Mapping the Geometry of Metal Three-Coordination Using Crystal Structure Data: Reaction Pathway for Ligand Addition to Linear Hg^{II} Species

by **Frank H. Allen***[a]), **Raju Mondal**[b]), **Nigel A. Pitchford**[b]) and **Judith A. K. Howard***[b])

[a]) Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK
[b]) Department of Chemistry, University of Durham, South Road, Durham DH1 3LE, UK (Fax: +44-1223-336033 (FHA), e-mail: allen@ccdc.cam.ac.uk; j.a.k.howard@durham.ac.uk)

Dedicated to Professor *Jack D. Dunitz* on the occasion of his 80th birthday

Symmetry-modified principal-component analysis has been used to visualise the geometrical distortions of three-coordinate metal centres observed in crystal structures retrieved from the *Cambridge Structural Database*. It is shown that compounds of Cu, Ag, Hg, Zn, and Au dominate the dataset, and exhibit distortions away from the trigonal planar archetype towards *a*) T-shaped and *b*) Y-shaped geometries. A small number of compounds, principally of Cu and Ag, also show distortions towards trigonal pyramidal geometries. The interconversions from Y-shaped geometries, through the trigonal planar form to T-shaped geometries are clearly mapped by the PC analysis. For Hg^{II} complexes, it is possible to interpret the transition from T-shaped geometries to the trigonal planar form in terms of a reaction pathway for ligand addition to linear $L_1-Hg^{II}-L_2$ species.

Introduction. – Metal centres that exhibit coordination number three are relatively rare [1], occurring mainly for the elements Cu, Ag, Au, Zn and Hg. The archetypal coordination geometries are trigonal planar (D_{3h}) and trigonal pyramidal (C_{3v}), with the former dominant. However, even a cursory inspection of crystal structures containing three-coordinate metal centres reveals the existence of significant angular distortions from ideal symmetry. Thus, in planar examples (*Fig. 1*) there exist distortions from the trigonal archetype (*Fig. 1,a*) towards a T-shaped geometry with $\theta_1 \rightarrow 180^\circ$, and θ_2, $\theta_3 \rightarrow 90^\circ$ (*Fig. 1,b*), and the converse distortion to a Y-shaped geometry having $\theta_1 \rightarrow$ *ca.* 60° and θ_2, $\theta_3 \rightarrow$ *ca.* 150° (*Fig. 1,c*).

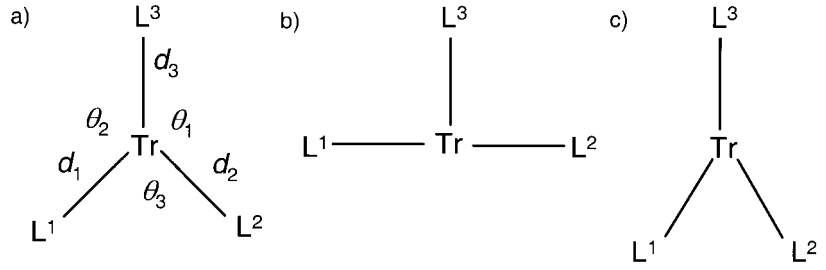

Fig. 1. a) *Trigonal planar* D_{3h} *coordination geometry, and the bond lengths and angle parameters used in this analysis.* b) *T-Shaped deformation from* D_{3h} *symmetry.* c) *Y-Shaped deformation from* D_{3h} *symmetry.*

In this paper, we begin by examining the coordination geometry of all three-coordinate species TrL_3, where Tr is any transition element, and L is a ligand connected

to Tr *via* bonds to the elements C, N, O, F, Cl, Br, I, Si, As, Te, S, P, Se. We have used the *Cambridge Structural Database* (*CSD*) [2] for this study, and accepted the definitions of connectivity stored therein, which reflect those of the authors of papers recorded in the database. We then map the angular distortions present in the retrieved data set by principal-component analysis [3][4], and also record statistics for the major Tr-contributors.

In the second part of the analysis, we examine the differences in the geometric distortions exhibited by some individual TrL_3 systems for Tr = Cu, Ag, Au, Zn and Hg, and, using the principle of structure correlation [5–8], we interpret the distortions exhibited by HgL_3 species in terms of ligand addition to the more common and approximately linear HgL_2 system.

