Durham E-Theses

Systematic structural studies in metal complex chemistry

Yao, Jing Wen

How to cite:

Yao, Jing Wen (1998) *Systematic structural studies in metal complex chemistry*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/5055/

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, Durham University, University Office, Old Elvet, Durham DH1 3HP
e-mail: e-theses.admin@dur.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

Systematic Structural Studies in Metal Complex Chemistry

by

Jing Wen Yao

The copyright of this thesis rests with the author. No quotation from it should be published without the written consent of the author and information derived from it should be acknowledged.

A thesis submitted for the degree of Doctor of Philosophy

Department of Chemistry

University of Durham

November, 1998

16 APR 1999

Contents

List of Figures

List of Tables

Acknowledgements

Firstly, I would like to thank my supervisors, Professor Judith A. K. Howard and Dr. Frank H. Allen (CCDC) for their help, encouragement and support during these last three years, without whose guidance this work would not have been possible.

I am grateful to the Cambridge Crystallographic Date Centre for providing a financial support and to the CVCP for an ORS awards and Department of Chemistry, University of Durham for the tuition fees.

I wish to thank Dr. Royston C. B. Copley for some useful suggestions used in systematic study for metal seven-coordination system. My thanks are also due to all the members of the crystallography group in Durham for their help and friendship.

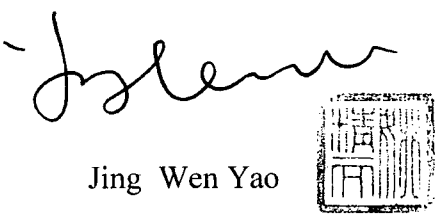
I would like to take this opportunity to thank Professor Vernon C. Gibson, Dr. Andrew K. Hughes and their former students, Carl Redshaw, Leela Sequeria and Gary L. P. Walker and Brian Bridgewater, in the department, who supplied the crystal samples so that I can perform the X-ray diffraction experiments and obtain the valuable structural data.

I would also like to thank my wife and son, for their support and understanding.

Declaration

I declare that the work presented in this thesis, except where otherwise stated, is based on my own research and has not been submitted previously for a degree in this or any other University.

Signed



Jing Wen Yao

The copyright of this thesis rests with the author. No quotation from it should be published without their prior written consent and information derived from it should be acknowledged.

Abstract

A procedure is presented for detecting geometrical preferences, deformations and interconversion pathways between different geometries for the transition metal coordination sphere ML_n. A discrepancy index [$R_{ang}(x)$] was proposed initially to address the problems of dimensionality and permutation complexity in the systematic analysis of coordination sphere geometry with higher coordination numbers ($n \geq 7$). But it can also be used generally for the lower coordination numbers.

A set of standard geometries for coordination numbers 2-9 are presented and the angles between the center point and each vertex for the polyhedra which are used to describe the coordination sphere geometries for coordination numbers 7-9 are idealised. These angles correspond to the metal-ligand valence angles in the coordination complex and are used as the standard values to measure the deviation of a real coordination sphere in the complex from these standard polyhedra.

Geometry of each coordination sphere (ML_{7-9}) from the Cambridge Structural Database (CSD) is identified by the calculations of $R_{ang}(x)$ values. Also the unique enumeration numbers of the ligands corresponding to each geometry can be derived over the *n!* ligand permutations. The different geometrical clusters and interconversion pathways from one to another are mapped in a designed two-dimensional plot.

The symmetry coordinates and principal component analysis are initially applied in these higher coordination number systems. They not only map the clusters represented to those standard geometries in the different symmetric point groups but also provide and confirm the interconversion pathways between the different geometries.

The other systematic study involves the analysis and correlation of the metal σ-π bond in the transition metal alkyne and alkene complexes from the CSD. Geometrical features of this specific bond are examined in the view of structure and some useful correlation between the key geometrical parameters are defined.

Finally, X-ray crystal structure determinations are briefly described and the crystal structures of ten transition metal compounds in coordination numbers 4-6 are presented.

Chapter 1

Introduction

The geometry of metal coordination spheres in metal complexes is a fundamental aspect of coordination chemistry. Knowledge of the coordination geometry provides the basis of orbital energies, bond properties of metal and ligand atoms and electron configurations of the metal atoms or ions etc.

