Crystal Structure Prediction for δ-HMX

Jose Arlo Gabriel Lopez Wui[1], Yi Yang[2], Dana O'Connor[2], Noa Marom[1, 2]

1 Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA

2 Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

Genarris is an open source Python package for random molecular crystal structure generation with physical constraints for seeding crystal structure prediction genetic algorithms and training machine learning models[1], [2].

Genarris 3.0 is an updated version of Genarris 2.0. The workflow has been changed. It also uses a Message Passing Interface (MPI) to achieve parallelization. First, Genarris randomly generates a pool of crystal structures for the given molecule. Then, geometry optimization using Rigid Press is performed on the randomly generated structures. Energy evaluation is then performed on the structures using the Density Functional Based Tight Binding approximation (DFTB). Then, clustering is performed, further down-selecting on the initial pool of structures. Finally, geometry optimization using the Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (LBFGS) is performed for relaxing the geometry of the structures.

As a case study, we used Genarris 3.0 to predict the crystal structure of the delta polymorph of HMX.

Introduction

Due to weak van Der Waals interactions, molecular crystals have the ability to exhibit polymorphism, a phenomenon where a given molecule crystallizes into multiple distinct forms. The crystal structure of a molecule affects its physical and chemical properties. In pharmaceuticals, the bioavailability and/or the mechanical properties of a drug are influenced by its crystal structure. In organic electronics, the charge carrier mobility may be affected. In high explosives, the stability and/or power of an explosive varies depending on the polymorph used.

Synthesizing polymorphs can be time-consuming, costly, and/or dangerous. A potential solution is computational crystal structure prediction: a growing field in which the goal is to computationally determine all possible crystal forms of a given molecule.

Genarris is an open source Python package for generating random molecular crystal structures. We used Genarris 3.0 to predict the crystal structure of **δ**-HMX.

HMX is a powerful high explosive with a variety of applications, such as being used as a detonator for nuclear weapons or as solid rocket propellant. Thus, its crystal structure is of interest as it is extremely useful. Since it is also used in explosives, determining its crystal structure computationally is a safer alternative as opposed to determining it experimentally. HMX has 4 polymorphs: **α**, ***β***, **γ**, and **δ**. It is worth noting that the delta polymorph, **δ**-HMX, is the most sensitive, which further exemplifies the importance of computational crystal structure prediction as a safer alternative to empirical methods.

Methodology

We first set the user defined parameters in a configuration file. We then ensured that the input molecule file which contains data about the HMX molecule is correct. Then, we ran Genarris on Bridges-2. Genarris runs through the entire workflow automatically.

Genarris 3.0 Workflow

The workflow of Genarris is as seen in Figure 1.

Figure 1: The workflow for Genarris 3.0

First, Genarris randomly generates a raw pool of molecular crystals for a given molecule. It does this by first estimating the unit cell volume given the desired number of molecules per unit cell, Z. It then randomly places molecules within the unit cell ensuring that the molecules are not placed too unphysically close. The user can also define the unit cell volume, preferably larger than that of the experimental structure's to allow Genarris ample room to randomly place molecules. Generation begins by first determining all compatible space groups with Z molecules in the unit cell. Genarris moves through the list of compatible space groups, checking once again that no two molecules are placed too close to each other. If the user-defined maximum number of failed attempts is reached for a space group, Genarris will proceed to the next compatible space group.

Then, geometry optimization is performed on the structures using Rigid Press. Energy evaluation on the structures follows next, using DFTB. Clustering is then performed by selecting the lowest energy structures, computing their atom-centered symmetry functions (ACSFs) to perform descriptor evaluation, followed by further down-selection. Finally, geometry optimization is performed on the structures using LBFGS.

Results

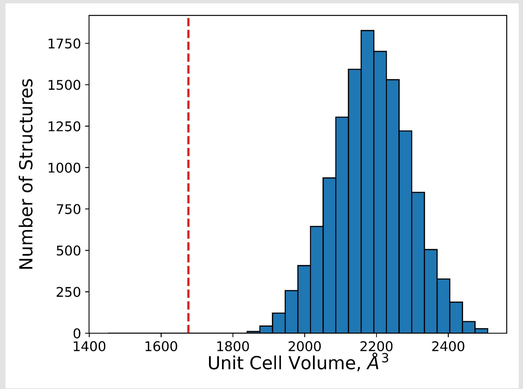

Figure 2: Unit cell volume distributions obtained at generation stage for **δ**-HMZ with Z = 6.

