

Cluster analyses of metal-organic fragments using the *dSNAP* software

Anna Collins, Chick C. Wilson
and Christopher J. Gilmore*

WestCHEM Research School, Department of
Chemistry, University of Glasgow, University
Avenue, Glasgow G12 8QQ, Scotland

Correspondence e-mail:
c.gilmore@chem.gla.ac.uk

The *dSNAP* computer program has been used to classify searches of the Cambridge Structural Database for two ligands: $-\text{O}-\text{CH}_2-\text{CH}_2-\text{O}-$ and $\text{N}(\text{CH}_2\text{CH}_2\text{O}-)_3$ commonly found in metal-organic systems. The clustering method used is based on total geometries (*i.e.* all the lengths and angles involving all the atoms in the search fragment, whether bonded or not) and proved capable of distinguishing in a wholly automatic, objective way between different types of metal complex purely on the basis of the geometry of the ligand and the relative positions of the O atoms to the metals.

Received 24 July 2009
Accepted 20 September 2009

1. Introduction

The Cambridge Structural Database (CSD; Allen, 2002) provides excellent search and analysis tools including: *ConQuest* (Macrae *et al.*, 2006), *ISO STAR* (Bruno *et al.*, 1997), *Mogul* (Bruno *et al.*, 2004), *Vista* (Cambridge Crystallographic Data Centre, 1994) and *Mercury* (Macrae *et al.*, 2006), but even with this software extracting and analyzing chemical information from the CSD can be daunting because of the volume of data available.

The use of cluster analysis to classify the results of searches of the CSD on the basis of total geometries (*i.e.* all the lengths and angles involving all the atoms in the search fragment, whether bonded or not) has been the subject of a number of recent papers. These include conformations of enones and enamines (Collins, Barr *et al.*, 2007), transition metals with salicylaldiminato ligands (Parkin, Barr, Collins *et al.*, 2007), conformations of pyranose sugars (Collins *et al.*, 2008), sulfonamide conformation (Parkin *et al.*, 2008), intermolecular interactions between carboxylic acids and secondary amides (Collins, Parkin *et al.*, 2007), and the comparison of whole crystal structures (Parkin, Barr, Dong *et al.*, 2007; Collins *et al.*, 2009). There is a review of the method and its applications by Parkin (2008). The majority of these studies have involved organic materials, but the same principles apply to metal-organic systems. There are, however, key differences between investigating metal-organic ligand systems and purely organic ones. Larger metal-organic systems frequently contain several instances of a particular ligand within a single molecule, and this may result in several hits in searches for the ligand or ligand fragment if the molecule is in a low-symmetry space group or only one if the molecule is highly symmetrical. The nature of the metal centre is also important. For example, how much of an effect does it have on the geometry of the organic ligand? The ligands themselves often exhibit high symmetry, and this must be dealt with effectively by the clustering program. Given the high interest in metal-ligand complexes in

a wide range of systems, rationalizing metal–ligand interactions offers underpinning value in developing and understanding new metal–organic materials.

In the existing literature, Orpen (2002) provides a review of applications of the CSD to molecular inorganic chemistry; there is also a review of developments in inorganic crystal engineering by Brammer (2004). Other work in this area includes a study on the coordination of carboxylates by Hocking & Hambley (2005), Fey, Harris *et al.* (2006) and Fey, Tsipis *et al.* (2006). Harris *et al.* (2005) have developed knowledge bases of transition-metal geometries and their associated ligands; Mínguez Espallargas *et al.* (2006) have investigated the interface between inorganic and organic fragments, and have described intermolecular halogen–halogen contacts, whilst Dance (2003) has correlated observed inorganic intermolecular motifs with their calculated energies.

In this paper the *dSNAP* (Version 1.0) computer program (Barr *et al.*, 2005) has been used to investigate two metal–organic systems using structural information mined from the

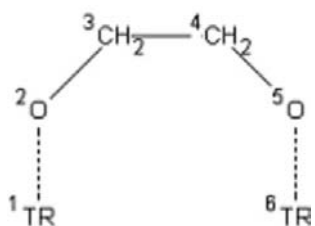


Figure 1
The first search fragment input into the CSD.