Crystal structures of transition metal coordination compounds determined by X-ray diffraction techniques show the most direct picture of the coordination geometry and various geometrical polyhedra have been used to describe the structure of the coordination sphere for different coordination numbers. Then, what factors determine a given metal coordination geometry and a detailed mechanism for metal-ligand bonding can be modelled? Acquiring and accumulating knowledge from known 3D structures may lead to modification and development of the original basis and reiteration of the modeling experiments in the absence of adequate theoretical models.

With the developments of X-ray diffraction techniques, much structural data on transition metal coordination complexes is available in the Cambridge Structural Database (CSD) (Allen and Kennard, 1993[1]). Recently, structural systematics on various organic molecular assemblies, ensembles, and all manners of structure aggregates have formed the basis of the subject of crystal engineering to predict and design new functional structures. The main aspects of these applications are presented in Chapter 2.

Identifying recurring geometrical preferences of coordination spheres, deformations of metal complexes in various coordination environments from the standard geometry, and interconversion pathways between different geometries are major objectives of this work. The CSD can provide the source of the known crystallographic metal complexes structural data. However, the conventional methods used in organic systems are inadequate in coordination systems, especially for the higher coordination numbers, since the numbers of parameters to describe the geometry (metal ligand valence angles, L-M-L) rise rapidly with the increase of coordination number *n*. The research project presented in this thesis develops and programs a simple

[1] References appearing in this Chapter are given together with the references of Chapter 2

and general method to address the problems of complexity and the added difficulty of ligand permutations. In Chapter 3, the full definitions of standard geometries for coordination number from 2 to 9 are given. The difficulties and restrictions in using conventional techniques in multivariate analysis compared with the flexible organic systems are stated.

In Chapter 4, the successful results of using the proposed method and combination of multivariate analyses are illustrated, which originally reveal the geometrical preferences, deformations and interconversion pathways in 7, 8 and 9-coordination.

In Chapter 5, systematic analysis is applied in metal alkyne and alkene complex systems. A large number of metal π-bond complexes have been retrieved from the CSD. The metal π-bond interactions for C≡C and C=C are investigated using database techniques. The resultant plots demonstrate the correlation of some key parameters that characterize the geometry of these complexes, strength of the metal-ligand interactions, etc. The coordination environmental effects to form the special σ-π bond are systematically discussed.

Some practical transition metal coordination crystal structures determined by X-ray diffraction are presented in Chapter 6. Accurate structural data forms the basis required for structure correlation. In addition to the structure diagrams, the major geometrical parameters for the coordination sphere are listed. Special attention is drawn to the geometry of coordination spheres that are related to the interests described in the earlier chapters.

Finally, further work that might be carried out with the study of higher coordination number sphere geometry is discussed in Chapter 7. Based on the obtained results, more interests are in interconversion pathways from one geometry to another, i.e. from $n \rightarrow n+1$ coordination. Such studies have been seen in lower coordination systems, e.g. reaction pathways from 4-coordination to 5-coordination, by investigating the geometrically similar reaction pathways involving the formation of a five-coordinate species (Cross, 1985). But, because of more variety in higher coordination geometries, it is more difficult to correlate the characterized geometrical parameters in the same manner. Preliminary work to explore such a reaction pathway from 6-coordination to 7-coordination is investigated and given in this chapter. More effort in defining sensible geometrical parameters is needed and more effective analytical methods need to be developed.

Chapter 2.

Structural Systematics: the Cambridge Structural Database

2.1 From X-ray Structures to Database Research

Determinations of crystal and molecular structures by X-ray diffraction techniques have become a routine method, with modern developments of computer techniques and a variety of methods for the solution of structures. X-ray crystal structures provide the three-dimensional geometry of molecules by giving bond lengths, valence angles and other parameters. It has proved to be a very successful method for chemists to understand the geometry of the three-dimensional atomic arrangements of molecules since this technique appeared in 1910's. A brief description of X-ray structure determination and some crystal structures will be presented in Chapter 6.

From the crystallographic results using X-ray diffraction methods, the positions of the atoms in the crystal unit cell can be located. Thereafter, the structural parameters describing the molecular geometry can be derived easily. These parameters include bond lengths, bond angles, torsion angles and any other geometric parameters, which can be used to describe the molecular structure. In addition, non-bonded distances and angles describe inter- or intra-molecular non-covalent interactions.