Database Search and Data Analysis Methodology. – Version 5.24 of the CSD (November 2002 release containing 272,066 entries) has been used in this study. Searches and geometric data retrieval have been carried out with the CSD System program Quest3D [9], with Vista Version 1.2 [10] and Mercury Version 1.1 [11] being used for data analysis and structure visualisation, respectively. Analysis of geometric distortions in the data set has been made in terms of the intraligand valence angles θ_1, θ_2, θ_3 indicated in *Fig. 1,a*, and calculated for each TrL_3 substructure located in the CSD. For each substructure, we also calculated (*Fig. 1,a*) the bonded distances d_1, d_2, d_3, and the distance, OOP, of the metal from the plane of the three ligating atoms. In each case, the covalent radii, $r_{cov}(n)$, of the central metal and each of the ligating atoms were also included in the table of retrieved data, where n is the node number of each atom as shown in *Fig. 1,a*. These are the standard radii used routinely by CSD software in assembling the bonded connectivity of molecules in all crystal structures stored in the CSD.

All searches reported in this paper were subject to secondary search criteria, which ensured that retrieved CSD entries were *a*) not subject to disorder, *b*) not polymeric (catena) structures, and *c*) were error-free after structure validation. Various crystallographic *R*-factor criteria were also applied at different stages of the study, as reported at appropriate points in the *Results and Discussion* section.

Because the substructure of *Fig. 1,a*, is topologically symmetric, it is not possible for the CSD search software to superimpose the substructural query onto successive CSD structures by a unique atomic enumeration. Thus, it is not possible, *e.g.*, to ensure that the search results always have atom 1 opposite to the largest θ angle, atom 2 opposite to the next largest, and atom 3 opposite to the smallest angle. If it were possible, such an action would cause the software to generate θ_1, θ_2, θ_3 such that they fell within a single asymmetric unit of valence angle space. Instead, we have used the 'permute' function in Quest3D to include geometric data for all six permutational isomers of each substructure (*Fig. 1,a*, with the atom numbers adopting all six possible permutations) in the resulting data matrix. The result of this operation is to fill valence angle space completely. In practice, this topological symmetry-modification technique is particularly valuable in visualising and interpreting the results of principal-component analysis, as described more fully in [4].

Principal-component analysis (PCA) itself performs an eigenanalysis of the matrix of three valence angles for each of the occurrences of the substructure located in the database, with the matrix being expanded if necessary by the topological symmetry of

the substructure. In the text and tables that follow, we use N_{frag} to denote the unique number of fragments located in the CSD; however, the symmetry-expansion will give rise to $6 \cdot N_{frag}$ entries in the data matrix used by the PCA algorithm incorporated into the Vista program. In PCA is analysed the variance in a p-dimensional data set (here $p = 3$), and re-expresses the original data with respect to q mutually orthogonal principal component axes (PCs), which are linear combinations of the original variables, and which together account for a very large proportion (*ca.* 99%) of the variance. If $q < p$, the technique has achieved a reduction in the dimensionality of the original dataset, reducing the number of PCx *vs.* PCy scatterplots required to map and visualise the data.

Results and Discussion. – *Statistics of Three-Coordination* (*Search 1*). An initial CSD search was run to locate all TrL_3 species, with no *R*-factor restriction. Results are presented in *Table 1*, in terms of N_{ent}, the number of CSD entries located, and N_{frag}, the number of unique TrL_3 substructures located within those entries. *Table 1* shows that 90% of three-coordinate species are accounted for by the five elements Cu, Ag, Hg, Zn and Au. The remaining 10% comprise complexes of a further 17 elements.

Table 1. *Statistics of TrL_3 Substructures in Crystal Structures Recorded in the CSD.* N_{ent} is the number of CSD entries, N_{frag} is the number of discrete substructures in those entries.

Tr-Element	N_{ent}	N_{frag}	% Total N_{frag}
Cu	453	793	48.0
Ag	216	297	18.0
Hg	148	204	12.3
Zn	80	109	6.6
Au	60	80	4.8
Fe	29	48	2.9
Mn	20	31	1.9
Pt	13	15	0.9
Cd	12	20	1.2
Ni	12	15	0.9
Other	33	40	2.4
Totals	1076	1652	100.0

PCA Results – All TrL_3 Species (*Search 2*). *Search 2* used *Search 1* criteria, but with the added restriction that $R < 0.070$. This yielded 1032 TrL_3 substructures from 611 CSD entries, giving 6192 parameter sets in the final data matrix after symmetry expansion.

