Figure 1. ZMOFs design towards controlled drug delivery (adapted from Rocca et al, Acc.Chem.Res., 44(2011), 957)

Keywords: BioMOFs, pharmaceuticals, drug delivery

MS9-P6 PANDDAs: multi-dataset methods for finding hits from fragment screening by X-ray crystallography

Nicholas M. Pearce[1], Sebastian Kelm[2], Jiye Shi[2], Charlotte M. Deane[1], Frank von Delft[1,3]

1. University of Oxford
2. UCB Pharma
3. Diamond Light Source

email: nicholas.pearce@stx.ox.ac.uk

We present novel PAN-Dataset Density Analysis (PANDDA) methods for processing the data obtained from Crystallographic Fragment Screen (CFS) campaigns.

CFS experiments are ideal for identifying small "fragment" molecules that bind to a protein, but historically this method has been disfavoured due to low throughput and high complexity. With increasing automation and other technical advances, it is now a streamlined, routine experiment at beamline I04-1 at Diamond Light Source.

However, bottlenecks remain in the processing of the resulting data. The small compounds frequently bind with low occupancy, leading to subtle or ambiguous signal in the observed electron density. Datasets must be inspected manually to ensure that binding fragments are not missed. Manual inspection is time consuming and inaccurate, in particular when there are many binding sites, or when fragments bind outside of known binding sites.

CFS experiments using soaking protocols lead to datasets that are broadly isomorphous (near-identical). In this case, the interpretation of CFS data requires the examination of hundreds of similar datasets to identify binding events. Rather than processing and inspecting each dataset separately - the current state-of-the-art - there is a scientific opportunity to analyse these datasets simultaneously.

The nature of the screening experiment results in most of the datasets displaying no evidence of the fragment binding – 'empty' datasets. Rather than discarding these datasets, we instead average them to generate a very accurate picture of the fragment-free crystal electron density. The variation in the electron density at each point in the aligned datasets is then analysed to identify outliers in individual datasets. This characterization of the fragment-free crystal creates a reliable baseline with which we can identify the areas of individual datasets where a fragment binds, or where a structural shift is observed.

Preliminary results show the PANDDA approach is capable of accurately identifying bound fragments in CFSs. On a training set of 200 structures, all 3 known hits are identified by the method, whilst 80% of datasets are correctly rejected as empty (false negative rate = 0). The PANDDA method also increased the number of hits: 3 extra hits were identified that were missed in the manual analysis.

The PANDDA method provides an objective indication of the presence of bound fragments, and can detect the binding of fragments with weak affinities.

Keywords: Fragment Screening, Crystallographic Methods