Supporting information for Convergence properties of crystal structure prediction by quasi-random sampling

David H. Case, Josh E. Campbell, Peter J. Bygrave, and Graeme M. Day*

School of Chemistry, University of Southampton, Southampton, United Kingdom

E-mail: g.m.day@soton.ac.uk

Supporting Information Available

In this supporting document we include the data which was calculated but not included within the main text. These graphs, and the appropriate interpretation pertaining to each, follow on from the discussions in the main section of the paper. The reader will find all relevant discussion of each type of graph therein.

All data supporting this study are openly available from the University of Southampton repository at http://dx.doi.org/10.5258/SOTON/385297

Revised Williams99 parameters

All parameters describing interactions between C, N, O and H atoms are described using Williams' original Williams99 forcefield parameters, apart from hydrogen bond H...A interactions, which were reparameterised to work more effectively with the atomic multipole

*To whom correspondence should be addressed

electrostatic model. For H...A interactions, the pre-exponential parameter of the exp-6 model was modified from the Williams99 value. The parameters are given in the following table.

Table 1: Revised H...A parameters in the exp-6 intermolecular model used. C = 0 for all interactions.

hydrogen	acceptor	A (eV) with 6-311G(d,p) basis set	A (eV) with 6-31G(d,p) basis set
H2	N1	149	131
H2	N2	166	124
H2	N3	163	115
H2	O1	129	101
H2	O2	105	77
H3	N1	70	102
H3	N2	118	118
H3	O1	127	58
H3	O2	133	89
H4	N1	141	96
H4	N2	77	80
H4	N3	34	43
H4	N4	198	196
H4	O1	56	48
H4	O2	112	116

Structure generation figures

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

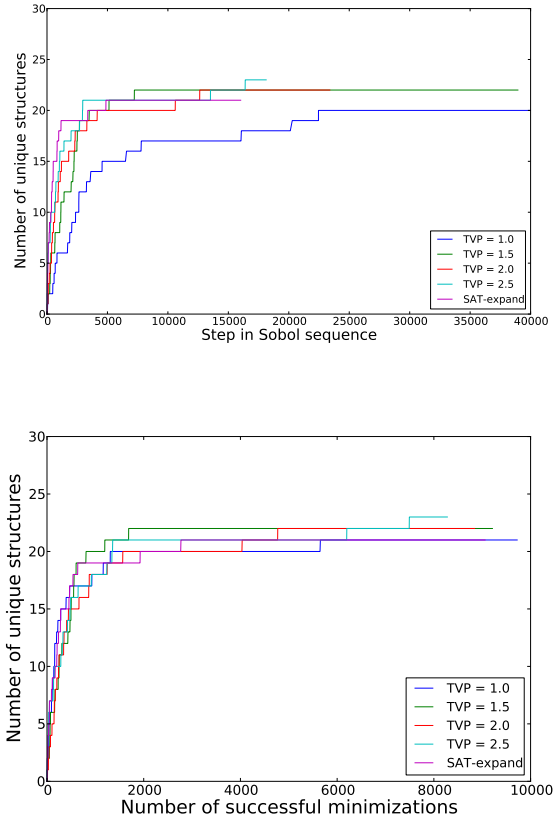
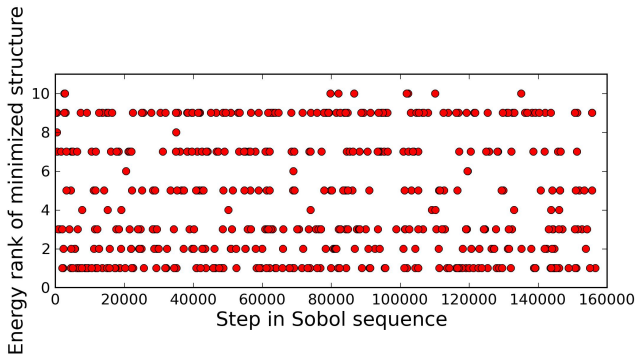
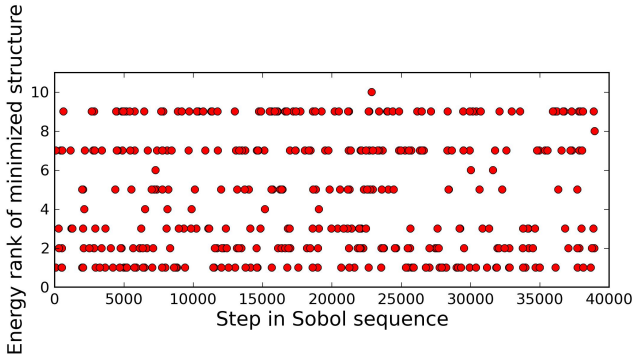