Table 2. *Results of PC Analyses for Various Datasets in Terms of the Percentage Variance* (var %) *Accounted for by Each PC, Together with the Numbers of CSD Entries* (N_{ent}) *and Symmetry-Expanded Substructures* (SEN_{frag}) *Located by Each Search within the* R*-Factor Limit Shown*

Dataset	*R*-Factor limit	N_{ent}	SEN_{frag}	Var % (PC1)	Var % (PC2)	Var % (PC3)
TrL_3	0.07	611	6192	49.10	49.10	1.80
CuL_3	0.10	445	5970	49.40	49.40	1.20
AgL_3	0.10	214	2472	48.10	48.10	3.80
HgL_3	0.10	139	1308	49.97	49.97	0.05
ZnL_3	0.10	77	864	49.95	49.95	0.10
AuL_3	0.10	60	588	49.96	49.96	0.09

PC Analysis (*Table 2*) based on valence angles yielded a degenerate pair of PCs (PC1 and PC2), each accounting for 49.1% of variance in the dataset, with the remainder (1.8%) being accounted for by PC3. Scatterplots of PC1 *vs.* PC2 and PC1 *vs.* PC3 are

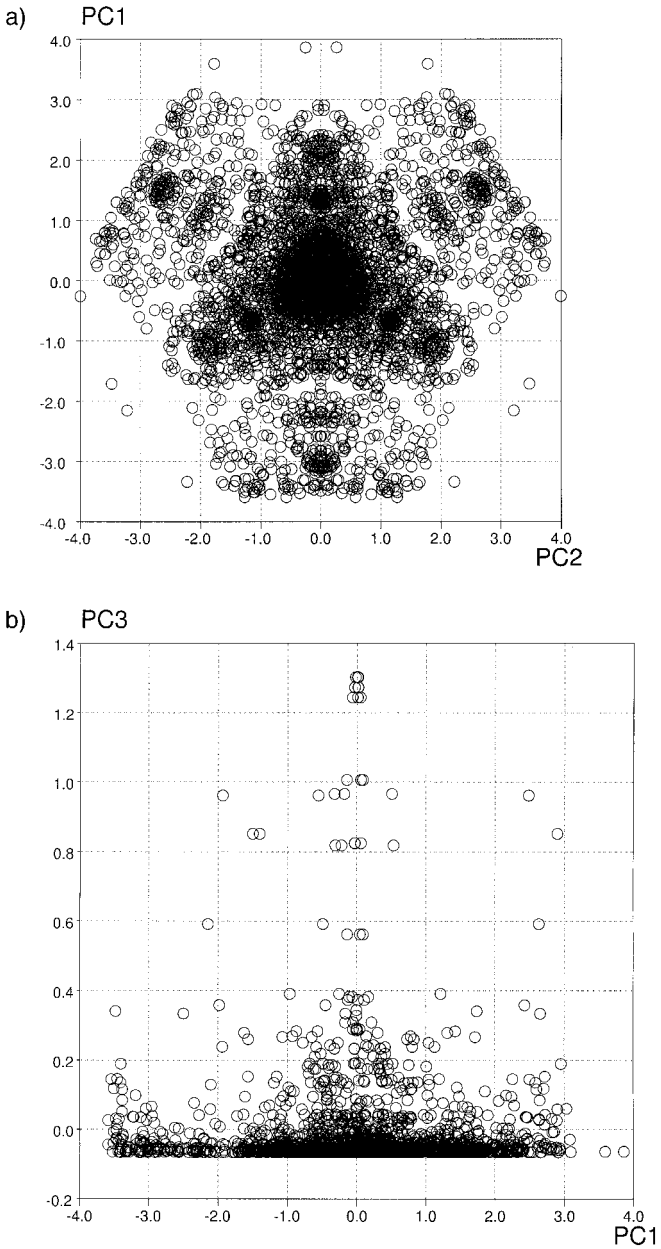

Fig. 2. a) *PC1* vs. *PC2 for all TrL_3 species located in* Search 2 (see text), b) *PC1* vs. *PC3 for all TrL_3 species located in* Search 2 (see text)