The properties of compounds depend upon their connectivity (two-dimensional or 2D) structure and on their geometrical (three-dimensional or 3D) structure. Atomic connectivities are described by certain fixed bonding formations, but more detailed information is contained in the geometrical details. Thus, further valuable structural information can be derived from systematic studies of geometric structure. Usually configuration, conformation and the supramolecular interactions of molecules and any structure-reactivity relationships are interesting to chemists. Biochemists meanwhile are interested in structure-bioactivity relationships. In summary, further investigations of the relationship between three-dimensional structure and molecular properties could be obtained by applied systematic analyses of structure parameters. Theoretical calculations based on quantum mechanical theory can also provide useful information on chemical

bonding and interpretations for some physical and chemical properties of compounds from the modeling of special functional groups. But in practice, such calculations are still limited to the small model molecule, because of the limitation of the computer and, in some cases, a lack of the suitable parameters for modeling purposes. An experimentally determined molecular structure presents a real geometric arrangement in three dimensions, so when a defined function group or a molecular fragment is investigated in different bonding environments from large numbers of known structures, a preferred geometric conformation or bonding form may be found which usually represents a low energy state. That is, comparison of several or many three-dimensional structures can also provide important insights. On the other hand, these sophisticated analyses of the three-dimensional properties of molecules are of great potential use to molecular modeling. Thus, it is necessary to establish fundamental tools for the study of structure-activity relationships from experimental data.

In the early stage of using X-ray crystallographic techniques, the numbers of compounds that could be characterized by X-ray crystal structure analysis were limited because of the restrictions of the equipment and the theoretical basis of structure solution. Therefore, it was possible to discuss nearly all of the determined structures in some detail. The earliest systematic study can be traced back to Pauling's "Nature of the Chemical Bond" (Pauling, 1940). All the information derived from this basis about the covalent bond, the ionic bond and the metallic van der Waals bond continued to be used for about half a century. Nowadays, the total number of structure determinations of small organic molecules has reached 200,000 and is increasing at the rate of about 15,000 per year. Such large sets of data definitely provide a valuable source through which to correlate the geometrical results systematically. But it also raises the difficulty of developing appropriate methods to search and organize the available data into a suitable form for systematic analysis.

As the Cambridge Structural Database (CSD) expanded (Allen, Kennard, and Taylor, 1983), it made systematic studies on large numbers of related structures possible. It not only provide automatic searching for a defined fragment or text etc., but also enabled statistical and numerical techniques to be applied to the large volumes of numerical data that may result from searches. Thus, use of the CSD has greatly facilitated the survey of large sets of X-ray derived geometrical data. Random errors are

averaged out, real errors are usually identified as outliers from individual structures and then patterns may become obvious in such a survey.

The combination of computational and experimental methods often provides a more complete result, particularly for simpler systems. In such studies, experimental observations are often used to compare with results from *ab inito* theoretical calculations. Often they gave consistent or similar distributions for the defined key parameters. For example, Allen, Harris and Taylor (1996) recently carried out a study of the conformations of some simple molecular substructures that contain alkyl chains about their central C-C bonds or some C-S bonds. In the CSD study, the torsion angle

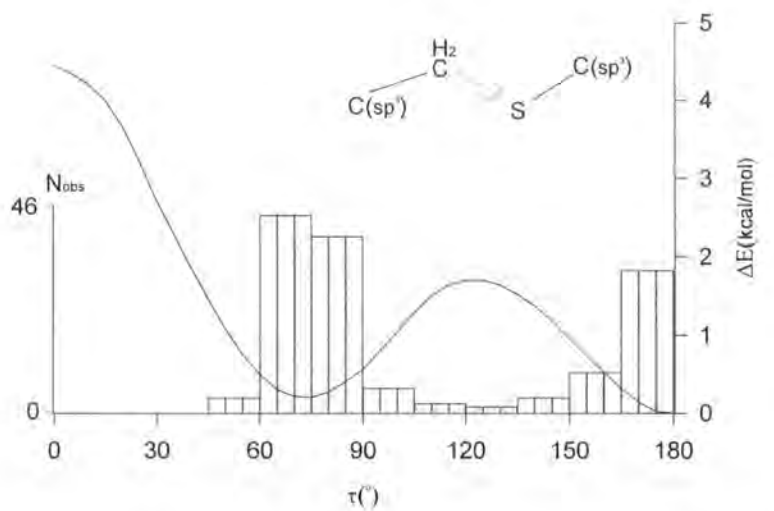


Figure 2.1 *ab initio* calculation of energy ΔE on the torsion angle τ and the frequency of appearance N_{obs} with this torsion angle in the CSD (Allen, *Harris and Taylor*, 1996).

