Describing polarization in hybrid ab initio/empirical force-field models for crystal structure prediction

Benjamin I. Tan[1], David H. Bowskill[2], Adam Keates[3], Constantinos C. Pantelides[1], and Claire S. Adjiman[1]

[1]Department of Chemical Engineering, Sargent Centre for Process Systems Engineering and Institute for Molecular Science and Engineering, Imperial College London, London SW7 2AZ, United Kingdom, [2] Syngenta Crop Protection, Breitenloh 5, Münchwilen CH-4333, Switzerland, [3]Process Studies Group, Syngenta, Jealott's Hill International Research Centre, Bracknell, Berkshire, RG42 6EY, United Kingdom

Communicating author email: bt817@ic.ac.uk

Crystal Structure Prediction (CSP) seeks to predict the putative polymorphs of a solid system and rank them by their relative thermodynamic stability. CSP models of the lattice energy that combine ab initio calculations with empirical force-fields have proved to be effective for this purpose [1, 2]. Such models typically describe the intramolecular and electrostatic contributions to the lattice energy using molecule-specific quantities derived from ab initio calculations performed on isolated molecules. Repulsion-dispersion interactions, and any other effects not accounted for in the ab initio-derived terms, are embedded in the empirical terms. However, such approaches naively neglect the effect of polarization, which hinders the accurate description of intermolecular interactions like hydrogen bonds [3, 4]. Here, we investigate approaches to incorporate polarization more rigorously into the hybrid ab initio/empirical force-field (HAIEFF) model, at different levels of fidelity.

To embed polarization into the ab initio-derived terms of the HAIEFF model, the ab initio computations used to derive molecular electrostatic features (e.g., distributed multipoles) are performed with two approximations of the crystalline electric field: a Polarizable Continuum Model (PCM) [5], and an electronic-embedding method termed the self-consistent electronic response to point-charge (SCERP) model [3, 4]. To appraise these induction models, complementary HAIEFF force-fields have been independently parametrized in this work. By fitting an additional force-field without either polarization method, a third model (denoted FF_{ind}) is considered, wherein the effects of polarization are completely borne by the empirical component of the HAIEFF model.

During the force-field parametrization, we find that a Mie potential is better suited for accommodating the effects of polarization, compared to the more widely used Buckingham potential. Overall, the training data are more accurately represented when using PCM or SCERP, rather than relying on the force-field parameters in FF_{ind} to absorb the effects of polarization. Between PCM and SCERP, differences in parameter estimation accuracy are marginal. The three modeling approaches are then applied in CSP studies of a rigid monohydrate system and a flexible organic molecule. In both cases, at least one polymorph is predicted to have an unreasonably high relative energy when using FF_{ind}; this spurious behavior is diminished when using PCM or SCERP. Generally, PCM and SCERP yield predictions of similar quality, but the former fails to capture differences in polarization between crystal structures, leading to some erroneous predictions. Conversely, these nuances are more precisely described using SCERP, leading to CSP predictions that conform better with results from periodic DFT-D calculations. Ultimately, we demonstrate that polarization models, like PCM and SCERP, should be more widely adopted in CSP. Currently, SCERP offers only modest improvement over PCM, but the more physically-sound foundations of SCERP offer greater potential for systematic improvement.

[1] Reilly, A. M. et al. Report on the sixth blind test of organic crystal structure prediction methods. *Acta Crystallographica Section B* **2016**, *72*, 439–459.

[2] Hunnisett, L. M. et al. The seventh blind test of crystal structure prediction: structure generation methods. *Acta Crystallographica Section B* **2024**, *80*, 517–547.

[3] Welch, G.; Karamertzanis, P.; Misquitta, A.; Stone, A.; Price, S. Is the Induction Energy Important for Modeling Organic Crystals? *Journal of Chemical Theory and Computation – J CHEM THEORY COMPUT* **2008**, *4*, 522–532.

[4] Zhang, Y. Crystal Structure Prediction for Cocrystals and Hydrates: Energy Models and Structure Generation. Ph.D. thesis, Imperial College London, 2023.

[5] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chemical Reviews* **2005**, *105*, 2999–3094.