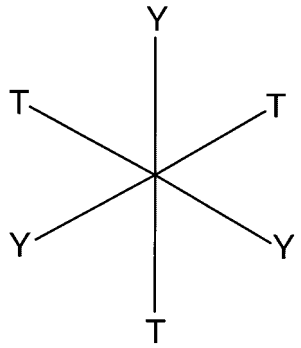

Fig. 3. *Schematic diagram indicating the deformation directions revealed in the PC1* vs. *PC2 scatterplots of* Fig. 2,a, *and* Fig. 4

shown in *Fig. 2*,*a* and *b*, respectively. Inspection of these plots with respect to the original data show that PC1 *vs.* PC2 map deformations away from trigonal planar (D_{3h}) geometry, represented by the large cluster of structures at PC1, PC2 = 0.0, 0.0, towards Y- and T-shaped geometries. The deformation directions in the PC1, PC2 plot are shown diagrammatically in *Fig. 3*, and these directions also apply consistently to the other PC1 *vs.* PC2 plots discussed later and shown in *Fig. 4*.

PC3 Maps the out-of-plane deformations of the metal atom towards pyramidal geometries. A subsequent analysis showed that the metal deviated from the ligand plane by $d > 0.2$ Å in 153 (14.8%) cases, and by $d > 0.4$ Å in 58 (5.6%) cases in the 1032 retrieved substructures. As in other PC analyses [4][12], the PCs, which are linear combinations of the basic variables, are directly correlated to the group-theoretic symmetry deformation coordinates (SDCs), which map deviations from D_{3h} symmetry [12] *via:*

$$\text{SDC1 } S_{4a}(E'): 1/\sqrt{6\ (2\theta_1 - \theta_2 - \theta_3)}\ \text{PC1: } 0.816\ \theta_1 - 0.408\ \theta_2 - 0.408\ \theta_3$$

$$\text{SDC2 } S_{4b}(E'): 1/\sqrt{2\ (\theta_2 - \theta_3)}\ \text{PC2: } 0.707\ \theta_2 - 0.707\ \theta_3$$

$$\text{SDC3 } S_5(A2''): 1/\sqrt{3\ (\theta_1 + \theta_2 + \theta_3)}\ \text{PC3: } 0.577\ \theta_1 + 0.577\ \theta_2 + 0.577\ \theta_3$$

Scatterplots of the SDCs (not shown) have exactly the same form as the PC plots illustrated in *Fig. 2*, but with the PC1 and PC2 axes rotated by 60° and 30°, respectively from SDC1 and SDC2.

PCA Results for Individual TrL_3 Species. TrL_3 Searches were carried out for the five most common elements in *Table 1*, *viz:* Cu, Ag, Hg, Zn, Au, with the criteria for *Search 1*, but with the restriction that $R < 0.10$. Results for these searches, including the PCA variance scores, are given in *Table 2*. One noticeable feature of the PC variances is the higher value of PC3 for Cu and Ag than for other elements. Thus, most of the out-of-plane displacements towards pyramidal geometries observed for the complete TrL_3 data set in *Fig. 2*,*b*, arise from complexes of these two elements.

The PC1 and PC2 plots for Cu, Ag, Hg, Zn, Au are compared in *Fig. 4*,*a – e*, with the Y- and T-shaped deformations arranged consistently to conform to the directions