about this central bond was chosen as the key geometrical parameter to express its conformational preference in the crystalline state. The geometrical results are presented as a torsional histogram for all observations of this fragment in crystal structures. When these distributions are compared with energy profiles computed for relevant model compounds by *ab initio* molecular orbital methods, it shows that the most frequent occurrence of real structures appears at similar angles with energy minimum from the distribution of energy versus the conformation angle. One of these substructures

involving a C-S bond is given in Figure 2.1. It shows the expected torsion angles representing *anti*, $\tau \sim 180^{\circ}$ and *gauche*, $\tau \sim 75^{\circ}$ ($\tau \sim 60^{\circ}$ for C-C bond) conformations. This indicates good agreement between theoretical calculation and observations. Crystallographic conformations obviously represent energetically accessible forms and the peaks in their torsional distributions usually correspond to minima in their potential energy surface.

Nevertheless, for more complicated systems, computational methods may only play a restricted role, due to the limitations stated above. Thus, systematic analysis of existing structures stored in crystallographic databases becomes a more direct way to acquire knowledge which can contribute significantly to the studies of structure/activity relationships.

2.2 Cambridge Structural Database

The Cambridge Structural Database (CSD) stores experimental atomic coordinates for ~190,000 organic and organometallic compounds as well as containing software facilities for search, retrieval, analysis and display of these structure contents. There are about 10% of the total number added each year as new entries. In the 1995 version, the extraction of defined structure parameters and three-dimensional coordinates for a particular interest from the database is carried out by identifying the defined atomic symbols and geometric relationships to neighboring atoms; the fragments to be searched could be defined in two-dimensional (2D), three-dimensional (3D) or bibliographic (1D) etc. information according to the needs of problems. After the search step is finished, the defined parameters may be printed or further analyzed.

2.2.1 Information Content of the CSD

To retrieve structure information rapidly and effectively, three distinct types of information for each crystallographic entry are stored in the CSD. These are conveniently categorized in terms of their "dimensionality" as illustrated in Figure 2.2.

