

Contents list available at JMCS

Journal of Mathematics and Computer Science

Journal Homepage: www.tjmcs.com



Genetic Algorithms for Structure Prediction of New Bithiazole Molecular Crystals: Methodology and Applications

AkramHossenian^{1,*}

¹Department of Engineering Science, College of Engineering, University of Tehran, P.O. Box 11365-45
Tehran, Islamic Republic of Iran
E-mail: Hoseinian@ut.ac.ir

Amin Ghodousian²

²Department of Engineering Science, College of Engineering, University of Tehran, P.O. Box 11365-4563
Tehran, Islamic Republic of Iran
E-mail: A.ghodousian@ut.ac.ir

Article history:

Received July 2014

Accepted September 2014

Available online September 2014

Abstract

This article describes the application of our distributed computing framework for crystal structure prediction (CSP), the modified genetic algorithms for crystal and cluster prediction (MGAC), to predict the crystal structure of bithiazole molecules. Using a distributed parallel genetic algorithm and local energy minimization of the structures followed by the classifying, sorting, and archiving of the most relevant structures. A genetic algorithm has been used to generate plausible crystal structures from the knowledge of only the unit cell dimensions and constituent elements. This strategy increases the efficiency of the DFT based GA by several orders of magnitude. This gain allows considerable increase in size and complexity of systems that can be studied by first principles. The Gaussian 03 package is used to perform the calculation of these atomic charges at the optimized geometry (HF/6-31G*level). Our results indicate that the method can consistently find the experimentally known crystal structures of bithiazole molecules. The structural of computational parameters are in agreement with the experimental data.

Keywords: Genetic algorithms, Crystal structure prediction, DFT, framework, bithiazole.

1. Introduction

Structure is the most important piece of information about a material, as it determines most of its physical properties. It was long believed that crystal structures are fundamentally unpredictable. Crystal structure prediction starting from the chemical composition alone has been one of the long standing challenges in theoretical solid state physics, chemistry, and materials science [1, 2]. Progress in this area has become a pressing issue in the age of computational materials discovery and design. In the past two decades several computational methods have been proposed to tackle this problem. As the stable structure corresponds to the global minimum of the free energy, several global optimization algorithms [3,4] have been devised and used with some success for crystal structure prediction (CSP) for instance, simulated annealing [5, 6], meta dynamics [7], evolutionary algorithms [8, 9], random sampling [10], basin hopping [11], minima hopping [12], data mining [13], particle swarm optimization [14, 15] and random structure search [16].

For inorganic crystals, in many cases it is already now possible to predict the stable structure at arbitrary external pressure. Towards the ambition of designing novel materials prior to their synthesis in the laboratory, reliable and efficient prediction of the structure of more complex (in particular, molecular) crystals becomes imperative.

Molecular crystals are extremely interesting because of their applications as pharmaceuticals, pigments, explosives, and metal-organic frameworks [17, 18]. The periodically conducted blind tests of organic crystal structure prediction. The tests show that it is now possible to predict the packing of a small number of molecules.

In these cases, efficiency of search for the global minimum on the energy landscape is not crucial. However, if one has to use expensive ab initio total energy calculations or study systems with a large number of degrees of freedom (many molecules, especially if they have conformational flexibility), the number of possible structures becomes astronomically large and efficient search techniques become critically important. In addition, the nature of weak chemical interactions makes it common that molecules have a wide variety of ways of packing with lattice energies within a few kJ/molecule of the most stable structure. Thus prediction of such large structures is certainly a challenge, especially if the number of trial structures has to be kept low to enable practical ab initio structure predictions.

While there has been steady progress in predicting crystal structures of elementary crystals, oxides, and binary alloys [10-15, 17-18], exploration of complex binary, ternary, and quaternary systems has required more advanced algorithms for configuration space exploration and faster but reliable methods for energy evaluation.

2. Computational Methods

Computer simulation techniques are now acquiring a truly predictive capability in structural chemistry. This is a significant development since it is not always possible to grow a large enough crystal to determine the crystallographic coordinates and properties of a new compound from its X-ray diffraction pattern alone.

A new methodology for the prediction of molecular crystal structures using only the atomic connectivity of the molecule under consideration is presented. The approach is based on the global minimization of the lattice enthalpy of the crystal. The modeling of the electrostatic interactions is accomplished through a set of distributed charges that are optimally and automatically selected and positioned based on results of quantum mechanical calculations.

The prediction of the structure of crystals formed by organic molecules is a problem of immense practical significance to many key sectors of the process industries. This is particularly so in view of the propensity of many such compounds to crystallize in multiple stable forms and the influence of crystal structure on key product characteristics such as the density, color, solubility, rate of dissolution, melting point, chemical stability and optical properties.

A four-step global optimization algorithm is used for the identification of the local minima of the lattice enthalpy surface.