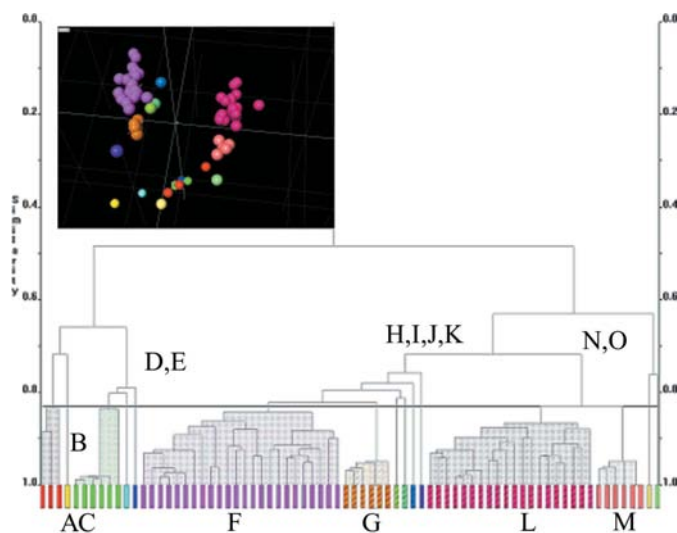


Figure 2
(a) The dendrogram and the metric multi-dimensional scaling (MMDS) plot as an inset derived from the total geometries of the $\cdots\text{O}-\text{CH}_2-\text{CH}_2-\text{O}\cdots$ ligand. The clusters are labelled A–O. B is a singleton yellow cluster; D and E are also singletons coloured light and dark blue; H, I, J and K are singletons lying between G and L, and N and O are singletons at the right-hand end of the dendrogram. At a cut level of 0.83 there are 15 clusters. The MMDS plot confirms the representation of the data as given by the dendrogram.

CSD. In the first example, the mode of coordination of the ligand component to the metal was defined. In the second system a common ligand was defined, but no metal component was specified. These examples highlight the importance of how the search is defined; this helps define what information is available, and the focus of the results.

2. The $-\text{O}-\text{CH}_2-\text{CH}_2-\text{O}-$ ligand

The search fragment for the $-\text{O}-\text{CH}_2-\text{CH}_2-\text{O}-$ ligand in metal–organic complexes was defined and numbered as shown in Fig. 1. TR represents any transition metal, and the bond type between the metal and O atoms is of type *any*. H atoms attached to C3 and C4 have been defined implicitly, *i.e.* the H-atom positions do not have to be recorded in the database structure, but the only atom type that can be found attached to these two C atoms is hydrogen. The connectivity of the O atoms is not completely defined; the ligand may exist as it is drawn, or O2 and O5 may have other atoms attached to them. The way in which this fragment was defined resulted in hit structures containing at least two metal atoms; structures where O2 and O5 were bonded to a single metal atom were not included in the resulting hit list. If only a single metal were included, the symmetry of the system would be reduced and so spurious multiple hit fragments might occur within a single structure. Appropriate care in setting up a search fragment and symmetry issues have emerged as important aspects of *dSNAP* input. Proper use of the CSD and symmetry make optimum use of the program. In this case a total of 26 hit structures were returned, with 74 hit fragments in searches from the CSD, Version 5.30 with updates dated November 2008 and February 2009. The conditions imposed on the search were: $R < 0.05$; not disordered, no errors, not polymeric, no ions, no powder structures, and only metal–organics. The fragment has topological symmetry, and the *dSNAP* program corrects for this to avoid the problem of artefacts and incorrect clustering. All the symmetry possibilities are explored for each fragment, and atoms are renumbered if necessary to ensure that all the fragments are optimally consistent with each other.


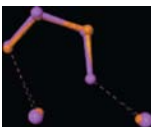
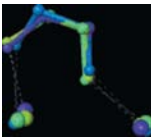
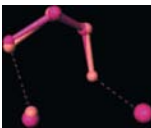
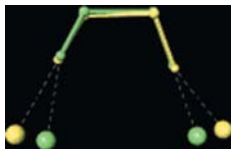
The number of hit fragments is bigger than the number of hit structures for several reasons: the search fragment can occur more than once in some of the structures; situations with $Z' > 1$, or multiple instances of the same fragment in a single molecule. The latter case is particularly common in metal–organic structures containing multiple metal centres linked by several ligands. The molecules may also exhibit high symmetry, and only symmetry-independent instances of the search fragment are counted in these cases. In the discussion that follows, when a structure contains more than one search fragment the refcode is appended by $_{nn}$, where nn is the number of the fragment in that structure.

The results of the cluster analysis are summarized in Fig. 2 using our usual representations of a dendrogram and a metric multi-dimensional scaling (MMDS) plot. The use of these techniques has been described fully elsewhere (see, for example, Parkin, 2008) so only a brief summary will be given

Table 1

The clusters identified by *dSNAP* with the relevant fragments superimposed.

The colours of each fragment are taken from the dendrogram. An asterisk (*) indicates multiple determinations of the same refcode.

Cluster (colour)	No. of fragments	Geometry of fragment
A (red)	3	
B (yellow)	1	
C (green)	6*	
D (cyan)	1	
E (blue)	1	
F (pink)	24	
G (brown stripe)	6	
H (green stripe)	1	
I (mint green stripe)	1	
J (blue stripe)	1	
K (purple stripe)	1	
L (pink stripe)	20	
M (peach)	6	
N (light yellow)	1	
O (light green)	1	

here. The dendrogram takes the form of a tree diagram, where each box at the bottom of the figure represents a single hit fragment. The boxes are joined by horizontal tie bars which link together fragments according to the calculated similarity between each connected branch. The vertical axis is a similarity scale, with zero similarity at the top, and a similarity of 1.0 at the bottom, *i.e.* if two fragments are joined by a tie-bar near the bottom of the dendrogram then they can be considered very similar, justifying their being grouped together. If two branches do not meet until near the top of the dendrogram the associated fragments are only loosely related to each other. The horizontal line that spans the width of the dendrogram marks the cut level. Any fragments that are linked at a higher level of similarity than the cut level belong to separate clusters; the principal of the cluster analysis using *dSNAP* analysis is that molecular fragments within a single cluster are geometrically similar, while those in different clusters will have some geometric feature or features that are quite distinct. The identification of clusters, and determining

the geometric features underlying the differences between clustered fragments, is the main structural chemistry driver for *dSNAP* applications. Raising the cut level will decrease the number of clusters and *vice versa*. For the data used here, there are 15 clusters at a cut level of 0.83.