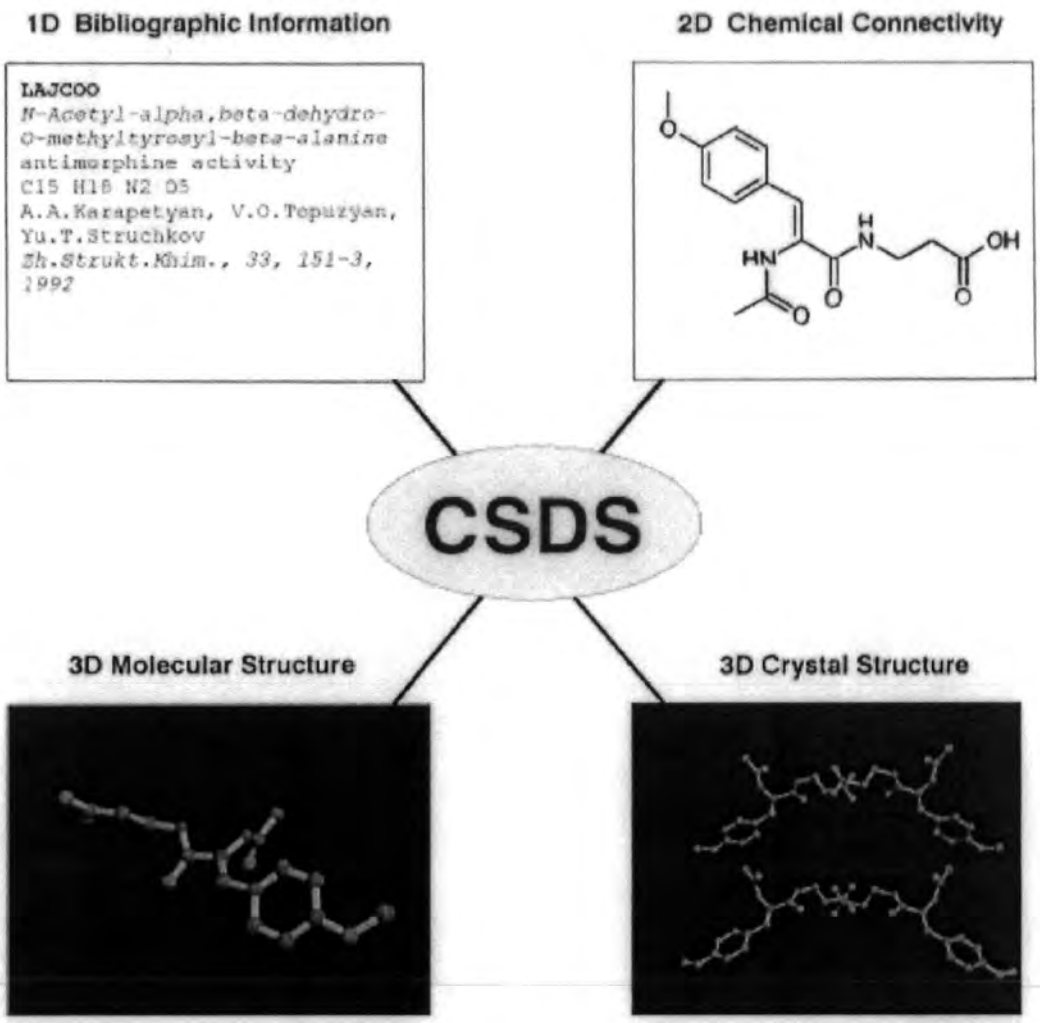


Figure 2.2 The Cambridge Structural Database System information fields (*reproduced by kind permission of the CCDC*)

1D Bibliographic Information

These data fields include all of the bibliographic material and some experimental information for the crystal structures, which concerns: chemical text, *e.g.* compound names, molecular formula etc., author's names and literature references, crystallographic cell dimensions and space group, and other simple text and numerical information for the particular entry.

2D Chemical Connectivity

The 2D diagram of the chemical connections is encoded as a connection table, in which atom and bond properties are defined. It records atoms as: atom sequence number, element type, the number of connected non-hydrogen and number of connected hydrogen atoms and the net charge; and indicates bond properties by different bond types, that is, 1 = single, 2 = double, 3 = triple, 4 = quadruple (metal-metal), 5 = aromatic, 6 = "catemeric link", 7 = delocalised double and 9 = π-bond.

The most important searches of the CSD involve the matching of chemical substructural queries to the 2D connection tables stored in the database. Queries are entered via a 2D chemical drawing interface.

3D Structural Data

The 3D atomic coordinates, the space group symmetry, the covalent radii and the crystallographic connectivity established by using these radii are contained in this information field. Using all this information, a 3D molecular diagram can be represented.

All of these three forms constitute the basis of storage of the CSD, which has been extracted from the primary literature or provided by authors, with over 800 journals represented in the database.

2.2.2 Software Systems of the CSD

The CSD provides a menu-driven interactive graphics software system for a variety of computer platforms.

- The **QUEST3D** has been designed for search and retrieval of fragments or molecules to be investigated. The search input is based on the principle of information being stored at three levels of "dimensionality". The details are defined according to different requirements.

The program provides multiple options for those 1D/2D/3D information fields. One of the most commonly used definitions for the search is the location of 2D substructure, which may then be further constrained using wide variety of 3D geometrical criteria. The search query can be constructed graphically and a number of separate queries can be combined.

Generally, the program will output files that contain user-defined geometrical parameters for any substructure found in the search process. Other types of files also contain the basic geometrical parameters and can be output selectively for connection to external software packages.