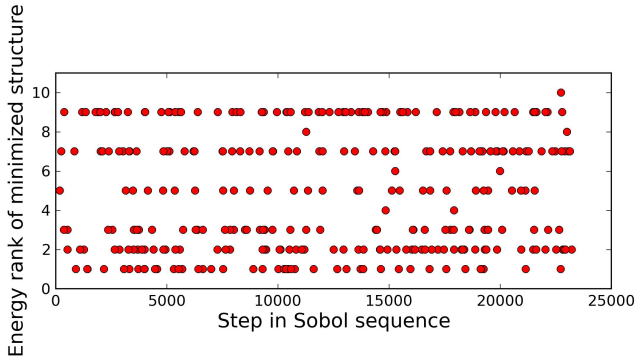
Figure 1: The convergence of the completeness of each set of structures which are 15 kJ/mol from the minimum in that set, for the case of CC1 in space group $P2_1/c$ is shown. That the data do not reach the same assymptote in all cases may be accounted for either by poor sampling, or structures whose energies lie on the energy cutoff.



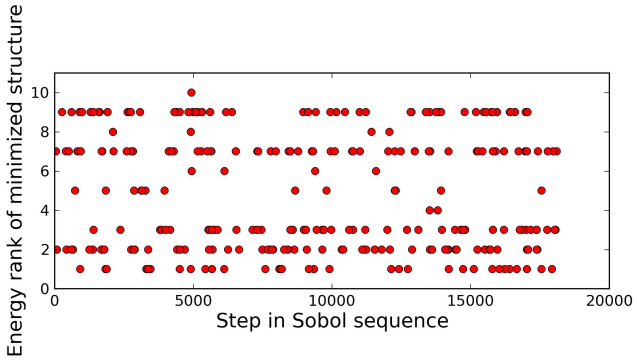
TVP = 1.0



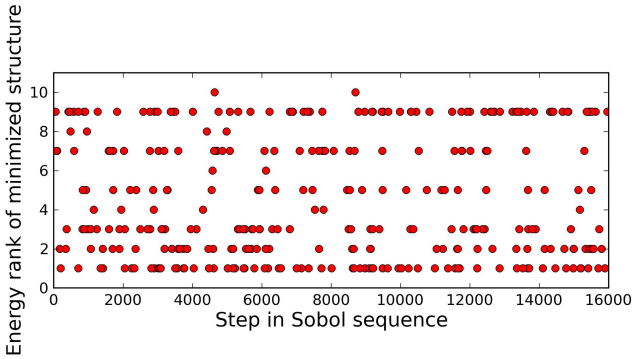
TVP = 1.5



TVP = 2.0

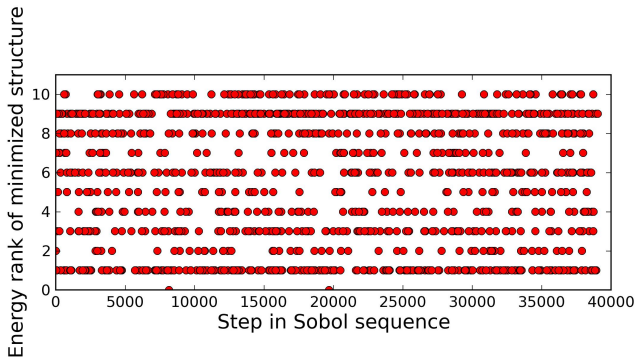


TVP = 2.5

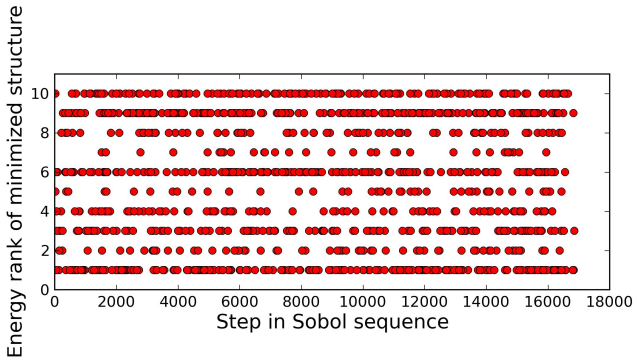


SAT Expand

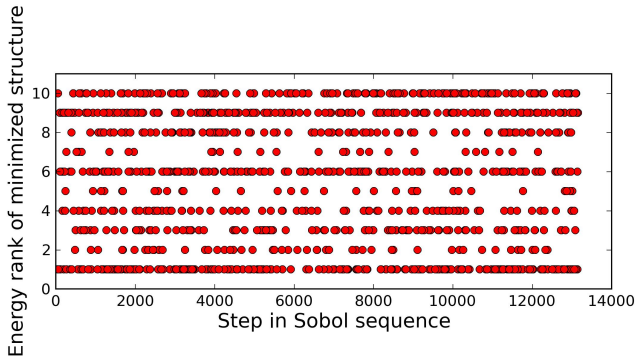
Figure 2: Occurences of the lowest ten structures of CC1, in the space group $P2_1/c$, as a function of the Sobol seed which generated each structure.