In the MMDS plot each sphere represents a hit fragment. Fragments that are located close together in space can be expected to have similar geometries. The colours are taken from the dendrogram; as the MMDS plot and dendrogram are calculated by different methods, this provides a check on the consistency of the clustering. Ideally, spheres of the same colour will be close together and will form clusters that are well separated in space from spheres of different colours; this indicates good agreement between the MMDS plot and the dendrogram and gives confidence in the distinct nature of the clusters and confirmation of an appropriate choice of cut-level in the dendrogram.

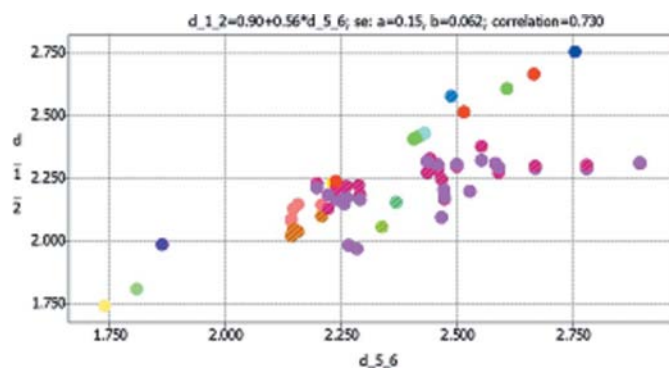
Table 1 shows the clusters with the relevant fragments superimposed. The colours of each fragment are taken from the dendrogram. The clustering is readily rationalized:

(i) In groups A to E the metal atom is either Cu, Zn, Cd or Ta with an open unrestrained ligand connectivity.

(ii) In the F cluster, the central metal is Mn, Ti or Cd, while in the G cluster it is Co, Ni or Fe. There are corresponding differences in other parameters, *e.g.* the O2–C4 distance and the O2–C3–C4 angle. The same pattern is observed between clusters L and M, and the differences relate to the relative positions of C3 to TM6.

(iii) Groups F, G, L and M are all found in structures with multiple metal centres. In groups G and M, the metals form distorted cube-like structures with opposing vertices occupied by O atoms. In groups F and L, the metals are found in a mixture of planar geometries and polyhedra.

(iv) Clusters F and G appear to have almost identical geometries, but the MMDS plot suggests that they are separate. This implies that there are geometric parameters distinguishing the two groups. On examining the relevant geometries, it is apparent that this difference is associated with the types of metal centre present in the groups and is reflected in the metal–oxygen distances. A scatterplot of the O2–metal

**Figure 3**

A scatterplot of the O2···metal (d_{1_2}) distance versus the O2···metal (d_{5_6}) distance. The colours of the plotted points come from the dendrogram.

distance (atoms 1–2 in Fig. 1; distance d_{1_2} in the scatterplot) with the O2–metal (atoms 6–5 in Fig. 1; d_{5_6} in the scatterplot) distance is shown in Fig. 3. It can be seen that group G (brown striped colours) has an O2–metal (d_{1_2}) distance in the range 2.0–2.1 Å and the O2–metal (d_{5_6}) distance in the range 2.1–2.2 Å, whereas in group F (purple) these have a much larger range of 1.9–2.3 and 2.2–2.9 Å.

(v) In three of the four structures in groups A and B (coloured red and yellow), the O and C atoms are found in a 1,4-dioxane ligand. All the structures in these groups have the two transition metals in a *trans* arrangement. In group A, the metal centres are Zn, Cu and Cd, while in group B it is Ti. In Groups C and E, the organic component of the hit fragment is found in the same ligand. In group D, the ligand is 1,2-dimethoxyethane.

2.1. The $\cdots\text{O}-\text{CH}_2-\text{CH}_2-\text{O}\cdots$ ligand: conclusions

A key conclusion of this first analysis is that the clustering method used, based on total geometries, is capable of distinguishing between different types of metal complex purely on the basis of the geometry of the ligand and the relative positions of the O atoms to the metals. In this case the search allowed for a variety of ligands bonding *via* appropriately disposed O atoms. However, it is easy to envisage how searches on more specific ligands could be performed to yield similar trends. For this example, a more restrictive search would result in too few hits for cluster analysis to be valuable. While small datasets are advantageous for illustrative purposes, the key strength of cluster analysis lies in its ability to facilitate the analysis of large data sets, consisting of hundreds or even thousands of structures.