- When the search process is done, hits that meet the requirements specified in the input search code are extracted from the CSD. The required geometrical parameters may output in the form of a table, and the structures can be re-displayed on a computer graphics screen by the software **PLUTO**. The program PLUTO takes the coordinate data from the search as its input file. When displaying the three-dimensional structure on the graphics screen, it provides rotation and translation functions so that the user can view a structure from the best direction and also the bond lengths and angles can be calculated and displayed on the screen by clicking on the required atoms. It is also capable of exploring the non-bonded networks presented in crystal structures.

- The geometric parameter analysis program, **VISTA** (VISTA 2.0), provides further facilities to display and manipulate three-dimensional structural images. VISTA reads table files that are automatically generated by QUEST when 3D parameters are defined for a CSD search. Some useful statistical procedures, such as histogram, scattergram, correlation/covariance and principal component analysis, are included. Thus, the CSD is not only a database to store structural information, but also provides a

fundamental tool for molecular geometric analysis. "Only in this way can we form classifications and recognize patterns of structural behavior that may be interpreted in chemical terms" (Allen, Kennard and Watson, 1994).

- **IsoStar** is a recently released program which provides experimental and theoretical information on non-bonded interactions.

The IsoStar knowledge base contains information derived from the CSD and Broohaven Protein Data Bank (PDB) and molecular orbital calculations. Thus, the geometries, energies and frequencies of occurrence of different types of non-bonded contacts are presented in the form of scatterplots by the selection of a "central group" and a "contact group" during the CSD non-bond search. It is therefore ideal for identifying unusual attractive interactions. This knowledge-based library of intermolecular interactions can provide as a useful tool in the field of crystal engineering and rational drug design.

2.3 An Overview on the Application of Cambridge Structural Database

The CSD provides a reservoir of crystal structures determined by X-ray or neutron diffraction methods, and each database entry is assigned by the reference code. It is usually simplified as "refcode". The refcode consists of six alphabetic characters and if more than one publication for the same compound under different experimental conditions can be found, such as low temperature or supersession of a full publication to the preliminary publication, a possible further two numeric digits is used to identify the differences. Based on features of the three-dimensional structures one can use this rich source to derive information which is important to chemical or biochemical reactivity, by comparing analogous molecular geometries in various different crystal structures. For a simple instance, for chloroethanoic acid [shown in Figure 2.3(a) (Kanters and Roelofsen, 1976)], many example fragments can be found from a CSD search. Each individual structure has its own fixed arrangement for this molecule. However, when all molecular fragments from different crystal structures are compared, we can see not only a certain three dimensional arrangement which most fragments adopt, but also some deviations from this favoured form. Figure 2.3(b) gives a superposed plot of twelve such

fragments (refcode: ACLMEN etc). It shows the possible conformational variability among these analogous molecules. It can be seen that the carboxyl group has smaller variability than the substituents on the α carbon, because of its more fixed vibration and rotation. For more complex systems, some other methods, e.g. statistical etc. can be used to assist in identifying and classifying the geometric preference and thereafter the optimal geometry is derived. Identifying such conformations is helpful to molecule modeling.