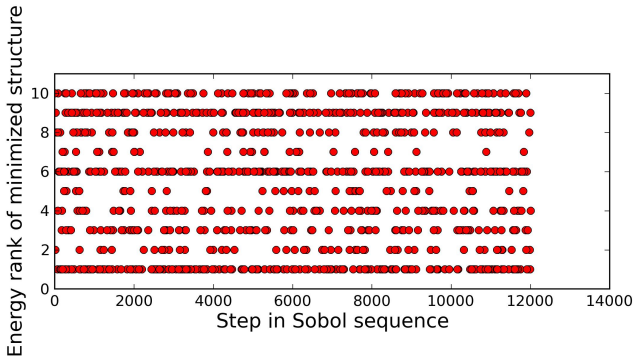
TVP = 1.0



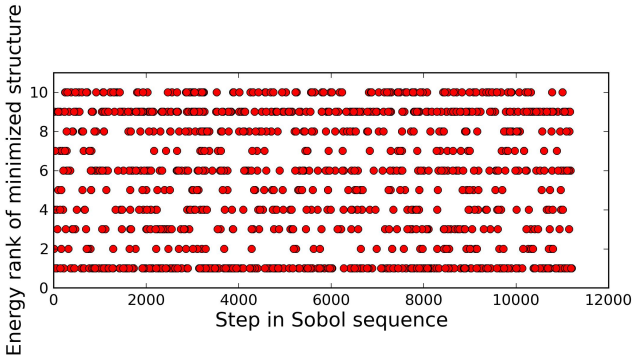
TVP = 1.5



TVP = 2.0

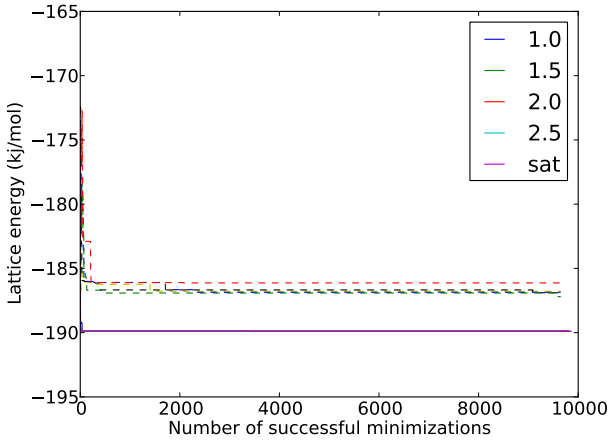


TVP = 2.5

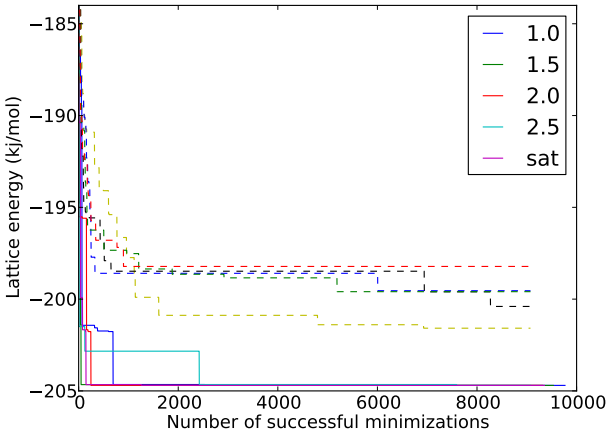


SAT Expand

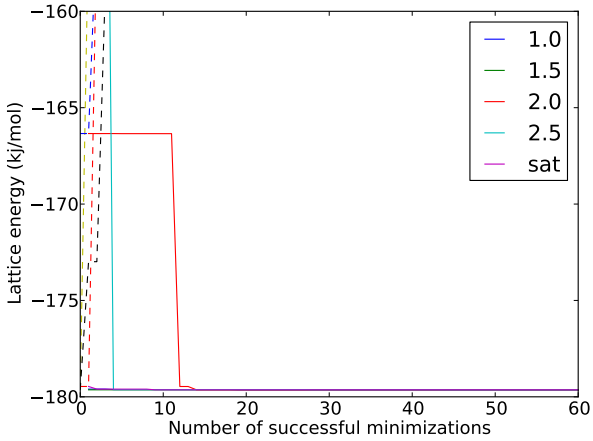
Figure 3: Occurences of the lowest ten structures of artemisinin, in the space group $P2_12_12_1$, as a function of the Sobol seed which generated each structure.



Quinacridone $P\bar{1}$

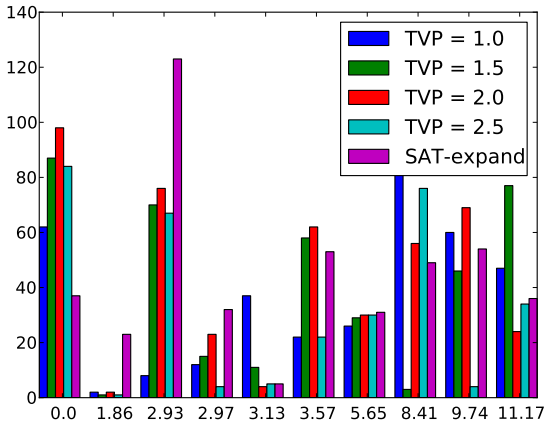


Quinacridone $P2_1/c$

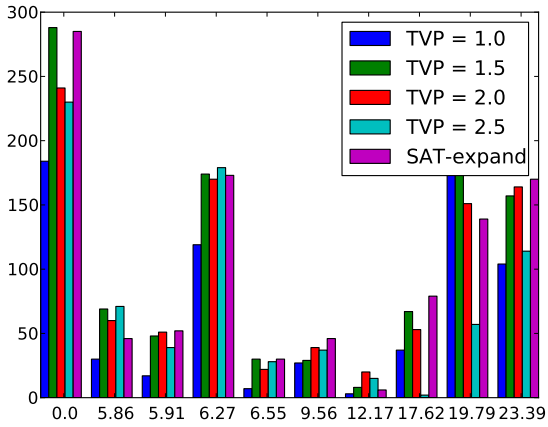


CC1 $P1$

Figure 4: The average lattice energy of the ten lowest energy structures is shown, as a function of the number of minimized structures generated in the experimentally observed space group for a) quinacridone in $P\bar{1}$, b) quinacridone in $P2_1/c$ and c) CC1 $P1$. The solid indicate the energy of the single lowest energy structure, where the color relates to the same method in the legend. The data had converged after 60 minimizations for c), so is not shown beyond this point for clarity.