3. The $\text{N}(\text{CH}_2\text{O})_3$ ligand

For this example, the ligand was defined but not the metal centre – the search fragment and the associated numbering are

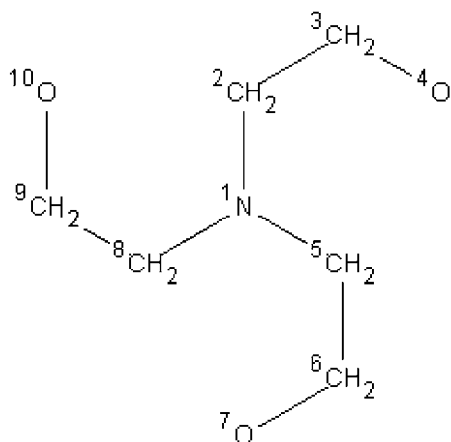


Figure 4
The $\text{N}(\text{CH}_2\text{O})_3$ ligand used in the CSD search and the numbering system employed.

shown in Fig. 4. To provide structural chemistry context, this ligand is commonly used in research into single-molecule magnets. The conditions imposed on the search were: $R < 0.05$; not disordered, no errors, not polymeric, no ions, no powder structures and only metal-organics.

This is quite a complex data set. There were 95 hit structures from the CSD search, yielding 108 fragments. Cluster analysis using *dSNAP* produced a dendrogram which, when cut at a level of 0.700, gave six clusters as shown in Fig. 5. Both the dendrogram, and the associated MMDS plot, show well defined clusters although they are somewhat diffuse when viewed *via* the MMDS display. At this choice of cut level the clusters are readily rationalized in terms of a few geometric parameters, as for the previous example, and this rationalization is summarized in Table 2:

(i) Groups A and C both contain the ligand as μ_3 -coordinated (*i.e.* bonding to a single metal centre through all three O atoms). The exception to this is the structure with refcode YAZZAB in group A, which is μ_2 -coordinated involving an oxygen-bridged di-lead compound: bis(μ_2 -*N,N*-bis(2-oxyethyl)-2-ethanolamine)-di-lead(II). The geometry in this case is such that it is indistinguishable from μ_3 ligands from the viewpoint of the clustering method used here. In group A the O atoms are approximately equidistant to each other, while in group C the O7 \cdots O10 distance is much larger than either the O4 \cdots O7 or O4 \cdots O10 distances.

(ii) Group E contains fragments where the ligand is μ_2 -coordinated. The structures in this cluster contain Co, Cu, Fe, Ni, Re and V metal centres.

(iii) In group B there are two fragments from two structures with a μ_2 -coordinated ligand, and there is one free hydroxyl group. In both structures one of the bound O atoms bridges

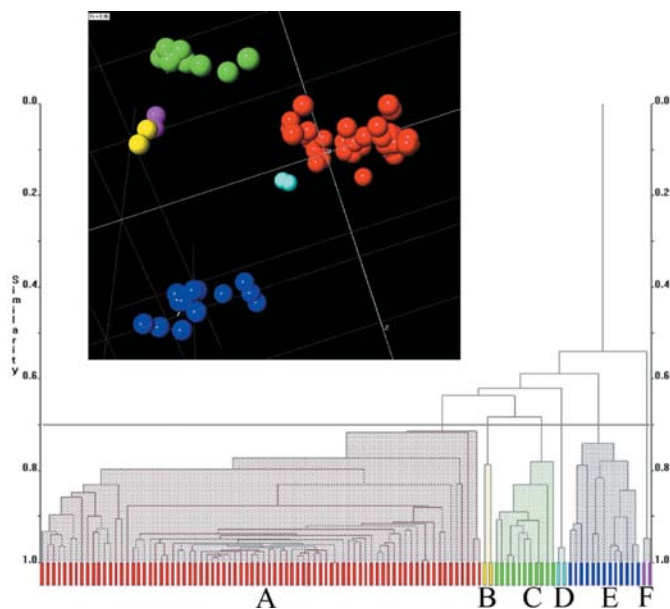

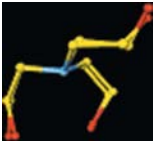
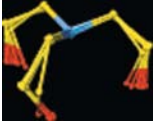
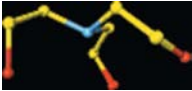
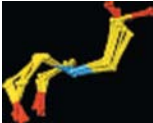
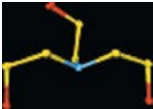


Figure 5
(a) Dendrogram with the MMDS plot as an inset for the $\text{N}(\text{CH}_2\text{O})_3$ ligand. The cut level is set to 0.70 giving seven well defined clusters.

Table 2

A summary of the seven clusters produced by *dSNAP* for the $N(\text{CH}_2\text{O}-)_3$ ligand.

The colours refer to the dendrogram in Fig. 6(a).

Cluster (colour)	Frequency (%)	Geometry of fragment
A (red)	78 (72.2)	
B (yellow)	2 (1.9)	
C (green)	11 (10.1)	
D (cyan)	2 (1.9)	
E (blue)	13 (12.0)	
F (pink)	2 (1.9)	

two metal centres, the other three metal centres; the metal is Mn in all cases.

(iv) Group D contains two fragments from one structure (which therefore contains two ligands) where only O10 is bound to the metal centre (μ_1). The two other O atoms are hydroxyl groups, forming hydrogen bonds to the bound O atom of the other ligand.

