

## A Computationally Efficient Algorithm for Accurate Local Energy Minimization of Crystal Structures Containing Flexible Molecules

A.V. Kazantsev, P.G. Karamertzanis, C.S. Adjiman \*, C.C. Pantelides

Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London SW7 2AZ, United Kingdom.

\* poster presenter

Crystal structure prediction of organic molecules from their atomic connectivity diagram has been a key problem in computational chemistry for over a decade. Following recent advances, there exist reliable search methodologies that manage to identify most low-energy minima (including the experimentally determined polymorphs) for crystals of both rigid [1] and flexible [2] molecules. However, a major problem with these techniques is that, in order to be able to search efficiently what is a very large space of possible crystal structures, they have to rely on approximate estimates for the inter-molecular and intra-molecular energy contributions:

- Inter-molecular electrostatic contributions are represented using atomic (and sometimes other) charges derived from quantum mechanical calculations of the molecule's electrostatic field. Albeit computationally efficient, such representations are not as accurate as those based on multipole moments.
- In the case of flexible molecules, intra-molecular energy is computed by interpolating the results of quantum mechanical calculations over pre-computed grids of the molecule's flexible torsion angles [2]. The atomic charges may also be expressed as functions of these torsion angles in a similar fashion. Although the interpolants themselves are quite accurate, the distinction between "flexible" and "rigid" degrees of freedom in the molecule introduces a degree of approximation which miscalculates the actual positions of the atoms in the molecule within the crystal lattice, and therefore the corresponding inter-molecular interactions.

Because of the above reasons, the absolute energy and, consequently, the ranking of low-energy crystal structures identified by state-of-the-art search methods are not reliable. They, therefore, have to be corrected by re-minimization of the crystal lattice energy using more accurate models. One of the most advanced algorithms available for this purpose is DMAFlex [3]. This incorporates a two-level optimization procedure:

- The outer optimization manipulates the values of the flexible degrees of freedom (such as torsion angles, bond angles, etc.) and performs a quantum mechanical calculation to determine the minimum energy molecular conformation corresponding to the current set of values.
- The inner optimization identifies the crystal structure that minimizes the inter-molecular energy based on the molecular conformation currently fixed by the outer optimization. The electrostatic contributions to the inter-molecular energy are derived via an anisotropic distributed multipole model (up to hexadecapole level) computed [4] from the results of the quantum mechanical calculation at the outer level.

The above approach is quite accurate as demonstrated by its applications to several systems of interest [5]. However, the incorporation of a complete quantum mechanical calculation within an optimization loop results in high computational cost. Moreover, the derivation of a new set of multipoles after each conformational change calculation is also computationally demanding. The use of a gradient-free (simplex) algorithm for the outer optimization, which can overcome small discontinuities introduced by the multipole expansion of the isolated molecule charge density, limits the extent of molecular flexibility that can be handled with a reasonable number of outer iterations.

This paper presents a novel energy minimization algorithm that is significantly more computationally efficient than the one presented above, whilst being of equivalent accuracy. The algorithm employs an identical two-level optimization scheme but without the need for a quantum mechanical calculation at every outer iteration:

- Intra-molecular energies are estimated using a local approximate model based on a quadratic Taylor expansion of intra-molecular energy in terms of the flexible molecular degrees of freedom. The expansion is constructed around a base point derived from the intra-molecular energy and its first and second derivatives computed via an isolated-molecule wavefunction calculation.
- The above wavefunction calculation is also used to compute a distributed multipole expansion of the isolated-molecule charge density, where moments up to hexadecapole level are expressed with respect to the local axis system of each atom. The multipoles are then used to model the conformationally dependent intermolecular electrostatic interactions by assuming that they remain invariant with respect to the local axis system.
- The above define local approximate models which are valid within a certain distance of the point at which they were derived. The frequency with which these local models are updated is automatically and independently adjusted by monitoring their molecule-dependent sensitivity to molecular conformation as the minimization progresses.
- A quasi-Newton minimization algorithm is used for both the inner and outer optimization problems. This ensures rapid convergence even when there are many flexible degrees of freedom.

The validity of the proposed methodology has been tested by minimising an extensive set of experimentally determined and computer generated crystal structures containing flexible molecules. The results obtained are essentially identical (within numerical error) with those of the earlier scheme, but obtained at a fraction of the computational cost. In practice, it has been found that many optimizations can be done using a single quantum mechanical calculation carried out at the initial point.

As a result, the proposed methodology addresses the limitations identified in previous algorithms. It also opens some interesting possibilities for future work, e.g. in terms of removing the distinction between “flexible” and “rigid” degrees of freedom, and also, potentially, incorporating these more rigorous energy calculations directly within the global search techniques.

#### References:

- [1] Karamertzanis, P. G. and Pantelides, C. C. *J. Comput. Chem.* 2005, 26, 304-324.
- [2] Karamertzanis, P. G. and Pantelides, C. C. *Mol. Phys.* 2007, 105, 273-291.
- [3] Karamertzanis, P. G. and Price S. L.; *J. Chem. Theory Comput.* 2006, 2, 1184.
- [4] Stone, A. J. *J. Chem. Theory Comp.* 2005, 1, 1128-1132.
- [5] Karamertzanis, P. G. et al.; *J. Phys. Chem.* 2007, B, Vol. 111, 19, 5326-5336