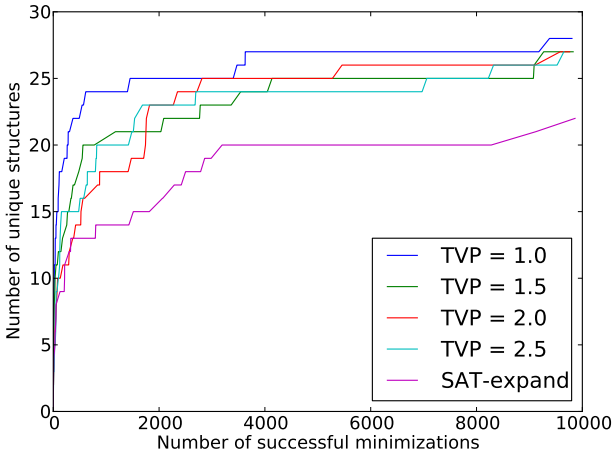


Quinacridone $P2_1/c$

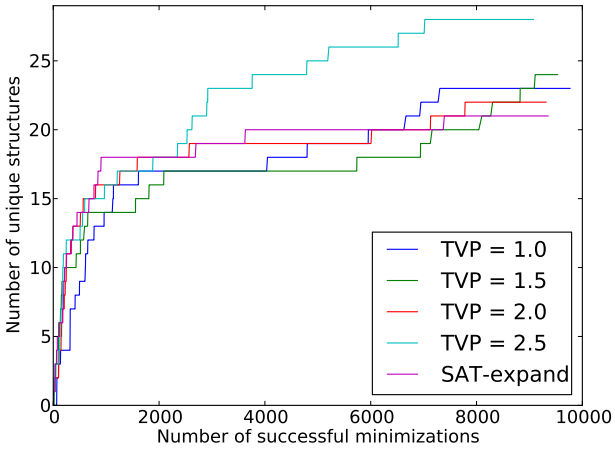


Quinacridone $Z' = 2$, $P\bar{1}$

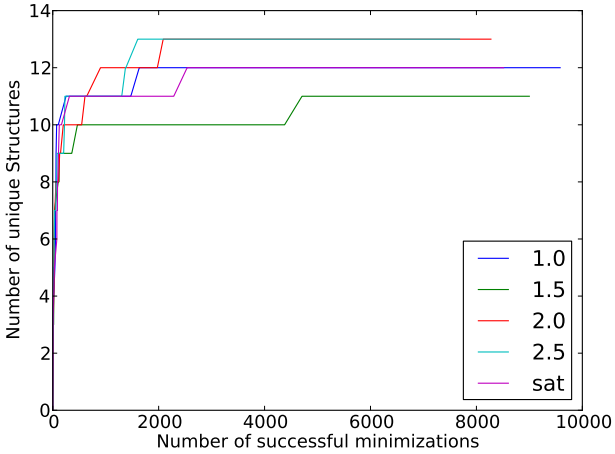
Figure 5: For each of the lowest 10 unique structures, for the denoted systems, the energy above the minimum in the set is displayed on the x-axis, and the number of times that it was found in the search is read from the y-axis.



Quinacridone $P\bar{1}$

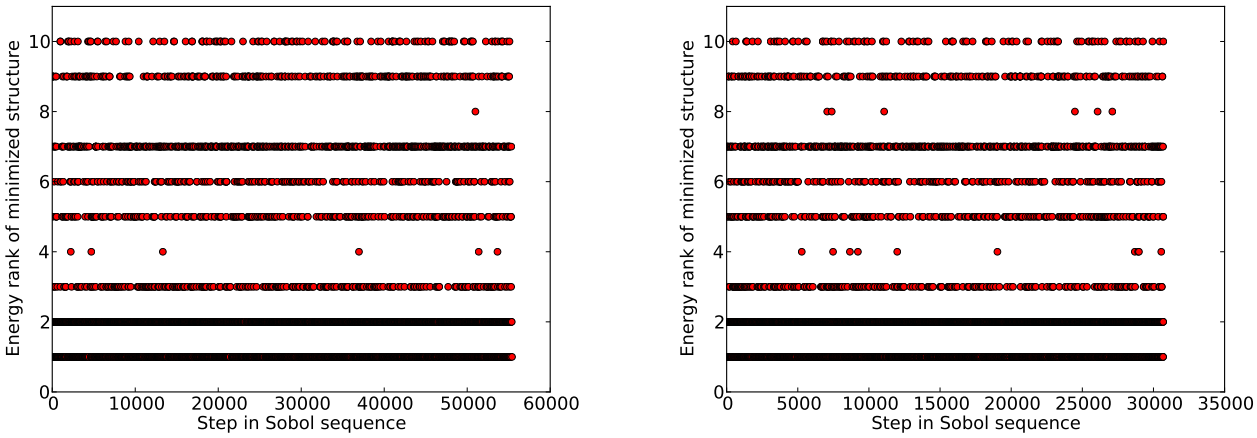


Quinacridone $P2_1/c$



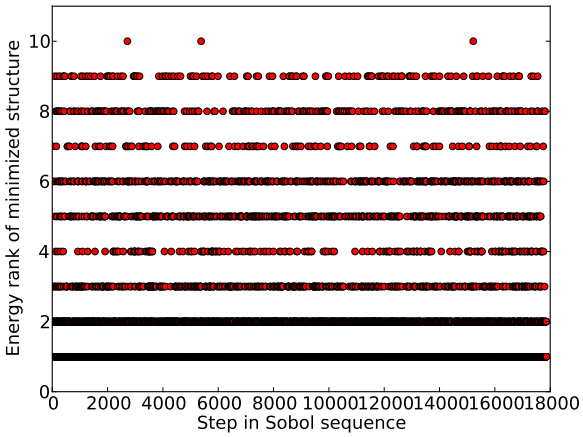
CC1 $P1$

Figure 6: The number of unique crystal structures within 15 kJ/mol (60 kJ/mol for CC1) of the global minimum displayed as a function of the total number of successfully energy minimized structures.