(v) Group F contains fragments from two structures containing Hg. The ligand is μ_2 -coordinated through O7 and O10, while O4 is not coordinated to the Hg atom; the O7...O10 distance is larger than for other fragments. In both structures the O atom is part of a phenoxy group.

(vi) Owing to the large number of members of Groups A, C and E, they were individually sub-clustered. It has been shown that there is often additional geometric information accessible from sub-clustering in such situations (Collins, Parkin *et al.*, 2007).

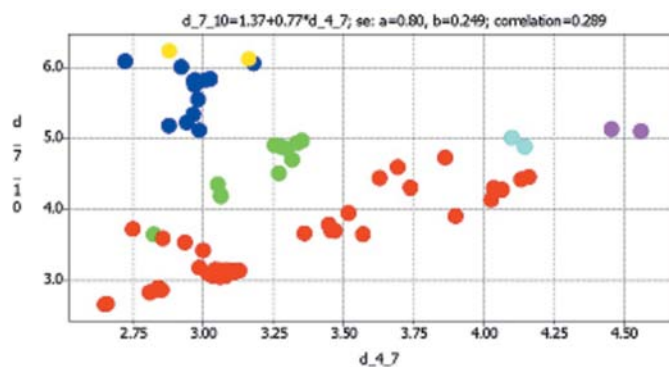
Because the distances between the O atoms are valuable for differentiating and rationalizing the results of the sub-clustering, a scatterplot is shown in Fig. 6 for the complete data set in which O7...O10 [$d(7_{10})$] is plotted against O4...O7 [$d(4_7)$]. It can be seen that most clusters are characterized by the range of these two variables except for group A where there is a very large variation. However, further detail and rationalization of cluster assignments can be revealed by re-

running the cluster analysis on selected groups. This is readily carried out in *dSNAP*.

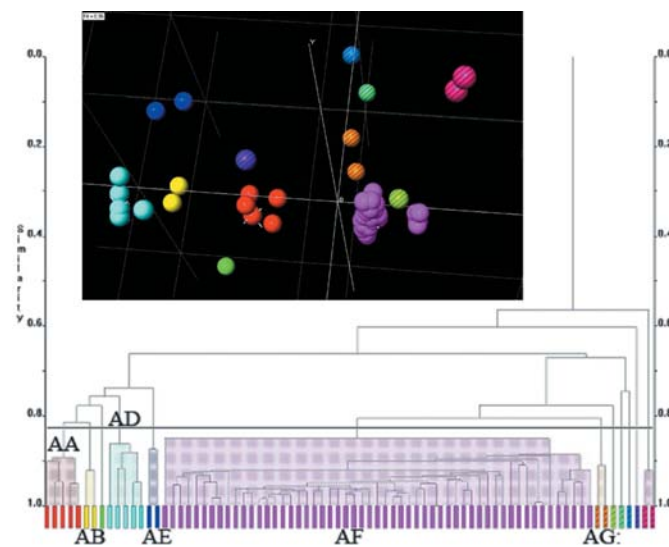
3.1. Sub-clustering group A

The 78 hit fragments in cluster A in Fig. 5 were re-clustered. The resulting dendrogram formed 12 clusters at a cut level of 0.826, as shown in Fig. 7. Both the dendrogram and the associated MMDS plot show clearly defined groups of structures. Fig. 8 shows the scatterplot of the O7...O10 versus O4...O7 distances. As can be seen in the scatterplot, O7...O10 and O4...O7 distances take approximately equal values in groups AC, AD, AF and AL, but not in other groups.

The largest group is AF, where none of the O atoms are bridging, and all O atoms are coordinated only to a single metal. In the MMDS plot a subgroup is apparent, which corresponds to cases where the central metal atom is silicon,

**Figure 6**

A scatterplot of O7...O10 [$d(7_{10})$] against O4...O7 [$d(4_7)$] for the $N(\text{CH}_2\text{O}-)_3$ ligand. The colours of the plotted points are taken from the dendrogram in Fig. 5. Groups B–F are well characterized by this distance although there is overlap between group B (yellow) and E (dark blue).

**Figure 7**

The dendrogram resulting from re-clustering group A in Fig. 6, and the corresponding MMDS plot as an inset.

although this difference is not very pronounced in the dendrogram. This is a further indicator of how valuable it can be to use more than one clustering method.

Group AB contains two hit fragments from refcode YAZZAB which, as noted above, are in fact μ_2 -coordinated. O10 is a hydroxyl group in each case, although this is not apparent from the geometry of the ligand fragment. The metal centre (Pb) is not equidistant from all three O atoms as is the case for the structures in group AF.

Groups AD and AE consists of structures where the ligand is larger than the search fragment used, and is coordinated to Na or K.

Group AL has the most noticeably different geometry, being almost planar at N1, and where the ligand is coordinated to Si.