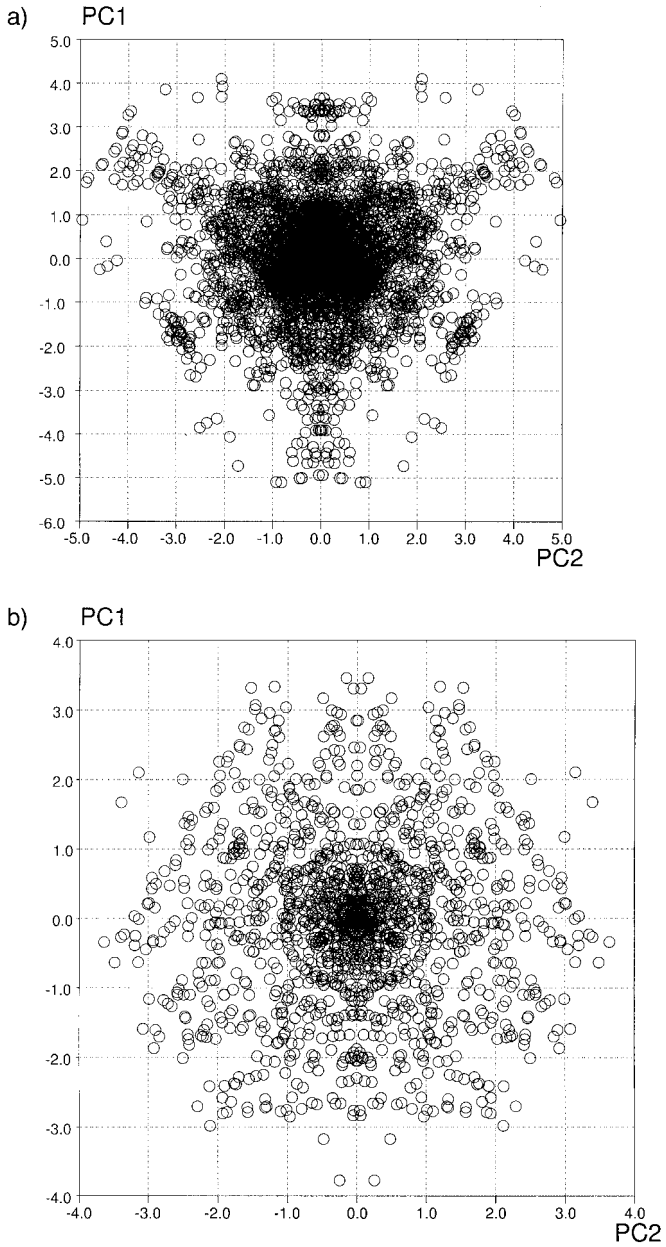

Fig. 4. *PC1* vs. *PC2 plots for the most common three-coordinate elements:* a) *Cu*, b) *Ag*, c) *Hg*, d) *Zn and* e) *Au*. The deformation directions are shown diagrammatically in *Fig. 3*.

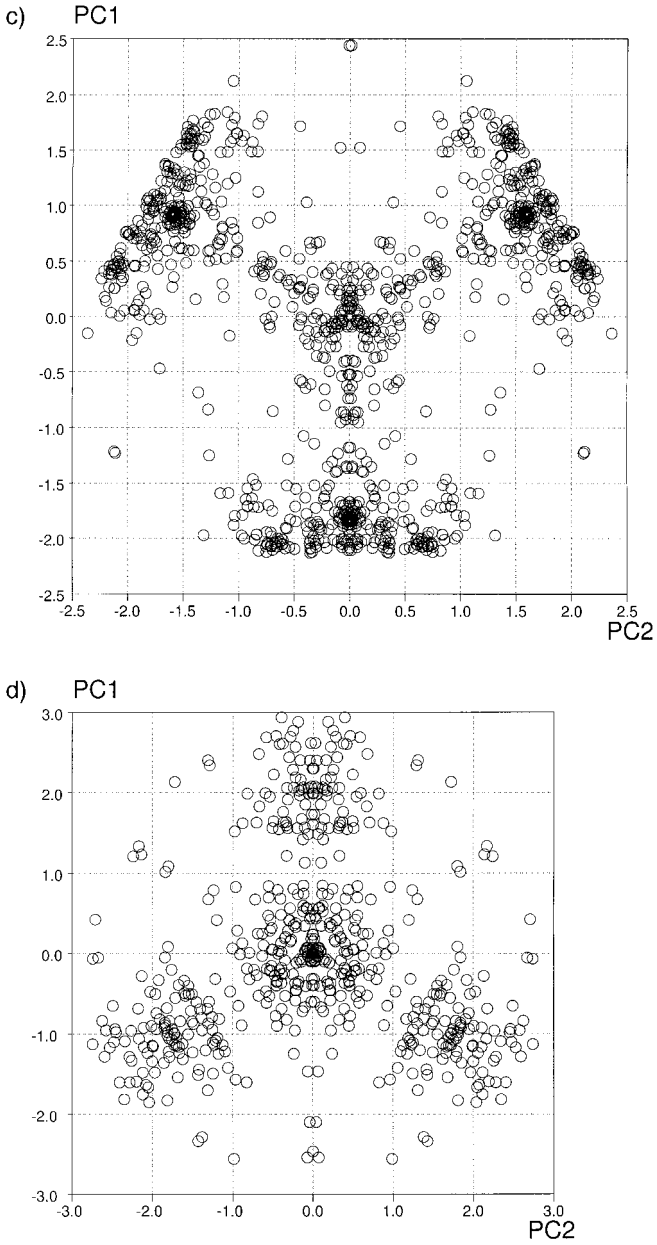
c)
PC1
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5
PC2
d)
PC1
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
-3.0 -2.0 -1.0 0.0 1.0 2.0 3.0
PC2