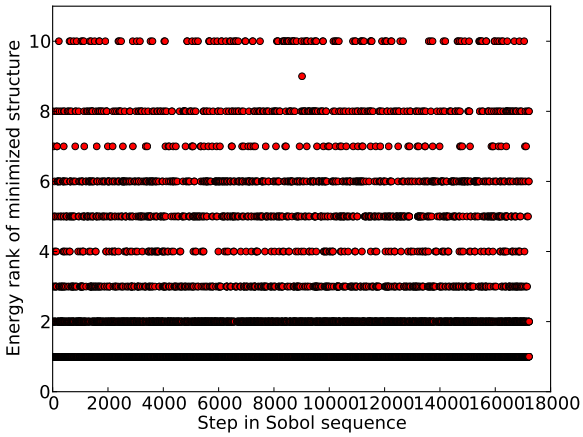


TVP = 1.5 TVP = 2.0

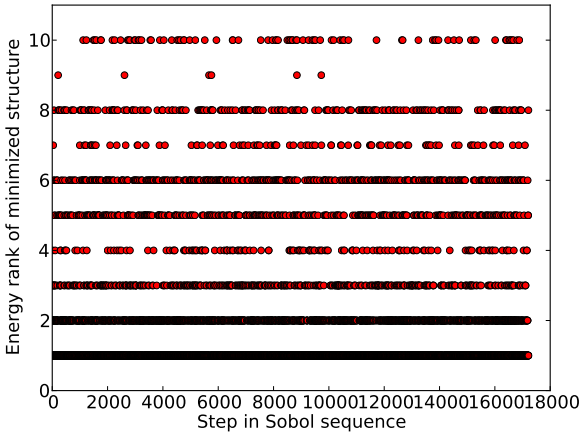
Figure 7: Occurences of the lowest ten structures of quinacridone, in the space group $P\bar{1}$, as a function of the Sobol seed which generated each structure.