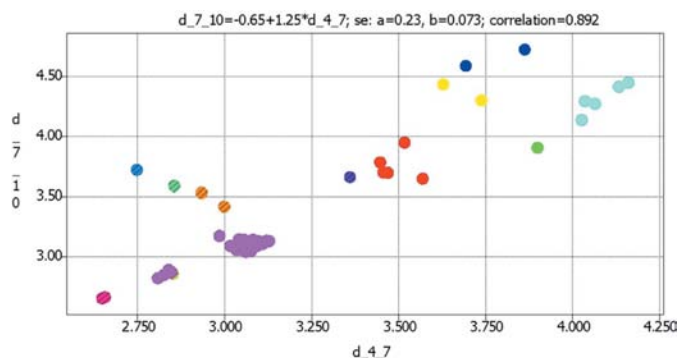


Figure 8
A scatterplot of O7...O10 [$d(7_{10})$] versus O4...O7 [$d(4_7)$] distances for the structures clustered in Fig. 7. As in the previous scatterplots, the distances characterize the groups effectively.

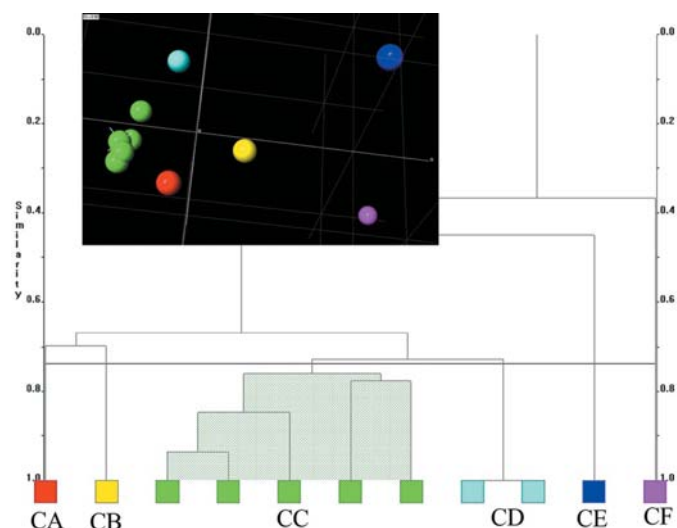


Figure 9
The resulting dendrogram of re-clustering the 11 hit fragments that formed group C in Fig. 5 with the MMDS plot as an inset. There are six clusters, four of which are singletons, at a cut level of 0.740. The two representations of the data are in agreement.

3.2. Sub-clustering group C

The 11 hit fragments forming group C in Fig. 5 were re-clustered. The resulting dendrogram and MMDS plot are shown in Fig. 9. There are six clusters at a cut level of 0.740.

The cluster assignment shows a large dependency on the bite size of the ligand, as indicated by the O...O distances, which are shown as a scatterplot in Fig. 10(a). Groups CB (yellow) and CF (magenta) have similar values for the O...O distances but differ in the geometry at C3. In the largest single group at this cut level, group CC, the metal is Cu for all fragments. The two fragments in group CD (cyan) are from two different determinations of the same structure; the metal is Sb. This can be inferred from the tie bar between them, which is essentially at a cut-level of 1.0 (Fig. 9). The other groups contain Nd (CA, red), Ho (CB), Ti (CE) and Zn (CF) as their metal centre. In Fig. 10(b) all the fragments are shown superimposed with colours taken from the dendrogram. It can be seen that quite subtle differences have been resolved automatically by *dSNAP*.

3.3. Sub-clustering group E

The 13 hit fragments that formed group E in Fig. 6 were re-clustered to give six clusters at a cut level of 0.645, as shown in Fig. 11.

There are two main areas of difference in the fragment geometries: there are two observed orientations at C6 and

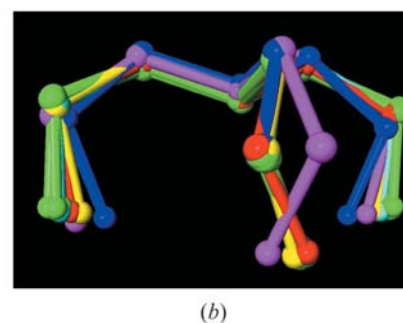
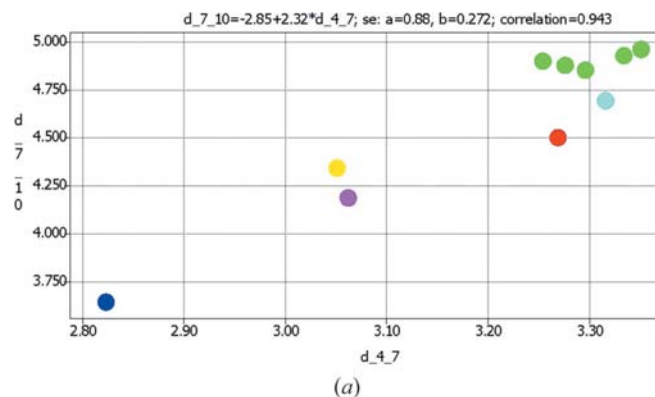


Figure 10
(a) A scatterplot of O7...O10 [$d(7_{10})$] versus O4...O7 [$d(4_7)$] distances for the structures clustered in Fig. 9. As in the previous scatterplots, the distances characterize the groups effectively. For example, the green group with five clusters has all O7...O10 distances in the range 4.8–5.0 Å, and an O4...O7 distance in the range 3.20–3.37 Å. (b) The fragments overlaid – the colours come from the dendrogram in Fig. 9. It can be seen that quite subtle differences have been resolved.

three observed geometries arising from the torsion angle around the C8—C9 bond, giving six geometric combinations, and hence six possible clusters. The fragments are overlaid in Fig. 12(a). These geometric differences can be rationalized by comparing the N1...C6 distance (reflecting the two orientations at C6) and the N1...O10 distance (which is related to the N1—C8—C9—O10 torsion angle) in the form of a scatterplot, and this is shown in Fig. 12(b). Each cluster identified in the dendrogram has a unique range of N1...C6 and N1...O10 distances.