A parallelized implementation of the algorithm permits a much more extensive search of the solution space than has hitherto been possible, allowing the identification of crystal structures in less frequently occurring space groups and with more than one molecule in the asymmetric unit. The potential energy function ($U(R)$) is shown below:

$$\begin{aligned}
 U(R) = & \sum_{\text{bonds}} K_r (r - r_{eq})^2 && \text{bonds} \\
 & + \sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2 && \text{angles} \\
 & + \sum_{\text{dihedrals}} \frac{V_n}{2} (1 + \cos[n\phi - \nu]) && \text{dihedrals} \\
 & + \sum_{i < j}^{\text{atoms}} \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} && \text{van der Waals} \\
 & + \sum_{i < j}^{\text{atoms}} \frac{q_i q_j}{\epsilon R_{ij}} && \text{electrostatic}
 \end{aligned}$$

Where r_{eq} and θ_{eq} are equilibrium structural parameters. K_r , K_θ and V_n are force constants, n is multiplicity, and c the phase angle for the torsional angle parameters. In addition, A , B , and q are parameters related to the nonbonded potentials. For the non bonded part, the electrostatic parameters (q_i , q_j) are calibrated using the restrained electrostatic potential fit (RESP) model.

While first-principles density functional theory (DFT) calculations offer accurate total energies, its computational cost imposes the bottleneck to the structure identification of complex materials with unit cells containing ~ 102 atoms and/or with variable stoichiometries. By contrast, calculations based on classical potentials are fast and applicable to very large systems but are limited in accuracy. For various systems, reliable classical potentials are not even available.

We present in this letter an adaptive-genetic algorithm [19-21] that combines the speed of classical potential searches and the accuracy of first-principles DFT calculations. It allows us to investigate crystal structures previously intractable by such methods with current computer capabilities. The flowchart of the adaptive-GA scheme is illustrated in figure 1.

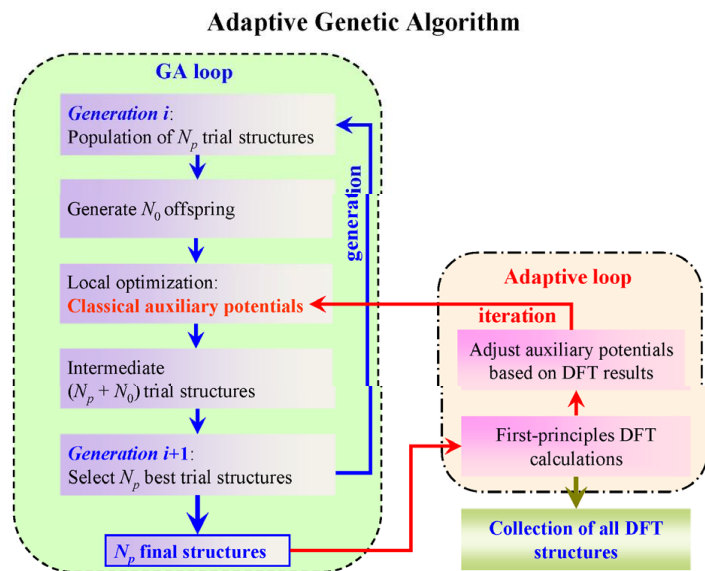


Figure.1. Flowchart of the adaptive genetic algorithm. The regular GA-loop is embedded in an adaptive loop. Optimization of offspring structures in the GA loop are performed using auxiliary classical potentials whose parameters are adjusted to reproduce DFT results obtained only in the adaptive loop.

Bithiazoles and their derivatives are of great interest because of their potential applications such as ligands for metals [22,23] and biological activities of their complexes, which show antibacterial, antifungal, anti-inflammatory, and antitumor activities [24-27]. They are also useful as electrochromic materials, photoluminescence, and semiconductor devices [28-34]. Geometry optimization of bithiazole molecules were performed using the B3LYP hybrid density functional method of theory with the STO-3g, 3-21G and 6-311G** basis sets. All the electronic structure calculations were carried out using the GAUSSIAN 98 program packages. The DFT calculations can give a remarkably good description of the molecular geometry.

5. Conclusion

Our study shows that it is possible to find good matches between predicted and experimental crystal structures of bithiazole molecules using a standard force field. The results suggest that increasing the effectiveness of the GA search has the undesirable consequence of bringing more physically non-plausible structures into the population. In this sense the GA is performing extraordinarily well in finding very unusual structures that according to the merit function, the crystal energy, are highly competitive. Clearly a single merit function is not sufficient because the energy given by the empirical potential function selects physically non-plausible structures. We believe that to correct this problem it will be advantageous to use a multi-objective approach to the GA optimization. A careful analysis of the structures shows that they have unit cell volumes much larger than those from the experimental crystals; this finding makes the unit cell volume a clear choice for a second optimization target. This is quite appealing because even in absence of experimental data the unit cell volume can be estimated using the empirical rules.

Acknowledgments

The authors would like to acknowledge the financial support University of Tehran for this research under grant number 28666/1/01.