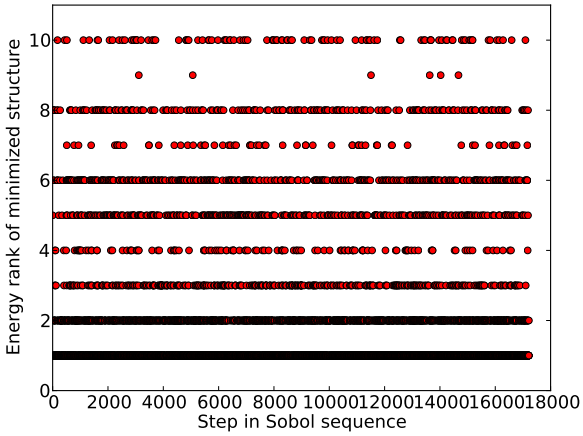
TVP = 1.0



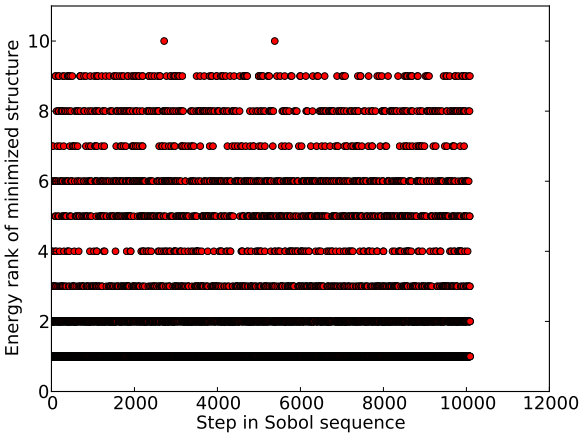
TVP = 1.5



TVP = 2.0

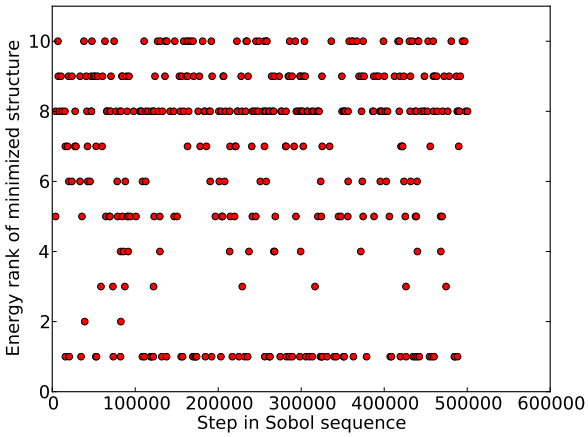


TVP = 2.5

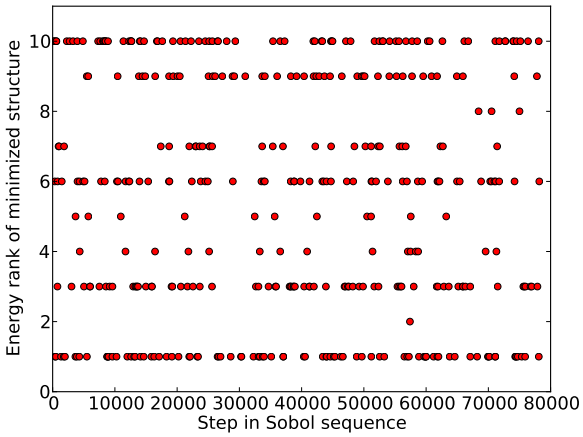


SAT Expand

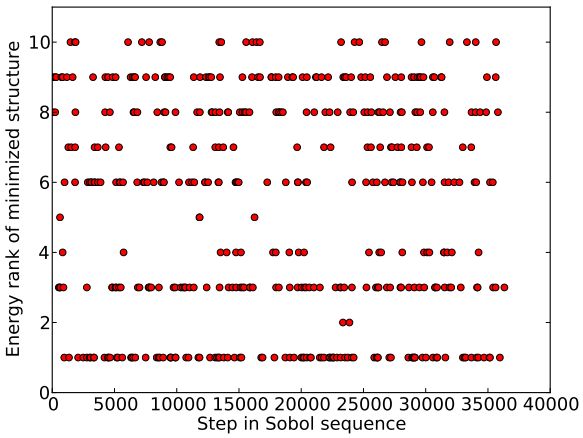
Figure 8: Occurences of the lowest ten structures of CC1, in the space group $P1$, as a function of the Sobol seed which generated each structure.