3.4. Conclusions concerning the coordination geometry of the N(CH₂O—)₃ ligand fragment

This example shows how information about the coordination geometry around the metal can be gleaned without the metal being included in the cluster analysis. The advantage of this is that the CSD search is not then limited to expected coordination geometries or will not contain many duplicate hits of the same ligand coordinated to a different metal (this would cause problems with this ligand if one or more of the O atoms is bridging). This duplicate-hits issue arose with the first example, but was easy to identify because the data set was much smaller; it is much more of a problem as the size of the data set increases.

Interestingly, although with sub-clustering it was possible to identify a large group of atoms coordinated to a single metal ion (group AE), in actuality the geometry of the ligand does not necessarily vary much according to whether the ligand is bridging or not.

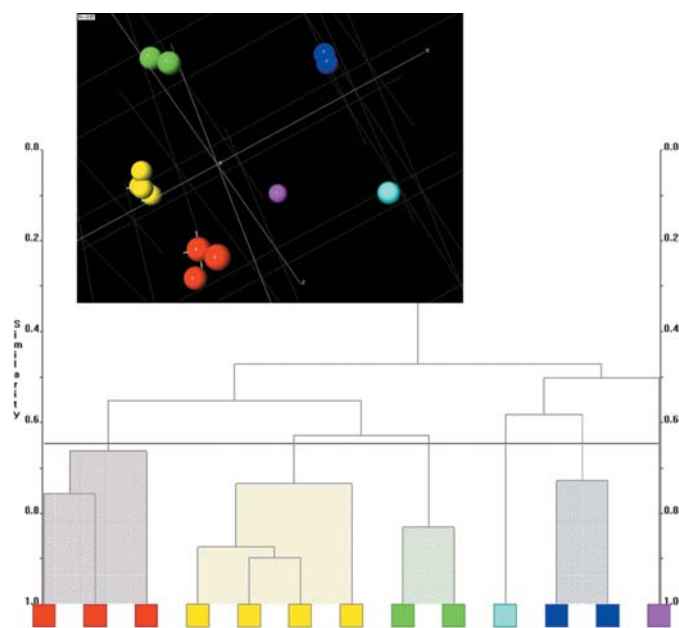


Figure 11

The results of re-clustering the 13 hit fragments that formed group C in Fig. 5. (a) The dendrogram and (b) the MMDS plot. There are six clusters, two of which are singletons, at a cut level of 0.645. Both representations of the data are in agreement.

4. Conclusions

Once again the *dSNAP* methodology and the associated software has shown how hits from CSD searches can be rapidly classified and rationalized using the methods of cluster analysis. In cases where there are small numbers of hits visual inspection may suffice, but this technique provides (a) objectivity, (b) an ability to process 1000s of hits, and (c) suitable visualization tools to allow rapid and intuitive assessment of the clustered geometries. In the first example, the mode of coordination of the ligand component to the metal was defined. In the second system, a common ligand was defined, but no metal component was specified. The clustering has proved capable of distinguishing in a wholly automatic, objective way between different types of metal complex purely on the basis of the geometry of the ligand and the relative positions of the O atoms to the metals, illustrating the potential of *dSNAP* analysis in providing important rationalization of geometries adopted in systems of interest in the