Fig. 4 (cont.)

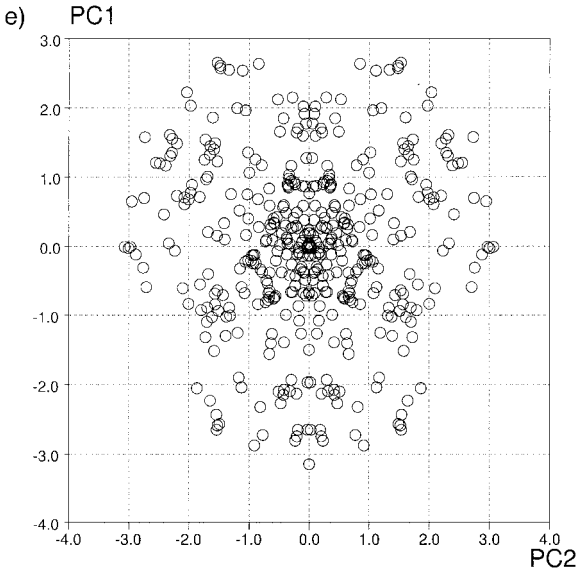

Fig. 4 (cont.)

illustrated in *Fig. 3*. The plots for Cu, Ag, Au are similar in form to *Fig. 2,a*, for the complete TrL_3 dataset. All three show that complexes approximating trigonal geometry are predominant, with small numbers of substructures showing deformations towards T- and Y-shaped geometries. The plots for Zn (*Fig. 4,d*) and Hg (*Fig. 4,c*) are, however, rather different. The former (Zn) shows two major features, the trigonal core and a discrete cluster of Y-shaped geometries, with almost no T-shaped examples. Examination of the Y-cluster shows that virtually all of the contributors arise from Zn^{II} substructures of the type shown in *Fig. 5* where, of necessity, one of the angles at Zn is forced towards 90° by the four-membered ring formation *via* X = O, N, Cl, *etc.* Indeed, examination of the more-general TrL_3 data set shows that the formation of four-membered rings involving Tr is by far the most common cause of Y-shaped geometries.

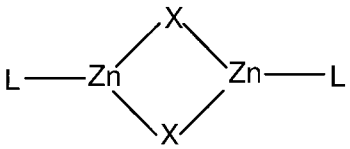

Fig. 5. *Substructure causing Y-shaped deformations of three-coordinate Zn species, the most common elements X are O, N, and Cl*

The PC1 *vs.* PC2 plot for Hg (*Fig. 4,c*) is the only one for which T-shaped geometries predominate, and with a well-defined pathway leading from the T-cluster to the relatively small central trigonal core cluster at the origin of PC1 *vs.* PC2 space. In light of the seminal studies by *Dunitz* and co-workers (*cf.* [5–8] and refs. cit. therein), it is reasonable to assume that this transition from T-shaped to trigonal Hg^{II} complexes