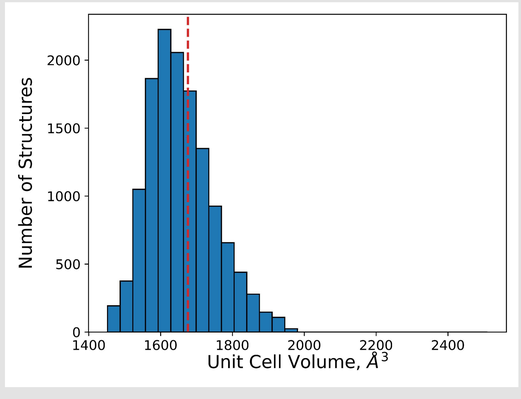

Figure 3: Unit cell volume distributions obtained at rigid press stage for **δ**-HMZ with Z = 6.

We were able to generate around 5000 structures containing 25 best-match structures. Table 1 shows the top 6 generated structures before and after Rigid Press according to their root mean square deviation of atomic positions of a cluster containing 20 molecules ($RMSD_{20}$ values) in Å. The $RMSD_{20}$ values were calculated using Mercury 4.0[3].

Table 1: Accuracy of the Generated Structures ($RMSD_{20}$ value)

Structure	CSD Refcode	Before Rigid Press	After Rigid Press
HMX	OCHTET03	No Best Match	**0.291**
		No Best Match	**0.293**
		No Best Match	**0.305**
		No Best Match	**0.391**
		No Best Match	**0.403**
		No Best Match	**0.491**

Results

We see that the structures in the initial pool generated by Genarris have larger initial unit cell volumes prior to Rigid Press. We see that Rigid Press produces more compact structures with smaller unit cell volumes on average.

We see that our generated structures after Rigid Press agree to within < 0.5 Å of the experimental structures.

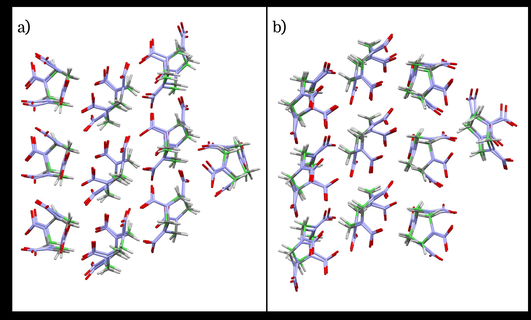

Figure 4: Overlay between the experimental structures (element-specific colors) and the structures from the pool generated by Genarris after Rigid Press for **δ**-HMX. a) $RMSD_{20}$ = 0.291 b) $RMSD_{20}$ = 0.293

Recommendations

We have generated the structure pool that includes the best matched structures, but we still need to improve it. Genarris 3.0 was able to complete all steps in the workflow except for the final geometry optimization step. We wish to have the generated structures as close as possible to the experimental structures, paying close attention to their the mean unit cell volume and $RMSD_{20}$ values. We will continue to experiment with the user-defined parameters in the configuration file under the generation and Rigid Press steps in the workflow, such as the number of structures per space group to generate, the mean unit cell volume, specific radius proportion, and the SR value.

Additionally, as the molecular structure of HMX is complex, we must still perform the final relaxation to achieve any further improvement. Another consideration would be to complete the final optimization step and determine the $RMSD_{20}$ values for the final pool of structures.

References

1: Li, X.; Curtis, F. S.; Rose, T.; Schober, C.; Vazquez-Mayagoitia, A.; Reuter, K.; Oberhofer, H.; Marom, N. Genarris: Random Generation of Molecular Crystal Structures and Fast Screening with a Harris Approximation. *The Journal of Chemical Physics* 2018, *148* (24), 241701. https://doi.org/10.1063/1.5014038.

2: Tom, R.; Rose, T.; Bier, I.; O'Brien, H.; Vázquez-Mayagoitia, A.; Marom, N. Genarris 2.0: A Random Structure Generator for Molecular Crystals. *Computer Physics Communications* 2020, *250*, 107170. https://doi.org/10.1016/j.cpc.2020.107170.

3: Mercury 4.0: from visualization to analysis, design and prediction
C. F. Macrae, I. Sovago, S. J. Cottrell, P. T. A. Galek, P. McCabe, E. Pidcock, M. Platings, G. P. Shields, J. S. Stevens, M. Towler and P. A. Wood, J. Appl. Cryst., 53, 226-235, 2020