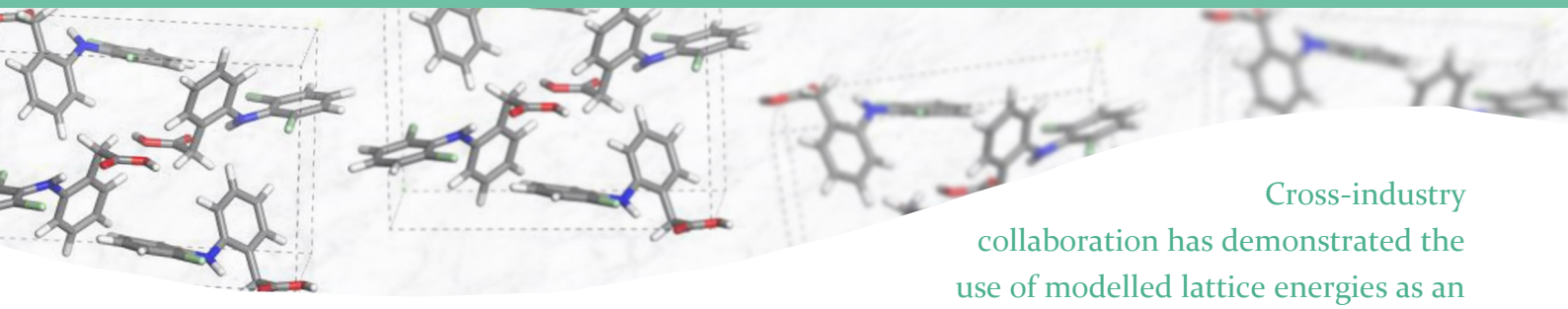
Lattice Energy Prediction using Big Data Approaches

A step towards understanding and predicting API solubility early in the development cycle



Cross-industry collaboration has demonstrated the use of modelled lattice energies as an early indicator of API solubility

Drivers

The sophistication of modern therapies has driven drug discovery towards higher molecular weight, lower solubility compounds. Understanding and predicting solubility early in the development cycle using lattice energy as an indicator property can provide business advantage by streamlining candidate selection, and linking up thinking between medicinal chemistry and drug development.

A new modelling tool allows lattice energies to be predicted for the crystalline form of a range of molecular variants. This gives an indication of the likely solubility directly from 2D structure, allowing candidates to be ranked with no need for experimentation.

Approach

Solubility of any crystalline material is the result of a balance of solvation and solid state packing effects, so effective modelling needs to consider solid-state as well as molecular properties. Work done under ADDoPT has affirmed the importance of the solid-state contribution to solubility and provided a practical model to account for them.

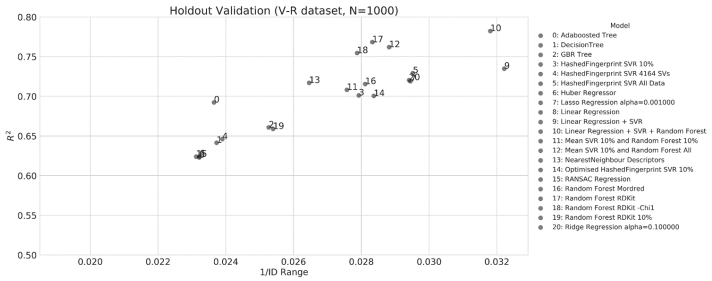
Key Features

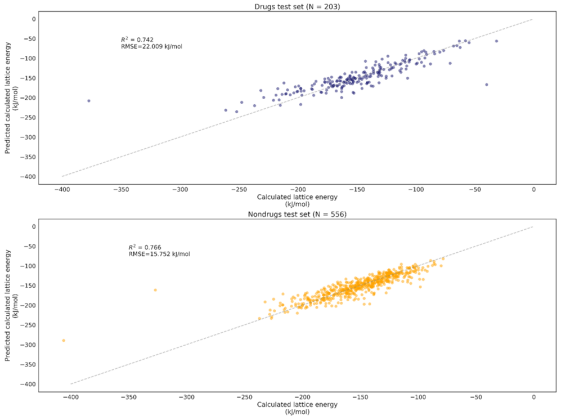
- The ability to predict lattice energies (a good indicator of solubility) from 2D structure alone developed and validated in an industry-relevant molecular space
- A big data, cross-industry approach maximizing the ability of model to cope with future evolution of drug development space
- Bridging the gap between medicinal chemistry and drug development

Fifty thousand molecules in the CSD database with known crystal structures have been analyzed using machine learning to relate 2D molecular descriptors to their lattice energies. The scarcity of directly measured lattice energies has been overcome by using calculated values derived from single crystal structure experiments, using best available lattice energy calculation methods developed in ADDoPT at the University of Leeds.

An ADDoPT Case Study featuring collaboration between Pfizer, the STFC Hartree Centre, the Cambridge Crystallographic Data Centre, and the University of Leeds

Model validation and testing





After training, each model was validated on a validation dataset for comparative purposes (above). A combined model was the best performing, performing well on both drug and non-drug test data (right). As expected, performance was slightly better on a nondrug test set as the training data was composed of all non-drug molecules.

This may be understood as the necessary thorough groundwork to enable what will ultimately be a simple and streamlined routine workflow: the user feeds in a 2D structure and receives a lattice energy prediction as a marker of likely solubility, to guide further experimentation.

Educating medicinal chemists in the importance of the solid state to likely solubility of candidate molecules, and aligning ways of thinking within drug development

Results and Benefits

Material scientists at Pfizer have evaluated and further developed the new predictive model, applying it to 1500 drug structures within the area of chemical space of interest to them. Results show good correlation between model-derived and crystal structure-calculated lattice energies.

Beyond showing that the predictive model is sufficiently robust to perform in the chemical space of specific interest to Pfizer, ADDoPT has enabled the combination of diverse expertise (at Leeds for lattice energy calculation and Hartree for data analysis) to uniquely build and analyse a data set with overarching coverage of the full chemical space of crystalline materials, and demonstrate its ability to perform in the focused space of drug-interest.

By educating and informing medicinal chemists of the importance of the solid state upon the likely solubility of candidate molecules, the model helps align thinking within, and "bridge the gap" into, drug development. The ultimate benefit of this work will be a more streamlined and accelerated development timeline.

Further Steps

Awareness raising of this new development is already underway in the medicinal chemistry community.

Ultimately, this is the first step towards achieving a vision of connected development where 2D structures are used to predict 3D crystal properties, crystal properties predict particle properties, and particle properties in turn predict bulk powder properties.

Transforming pharmaceutical development and manufacture

Addressing the pharmaceutical industry's desire to deliver medicines more effectively to patients, the ADDoPT project has developed and implemented advanced digital design techniques that streamline design, development and manufacturing processes.

ADDOPT

ADVANCED DIGITAL DESIGN OF PHARMACEUTICAL THERAPEUTICS

www.addopt.org

ADDoPT is a collaboration instigated by the Medicines Manufacturing Industry Partnership, and part funded under the Advanced Manufacturing Supply Chain Initiative, a BEIS initiative delivered by Finance Birmingham and Birmingham City Council.