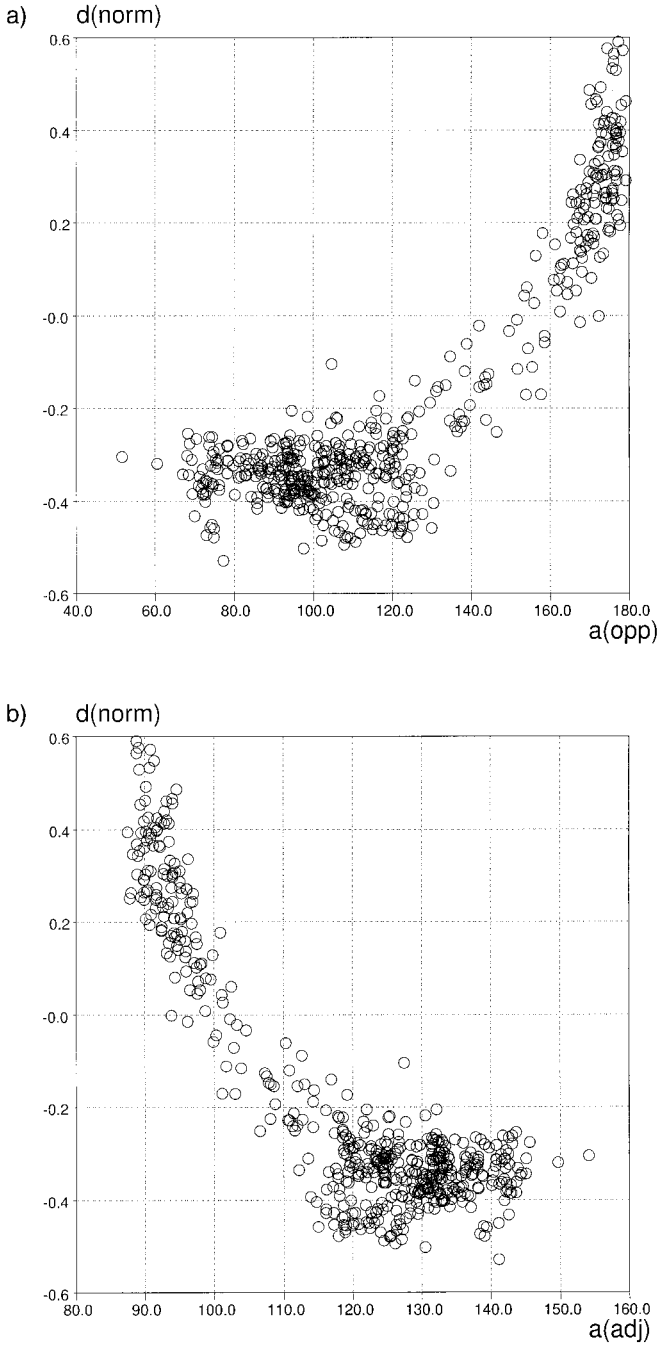

Fig. 6. a) *Plot of* d*(norm)* vs. a*(opp)*, b) *plot of* d*(norm)* vs. a*(adj)* (see text for parameter definitions)

maps the low-energy pathway for addition of L_3 (*Fig. 1*) to the linear $L_1-Hg^{II}-L_2$ system, remembering that linear two-coordination represents the most common geometry for the bulky Hg^{II}. Such an interpretation of *Fig. 4*,*c*, is in accord with the structure-correlation principle, *i.e.*, that the static distortions exhibited by a specific molecular substructure (fragment) in a wide variety of crystalline environments can be assumed to map the distortions that the fragment would undergo along a reaction pathway. The relationship between geometrical distortions and potential energy is then made as follows [5]: '*If a correlation can be found between two or more independent parameters describing the structure of a given (sub)structural fragment in a variety of environments, then the correlation function maps a minimum-energy path in the corresponding parameter space*'.

To examine this hypothesis further for the case of ligand addition to Hg^{II}, we have examined the coordination in terms of the normalised Hg–L bond lengths:

$$d(\text{norm}) = d(\text{Hg}-\text{L}) - r_{\text{cov}}(\text{Hg}) - r_{\text{cov}}(\text{L})$$

where $r_{\text{cov}}(\text{Hg})$ and $r_{\text{cov}}(\text{L})$ are the standard covalent radii of Hg (1.70 Å) and L, respectively, used in assembling CSD entries. The quantity d(norm) is calculated for each hit substructure during permutational expansion, so that d(norm) is calculated for each of the bond lengths d_1, d_2, d_3 in *Fig. 1*.