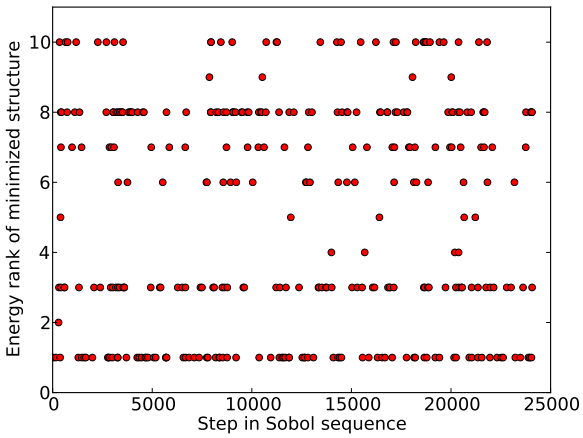
TVP = 1.0



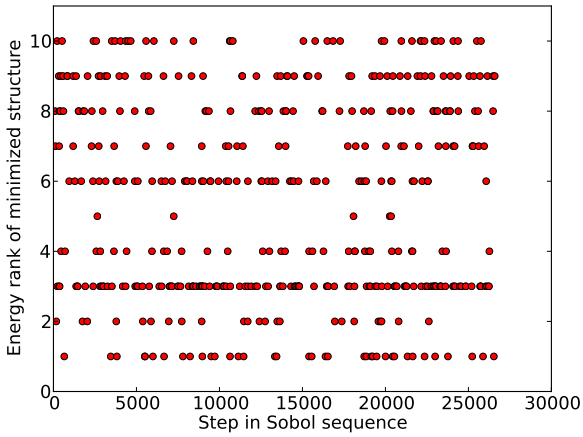
TVP = 1.5



TVP = 2.0

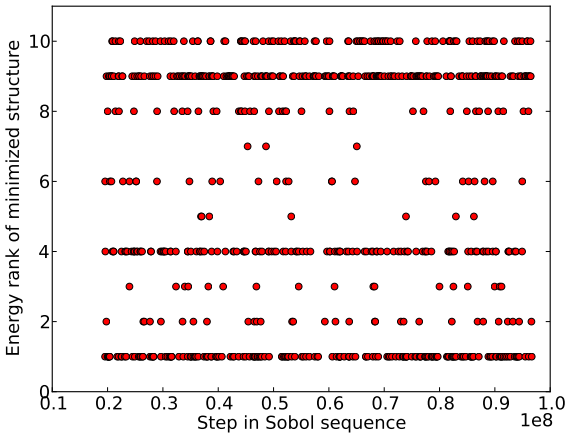


TVP = 2.5

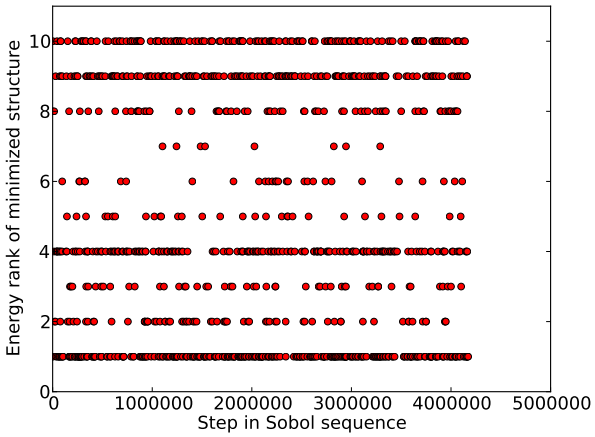


SAT Expand

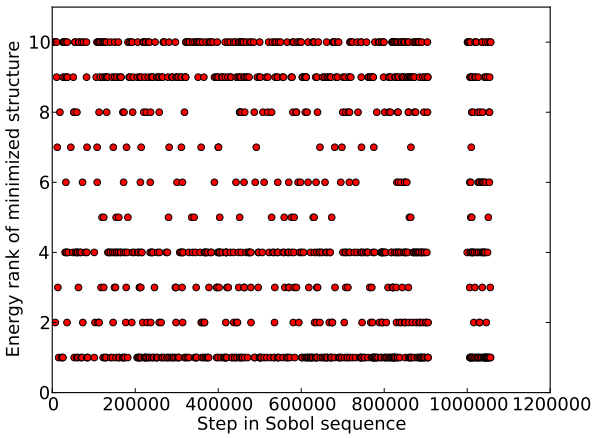
Figure 9: Occurences of the lowest ten structures of quinacridone, in the space group $P2_1/c$, as a function of the Sobol seed which generated each structure.



