Applying structural informatics approaches to pharmaceutical supply chain processes

Andrew G. P. Maloney[1], Mathew J. Bryant[1] and Neil Feeder[1]

[1]*The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ, UK.*

E-mail: maloney@ccdc.cam.ac.uk

Pharmaceutical development currently stands poised to benefit from the big data revolution, but not equipped with a platform to reap the potential rewards. The vision of the Advanced Digital Design of Pharmaceutical Therapeutics (ADDoPT) project is to obtain a more sophisticated definition, design and control of optimised pharmaceutical manufacturing processes using data analysis and modelling on a large scale.

The ADDoPT consortium is made up of a variety of pharmaceutical companies, small and medium enterprises and research organisations, each with vast repositories of drug product property and processing data. At the heart of this is the Cambridge Structural Database (CSD), the world's most comprehensive database of small molecule crystal structures. By linking this wealth of manufacturing data gained from across the ADDoPT consortium back to knowledge of molecular and crystal structures, an unprecedented perspective of drug product design can be obtained.

Working as part of the ADDoPT collaboration, the CCDC is making use of the thousands of drug crystal structures in the CSD to establish the key links between structural features and potential formulation issues. By applying a solid-form informatics approach, the CCDC is directing the development and application of predictive tools that exploit our understanding of crystalline structures. These tools will then be applied to support the design of more robust manufacturing processes and the identification of the most appropriate formulation decisions.

Using the range of software available in the CSD Materials suite, crystal structures of drug molecules have been evaluated at the molecular, intermolecular and supramolecular level. Alongside the application of new bespoke analytical tools developed using the CSD Python API, a robust structural analysis protocol for the drug formulation scientist is emerging, which offers valuable insights into the potential downstream behaviour of drug candidates during formulation.