

PXRD WITH RAMAN SPECTROSCOPY, DSC AND IR DATA

Christopher Gilmore, Gordon Barr and Gordon Cunningham, WestCHEM, Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland, UK.

Powder X-ray diffraction is usually considered the gold standard in high throughput studies designed to identify polymorphs, salts, co-crystals *etc.*, but other techniques such as Raman and IR spectroscopy, or differential scanning calorimetry (DSC) can have a major role to play also especially with poorly characterised samples. For example, PPXRD usually shows large differences between different polymorphs, but in Raman spectroscopy the differences are subtle and related to such things as H-bonding. Raman spectroscopy, on the other hand, can be very robust with poor quality samples. High throughput data can often also be of relatively poor quality and robust statistical procedures are also required when classifying it. Other numerical data may also be available. The problem arises, however, as to how best to use and combine such disparate data sources, and how best to visualize the data. The latter is particularly important with large data sets having multiple sources.

We have established a methodology used in the computer program, POLYSNAP 3, that is designed for such studies with up to 4 different data types [1, 2]. The PolySNAP methodology compares each pattern to every other on a point by point basis using the entire measured data set to generate a correlation matrix for each data type; the same procedure is used for all data types but with different pre-processing options. It is possible to weight each individual matrix and thus combine them to get an overall matrix, but how are the weights to be estimated? We use a dynamic weighting scheme based on the INDSCAL (INdependent SCALing of differences) method devised by Carroll & Chung [3] which is wholly automatic and very robust; it allows any combinations of data. For visualization we use techniques including dendrograms, metric multidimensional scaling, parallel coordinate plots and silhouettes.

Examples using combined PXRD, Raman, IR and DSC data to study polymorphs will be presented, along with other data sets of pharmaceutical interest. It can be seen that no one data type adequately resolves the contents of the samples and mixtures of them, but combining data types using the INDSCAL method is wholly successful.

1. G. Barr, G. Cunningham, W. Dong, C.J. Gilmore, and T. Kojima (2009). *J. Appl. Cryst.* **42**, 706-714.
2. G. Barr, W. Dong, and C.J. Gilmore (2009). *J. Appl. Cryst.* **42**, 965-974.
3. J.D. Carroll, J.D. & J.J. Chang, J.J. (1970). *Psychometria* **35**, 283-319.