[3] Fries, E.; Sieg, K.; Püttmann, W.; Jaeschke, W.; Winterhalter, R.; Williams, J.; Moortgat, G., *Science of the Total Environment*, accepted.
[4] Fries, E.; Starokozhev, E.; Haunold, W.; Jaeschke, W.; Mitra, S. K.; Borrmann, S.; Schmidt, M. U., *Atmospheric Environment* 2007, in press.
[5] Mayo, S. L.; Olafson, B. D.; Goddard III, W. A., *J. Phys. Chem.* 1990, *94*, 8897-8909.

MS15 O4

Acid-base systems, understanding the formation of co-crystals versus salts. <u>G. Ramon</u>, L.R. Nassimbeni, *Supramolecular Crystallographic Group, P.D. Hahn Building, Chemistry Departement, University of Cape Town, Rondebosch 7701, South Africa.*
E-mail: <u>gaelle.ramon@uct.ac.za</u>

Keywords: synthon, co-crystal, salt

We have employed the concept of supramolecular synthons[1] to prepare acid-base molecular aggregates and have studied their structures. The synthon chosen was the carboxylic-acid: 2-aminopyridine pair, which display mutually compatible hydrogen bonding interactions of the type N-H···O and O-H···N (Figure 1a). We observed that the hydrogen of the acid carboxylic group could be transferred to the nitrogen of the pyridyl group according to the difference in the pKa of the two compounds (Figure 1b).

From the results obtained, we correlated the O···N distances with the difference in pKa of the two compounds. The O···N distance decreases with larger values of ΔpKa.

In the meantime, we decided to study the relative stability of the structures obtained by calculating their lattice energy. Because of the proton transfer, the fragments of the structures are charged. From these calculations, we predict the results of competition experiments. These experiments consist in studying from which acid one base crystallises when placed in a mixture of two acids. The competition experiments are then repeated at different pH conditions so that the selectivity could be enhanced or inverted.
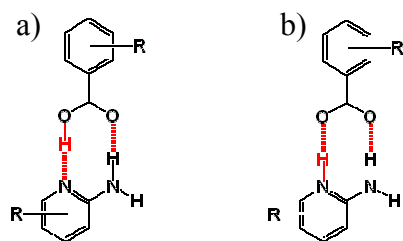

Figure 1: the carboxylic-acid: 2-aminopyridine synthon
a) formation of the co-crystal
b) formation of the salt with transfer of the proton (in red)

[1] Desiraju G.R., *Angew. Chem. Int. Ed. Engl.*, 1995, 34, 2311-2327)

MS15 O5

Complementarity of cocrystal formers in the CSD
<u>László Fábián</u>, Pfizer Institute for Pharmaceutical Materials Science, Cambridge Crystallographic Data Centre, Cambridge, UK. E-mail: fabian@ccdc.cam.ac.uk

Keywords: cocrystals; databases; statistical analysis

A database of 62109 reliable and complete organic crystal structures was extracted from the Cambridge Structural Database [1]. Sum formulae and InChI identifiers [2] were calculated for each residue. Using this data and a list of common solvents [3], 1914 structures were identified as cocrystals. Molecular descriptors, usually applied in QSAR studies, were calculated for all molecules in the cocrystal dataset. The resulting database describes pairs of molecules that form cocrystals with each other in terms of the calculated molecular properties.

The cocrystal properties database may serve a dual purpose. It can be a source of empirical rules and it can form the basis of a computer application to screen a database for likely cocrystal formers. Both potential applications require statistical methods to extract important and significant information from the large amount of data available (2 x 130 descriptors for each molecule pair).

The simplest approach to distinguish between likely and unlikely cocrystal formers is a classification of molecules, followed by an assessment of which classes of molecules crystallize together frequently. Since the applied descriptors are strongly correlated, principal component analysis was used to transform the 130 descriptors into a set of 20 independent variables. Ward's hierarchical agglomerative clustering method identified 15 clusters. The majority of the clusters can be described in chemical terms (*e.g.*, electron donors; zwitterionic amino acids). Preferences for cocrystal formation were investigated by computing relative frequencies (treated as probability estimates) of each possible group pairing (15x15). Strong preferences, *i.e.*, frequent combinations of groups were apparent from the data. They usually coincide with rules one would expect using chemical knowledge. For example, polyaromatic hydrocarbons frequently occur in cocrystals with charge-transfer acceptor molecules, perfluoro compounds or with each other. The successful identification of such preferences in a purely mathematical approach suggests that a molecular descriptor based classification can be a useful component in an automatic screening application.

A more detailed understanding of the compatibility of cocrystal formers can be obtained from the analysis of individual molecular properties. Is there a relationship between the size, shape, polarity or donor-acceptor imbalance of molecules in a cocrystal? Spearman's correlation coefficients (ρ) were calculated for all possible descriptor pairs (130x130) to check for the existence of such relationships. (The strongly biased distribution of descriptor values limits the usability of more conventional statistical descriptors.) Two-dimensional density plots were created for all variable pairs showing a correlation with $|\rho| > 0.25$. The statistical significance of such correlations was further elucidated using box plots. Representative examples of both strongly correlated variables (*e.g.* those related to shape) and expected but not observed correlations (*e.g.* difference between the number of donors and acceptors) will be presented and discussed.

[1] Allen, F.H. *Acta Cryst.*, 2002, B58, 380-388.
[2] http://www.iupac.org/inchi/ [3] Görbitz, C.H.; Hersleth, H.-P. *Acta Cryst.*, 2000, B56, 526-534.