References

- [1] Maddox, Crystals from First Principles Nature 335 (1988) 201-208.
- [2] A. Gavezzotti, Accounts of Chemical Research 27 (1994) 309-321.
- [3] M. Eslami, J. Vahidi, M. Askarzadeh, Journal of mathematics and computer Science 11 (2014) 291 – 299.
- [4] S. A. Moezi, E. Zakeri, Y. Bazargan-Lari, A. Zare, Journal of mathematics and computer science 11 (2014) 209-217.
- [5] J. Pannetier, J. Bassas-Alsina, J. Rodriguez-Carvajal, and V. Calgnaert, Nature 346 (1990) 343.
- [6] J. C. Schon and M. Jansen, Angew. Chem. Int. Ed. Engl. 35 (1996) 1286-1294.
- [7] R. Martoňák, A. Laio, and M. Parrinello, Phys. Rev. Lett. 90 (2003) 755-763.
- [8] A. R. Oganov and C. W. Glass, J. Chem. Phys. 124 (2006) 244-254.
- [9] A. R. Oganov, A. O. Lyakhov, and M. Valle, Acc. Chem. Res. 44 (2011) 227-231.
- [10] C. M. Freeman, J. M. Newsam, S. M. Levine, and C. R. A. Catlow, J. Mater.Chem. 3(1993) 531-550.
- [11] D. J. Wales and J. P. K. Doye, J. Phys. Chem. A 101 (1997) 5111-5119.
- [12] S. Goedecker, J. Chem. Phys. 120 (2004) 99-110.
- [13] S. Curtarolo et al., Phys. Rev. Lett. 91 (2003) 135-143.
- [14] Y.Wang, J.Lv, L.Zhu and Y.Ma, Crystal Structure Prediction via Particle-Swarm Optimization Phys. Rev. B 82 094116 (2010).
- [15] Y.Wang, J.Lv, L.Zhu and Y.Ma CALYPSO: A Method for Crystal Structure Prediction Comp.Phys. Commun. 183 2063-2070 (2012).
- [16] C. J Pickard and R. J. Needs Ab Initio Random Structure Searching J. Phys. Condens. Matter 23 053201 (2011).
- [17] S. L. Price, Adv. Drug Del. Rev. 56(2004) 301.
- [18] I. A. Baburin, S. Leoni, and G. Seifert, The Journal of Physical Chemistry B 112 (2008) 9437-9466.
- [19] M.Vatani, M.Asghari, G. Vakili-Nezhaad, *TJMCS*. 5 (2012) 60-66.
- [20] M.Moghaddas, M.R.Dastranj, N.Changizi, M.Rouhani, *TJMCS*.2 (2011) 572-579.

- [21] G. H. Mohebpour, A. Ghorbannia Delavar, *Journal of mathematics and computer science* 12 (2014) 282-294
- [22] A. Hosseinian, A. R. Mahjoub, *J. Mol. Struct.* 985 (2011) 270-287.
- [23] A. Hosseinian, S. Jabbari, H. R. Rahimipour, A. R. Mahjoub, *J. Mol. Struct.* 1028 (2012) 215-221.
- [24] Y. Aogagi, H. Suguaa, N. Murugesu, G. M. Ehernfeld, L.-H. Chang, T. Ohgi, M. S. Sherkhani, M. P. KieKup, S. M. Hecht, *J Am Chem Soc* 104 (1982) 5237-5278.
- [25] J. M. Riodan, T. T. Sakai, *Biochemistry* 21 (1981) 805-814.
- [26] S. Hideaki, S. Atsumi, N. Yasuyuki, *Chem. Pharm. Bull.* 45(1)(1997) 189-201.
- [27] T. Yan-Ni, Y. Pin, L. Qing-Shan, Z. Chun-Gui, *Chin. J. Chem* 14(5)(1996) 428-444.
- [28] Y. Ryuichi, M. Hitoshi, S. Hajimi, *Chem. Abstr.* 18184 (1997) 127-137.
- [29] J. K. Politis, M. D. Curtis, *Polym. Prepr.* 38(2) (1997) 379-393.
- [30] Y. Takakazu, M. Tsukasa, S. Hajime, A. Minoru, K. Dharma, S. Shintaro, *Chem. Lett.*, 2 (1997) 139-150.
- [31] Gonzalez, R. L.; Martin, D. C. *Macromolecules*, 30 (1997) 1524-1533.
- [32] Politis, J. K.; Curtis, M. D.; Yi, H.; Kaicki, J. *Macromolecules*, 32 (1999) 2484-2499.
- [33] A. Hosseinian, S. Jabbari, A. R. Mahjoub, M. Movahedi, *J. Coord. Chem.* 65 (2012) 2623-2633.
- [34] A. Hosseinian, H. R. Rahimipour, H. Haddadi, A. A. Ashkarran, A. R. Mahjoub, *J. Mol. Struct.* 1074 (2014) 673-678.