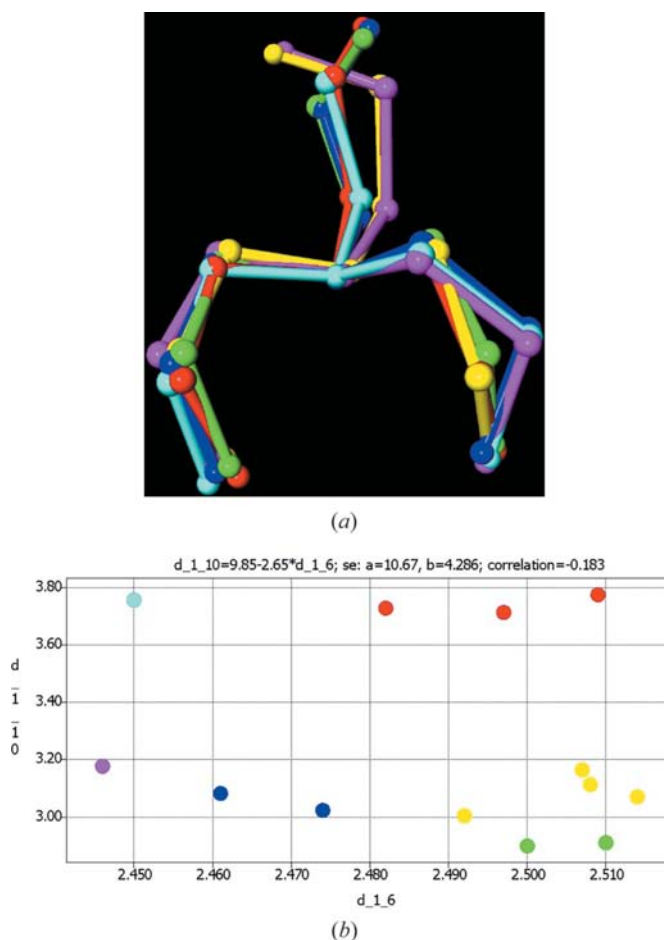


Figure 12

(a) The fragments resulting from re-clustering fragment C overlaid. The colours come from the dendrogram. (b) A scatterplot of O7...O10 [$d(7_{10})$] versus O4...O7 [$d(4_7)$] distances for the structures clustered in Fig. 11. As in the previous scatterplots, the distances characterize the groups effectively. For example, the green group with five clusters has all O7...O10 distances in the range 4.8–5.0 Å, and an O4...O7 distance in the range 3.20–3.37 Å. The colours also come from the dendrogram.

development of many types of metal-organic materials. These examples also highlight the importance of how the search is defined; this helps define what information is available, and the focus of the results. The *dSNAP* program is available free of charge from http://www.chem.gla.ac.uk/snap/dSNAP_beta.html.

Anna Collins was supported by the University of Glasgow.

References

- Allen, F. H. (2002). *Acta Cryst.* **B58**, 380–388.
- Barr, G., Dong, W., Gilmore, C. J., Parkin, A. & Wilson, C. C. (2005). *J. Appl. Cryst.* **38**, 833–841.
- Brammer, L. (2004). *Chem. Soc. Rev.* **33**, 476–489.
- Bruno, I. J., Cole, J. C., Kessler, M., Luo, J., Motherwell, W. D. S., Purkis, L. H., Smith, B. R., Taylor, R., Cooper, R. I., Harris, S. E. & Orpen, A. G. (2004). *J. Chem. Inf. Comput. Sci.* **44**, 2133–2144.
- Bruno, I. J., Cole, J. C., Lommerse, J. P. M., Rowland, R., Taylor, R. & Verdonk, M. L. (1997). *J. Comput. Aided Mol. Des.* **11**, 525–537.
- Cambridge Crystallographic Data Centre (1994). *Vista*. Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.
- Collins, A., Barr, G., Dong, W., Gilmore, C. J., Middlemiss, D. S., Parkin, A. & Wilson, C. C. (2007). *Acta Cryst.* **B63**, 469–476.
- Collins, A., Parkin, A., Barr, G., Dong, W., Gilmore, C. J. & Wilson, C. C. (2007). *CrystEngComm*, **9**, 245–253.
- Collins, A., Parkin, A., Barr, G., Dong, W., Gilmore, C. J. & Wilson, C. C. (2008). *Acta Cryst.* **B64**, 57–65.
- Collins, A., Wilson, C. C. & Gilmore, C. J. (2009). *CrystEngComm*, doi: 10.1039/b914683k.
- Dance, I. (2003). *CrystEngComm*, **5**, 208–221.
- Fey, N., Harris, S. E., Harvey, J. N. & Orpen, A. G. (2006). *J. Chem. Inf. Model.* **46**, 912–929.
- Fey, N., Tshipis, A. C., Harris, S. E., Harvey, J. N., Orpen, A. G. & Mansson, R. A. (2006). *Chem. Eur. J.* **12**, 291–302.
- Harris, S. E., Orpen, A. G., Bruno, I. J. & Taylor, R. (2005). *J. Chem. Inf. Model.* **45**, 1727–1748.
- Hocking, R. K. & Hambley, T. W. (2005). *J. Chem. Soc. Dalton Trans.* **5**, 969–978.
- Macrae, C. F., Edgington, P. R., McCabe, P., Pidcock, E., Shields, G. P., Taylor, R., Towler, M. & van de Streek, J. (2006). *J. Appl. Cryst.* **39**, 453–457.
- Minguez Espallargas, G., Brammer, L. & Sherwood, P. (2006). *Angew. Chem. Int. Ed.* pp. 435–440.
- Orpen, A. G. (2002). *Acta Cryst.* **B58**, 398–406.
- Parkin, A. (2008). *Crystallogr. Rev.* **14**, 117–142.
- Parkin, A., Barr, G., Collins, A., Dong, W., Gilmore, C. J., Tasker, P. A. & Wilson, C. C. (2007). *Acta Cryst.* **B63**, 612–620.
- Parkin, A., Barr, G., Dong, W., Gilmore, C. J., Jayatilaka, D., McKinnon, J. J. Spackman, M. A. & Wilson, C. C. (2007). *CrystEngComm*, doi: 10.1039/b704177b.
- Parkin, A., Collins, A., Gilmore, C. J. & Wilson, C. C. (2008). *Acta Cryst.* **B64**, 66–71.