

G-MolKnot: A grid enabled systematic algorithm to produce open molecular knots

F.E. Psomopoulos^{1,2}, P. Mitkas^{1,2}, C.S. Krinas³ and I.N. Demetropoulos⁴

¹ *Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54 124, Thessaloniki, Greece*

² *Intelligent Systems and Software Engineering Laboratory, Informatics and Telematics Institute/CERTH, 57 001, Thessaloniki, Greece*

³ *Department of Chemistry, Section of Physical Chemistry, University of Ioannina, Panepistimioupoli Dourouti, GR 45 110 Ioannina, Greece*

⁴ *Department of Information and Telecommunications Engineering, Section of Applied Informatics, University of West Macedonia, Park Agiou Dimitriou, Kozani, GR 50 100, Greece*

Emails: fpsom@danae.ee.auth.gr, mitkas@eng.auth.gr, me00599@cc.uoi.gr, idimitr@uowm.gr

Molecular intertwining entities are increasingly found in nature as knotted proteins and DNA while synthesized molecular knots along with catenanes and rotaxanes have been proposed as parts of potential molecular machines, entropy-driven molecular shuttles or scaffolds for drug carriers. The General Molecular Knotting algorithm (GMK) [1, 2] is capable to produce open knotted alkanes, polyethylene oxides and peptides with high or low steric energy at will. The method is quite efficient for small chain lengths, where all other methods fail. The results are validated by applying high level ab initio calculations for selected molecular knots.

Although the GMK algorithm is by design a sequential process, it can be decomposed to four distinct domains; molecule generation, knot evaluation, molecular mechanics optimization and quantum mechanics optimization. Moreover, each conformation generated during the first phase is processed independently, thus allowing multiple runs of the algorithm simultaneously. These aspects of the algorithm define it as embarrassingly parallel and specify the Grid as the ideal platform for execution.

The deployment of the GMK over a Grid Environment [3, 4], namely the EGEE grid [5], offers two major benefits: reduction of the time required for each run and transparency for the end user. The first can be achieved by creating an MPI-wrapper for the GMK algorithm which will enable the use of multiple CPUs simultaneously, allowing at the same time a scalability regarding the number of processors to be utilized. By registering the application as a free accessed resource on the Grid (by means of a guid) and providing a “default” job description file, the transparency for the end user can be assured together with complete independence from operational systems or libraries installation.

We believe that topological molecules are expected to gain high importance in the coming years so the design of these compounds should be accelerated; the grid version of the MolKnot program will facilitate these investigations.

References

1. C.S. Krinas and I.N.Demetropoulos, "A systematic algorithm capable to yield open molecular knots: Application to alkanes, polyethylene oxides and peptides", Chem. Phys. Lett. 433, (2007) p 422
2. C.S. Krinas and I.N.Demetropoulos, "C₂₂H₄₆: The Smallest Open 31 Knotted Alkane by Computer-Aided Design", J. Mol. Graph. Model, accepted 16 Oct 2007 (DOI 10.1016/j.jm gm.2007.10.010)
3. H. □ Polychroniadou, F. □ Psomopoulos and P. A. Mitkas: "G-Class: A Divide and Conquer Application for Grid Protein Classification", Proceedings of the 2nd ADMKD 2006: Workshop on Data Mining and Knowledge Discovery (in conjunction with ADBIS 2006: The 10th East-European Conference on Advances in Databases and Information Systems), 3-7 September 2006, Thessaloniki, Greece, pp 121-132.
4. C. Gkekas, F. Psomopoulos, P.A. Mitkas, "Modeling Gene Ontology Terms using Finite State Automata", Hellenic Bioinformatics and Medical Informatics Meeting, 4-5 October, 2007, Biomedical Research Foundation, Academy of Athens, Greece.
5. Enabling Grids for E-scienceE (EGEE), <http://www.eu-egee.org/>