Fig. 6,*a*, shows a plot of d(norm) *vs.* a(opp), where a(opp) is the valence angle 'opposite' the 'incoming ligand', *i.e.*, it is θ_n for normalised d_n ($n = 1-3$) in the nomenclature of *Fig. 1*,*a*. Our hypothesis is that as ligand L_3 approaches a linear $L_1-Hg^{II}-L_2$ system, with a(opp) close to 180°, it will have its longest d(norm) value. However, as the bonding interaction $Hg^{II}-L_3$ strengthens d(norm) decreases, and a(opp) will tend towards the 120° expected for a trigonal planar species. This behaviour is exactly reproduced in *Fig. 6*,*a*. An equivalent picture (*Fig. 6*,*b*) can, of course, be obtained by plotting d(norm) *vs.* a(adj), the mean of the angles subtended at Hg^{II} by the incoming L_3 with the existing $L_1-Hg^{II}-L_2$ system. This plot visualises the increase in a(adj) from an initial approach direction of 90° to the existing linear system, towards the 120° for the trigonal planar species which is the end point of the ligand addition reaction.

Conclusions. – This short paper has again shown the value of symmetry-modified principal-component analysis in visualising the deformations in metal-coordination geometry exhibited in crystal-structure determinations. While such deformations are well-known from examinations of limited numbers of structures, CSD analysis coupled with PCA can immediately provided a mapping of all available structures, which reveals the principal clusters of like geometries, and any interconversion pathways that connect these clusters. For the case of three-coordination, these techniques have identified the major deformations to Y-shaped and T-shaped geometries, together with a mapping of interconversions from trigonal-planar to pyramidal geometries. In this application, PCA has not achieved a dimension reduction in the dataset, but it has provided a mapping of the data that is more readily interpretable in structural terms than any mappings based on individual valence angles. It is also shown that the PCA mappings are equivalent to those generated by means of group-theoretic symmetry

deformation coordinates. In the case of Hg^{II}, the mappings have indicated a smooth interconversion from T-shaped to trigonal species, which has been interpreted here in terms of a reaction pathway for ligand addition to linear $L_1-Hg^{II}-L_2$ systems.

The authors would like to acknowledge the major contribution made by Prof. *Jack D. Dunitz* to *CSD*-based research and to the scientific development of the *Cambridge Crystallographic Data Centre*. This research includes the development, with his co-workers, of the structure correlation principle on which this work is based. The authors would like to thank the *Engineering and Physical Sciences Research Council*, UK, for the award of a Senior Research Fellowship (J. A. K. H.), and the CCDC and Durham University for financial support (R. M.).

REFERENCES

[1] N. N. Greenwood, A. Earnshaw, 'Chemistry of the Elements', Pergamon Press, Oxford, 1993.
[2] F. H. Allen, *Acta Crystallogr., Sect. B* **2002**, *58*, 380.
[3] C. Chatfield, A. J. Collins, 'Introduction to Multivariate Analysis', Chapman & Hall, London, 1980.
[4] F. H. Allen, M. J. Doyle, T. P. E. Auf der Heyde, *Acta Crystallogr., Sect. B* **1991**, *47*, 412.
[5] P. Murray-Rust, H. B. Bürgi, J. D. Dunitz, *J. Am. Chem. Soc.* **1975**, *97*, 921.
[6] J. D. Dunitz, 'X-Ray Analysis and the Structure of Organic Molecules', Cornell University Press, Ithaca New York, 1979.
[7] H. B. Bürgi, J. D. Dunitz, *Acc. Chem. Res.* **1986**, *16*, 153.
[8] H. B. Bürgi, J. D. Dunitz, 'Structure Correlation', VCH, Weinheim, 1994.
[9] F. H. Allen, J. E. Davies, J. J. Galloy, O. Johnson, O. Kennard, C. F. Macrae, E. M. Mitchell, G. F. Mitchell, J. M. Smith, D. G. Watson, *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 187.
[10] Vista – A program for the analysis and display of data retrieved from the CSD, Version 1.2, Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK.
[11] I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson, R. Taylor, *Acta Crystallogr.*, *Sect. B* **2002**, *58*, 389.
[12] S. F. A. Kettle, 'Symmetry and Structure', John Wiley, London, 1987.

Received February 5, 2003