating system.

Specific Aim 2: With the reconfigured cluster, implement, test, and benchmark Turbomole in parallel mode. Test calculations already performed on CPO using Turbomole in serial mode will provide a basis for comparison.

These tasks will have an impact on our research capability beyond the use of Turbomole (and Gaussian). The programs CHARMM, for MM and hybrid QM/MM calculations; and DOCK, for ligand binding to predict useful mutations, will also be used in the CPO research. We anticipate that reconfiguring the cluster will help us to use it more efficiently for all of these calculations.

References

- ¹M Sundaramoorthy, J Turner, TL Poulos. *Chemistry & Biology*. Vol. 5, 1998, 461-73.
- ²Hong-In Lee, Annette F. Dexter, Yang –Cheng Fann, Frederick J. Lakner, Lowell P. Hager, and Brian M. Hoffman. *J. Am. Chem. Soc.* **1997**, 119, 4059-4069
- ³COSMOlogic GmbH & Co. HG. Turbomole 5.6. Burscheider Str. 515 D-51381 Leverkusen. Germany.