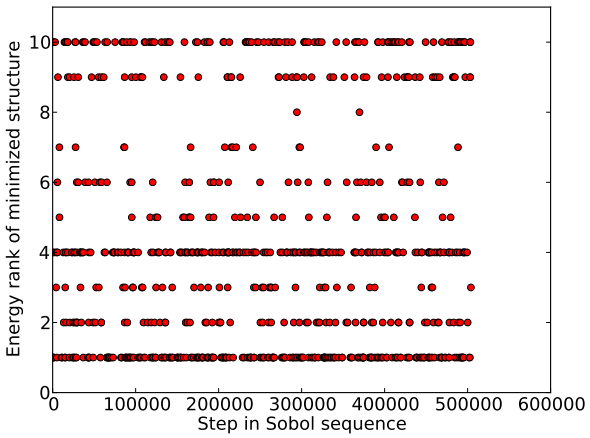
TVP = 1.0



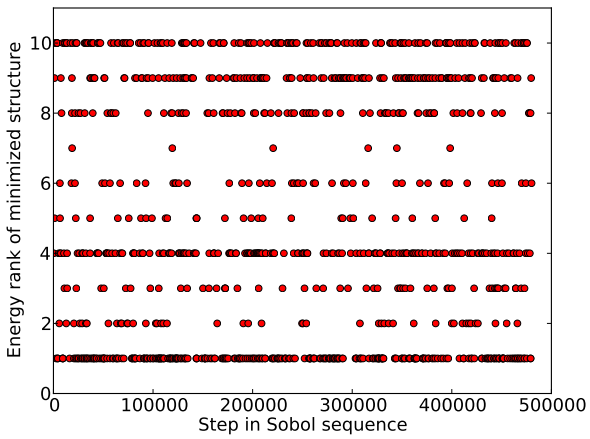
TVP = 1.5



TVP = 2.0



TVP = 2.5



SAT Expand

Figure 10: Occurences of the lowest ten structures of quinacridone, with $Z' = 2$, in the space group $P\bar{1}$, as a function of the Sobol seed which generated each structure. TVP = 1.0 has fewer structures overall due to the difficulty in generating $Z' = 2$ with a low TVP