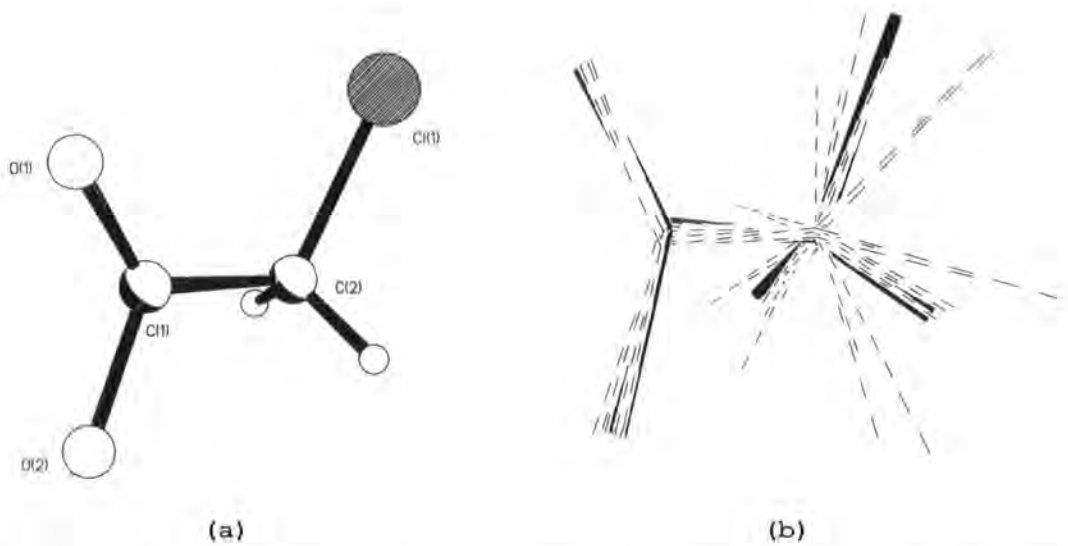


Figure 2.3 Conformational variability of chloroethanoic acid. **(a)** a diagram shows a single molecule (CLACET01) **(b)** a line drawing of twelve superimposed fragments from the CSD.

For a particular molecule of interest, a search procedure QUEST3D is carried out to find the defined full molecule or part of a chemical structure in the CSD crystal structures. Such a search procedure includes many search fields (CSD, 1992), but only some of them are commonly used. Usually, connectivity searches are used to search the precisely defined bonded fragment or molecule by drawing atomic connections in the drawing field. In this way, the exact structure described in the input code (such as in the 2D-CONSTRAIN sub-menu, select an ATOM or BOND to display the properties of bonded atoms) can be obtained and also a variety of ranges of bonding parameters can be defined or calculated (in the 3D-CONSTRAIN sub-menu). Of particular interest, the

geometric parameters of intermolecular interactions can be retrieved, e.g. hydrogen bond by specifying the hydrogen atoms 'bonded' to a given atom using distance criteria. The other method for search is to use text searches. Some search keywords are available in the TEXT sub-menu. As given above, they include compound names, formula, the names of authors or any item referring to the literature reference, such as year or journal name, and so on. Both these input searching codes can be used either independently and or together to give more restrictions to the search.

There are several ways to output or present the results from the search. Except for those used in VISTA and PLUTO, one of the possible outputs (GLIST) is to retrieve the geometric parameters, bond lengths and angles, in a formatted file *.gls, which not only give a conventional numeric values and also is ready for connection to an external program for further analysis, for example, for a seven-coordination tungsten complex (Umland and Vahrenkamp, 1982), $W(CO)_4I_2P(CH_3)_3$, it is represented by a Refcode, BOGBII in CSD. The geometric data are listed in Table 2.1.

As the facilities are available in database and with the improvements of the database software, the interest in the systematic application of chemical and crystallographic results by using the CSD as a basis, has been increased greatly since the late 1980's.

Table 2.1 An output form of example BOGBII in the *.gls file

BOGBII geometric data
Bond lengths

W1	I1	2.845	W1	P1	2.564	W1	C1	2.058	W1	C2	2.061
W1	C3	1.978	W1	I1C	2.845	W1	C1C	2.058	P1	C4	1.855
P1	C5	1.747	P1	C5C	1.747	C1	O1	1.078	C2	O2	1.120
C3	O3	1.161	C1C	O1C	1.078						

Bond angles

I1	W1	P1	90.5	I1	W1	C1	156.4	I1	W1	C2	77.3
I1	W1	C3	128.3	I1	W1	I1C	86.5	I1	W1	C1C	76.7
P1	W1	C1	76.3	P1	W1	C2	163.1	P1	W1	C3	121.0
P1	W1	I1C	90.5	P1	W1	C1C	76.3	C1	W1	C2	111.4
C1	W1	C3	73.3	C1	W1	I1C	76.7	C1	W1	C1C	115.7
C2	W1	C3	75.9	C2	W1	I1C	77.3	C2	W1	C1C	111.4
C3	W1	I1C	128.3	C3	W1	C1C	73.3	I1C	W1	C1C	158.4
W1	P1	C4	112.0	W1	P1	C5	119.9	W1	P1	C5C	119.9
C4	P1	C5	101.5	C4	P1	C5C	101.5	C5	P1	C5C	99.1
W1	C1	O1	175.0	W1	C2	O2	176.7	W1	C3	O3	180.0
W1	C1C	O1C	175.0	W1	C1C	O1C	175.0				

2.3.1 Correlation Between the Specified Structure Parameters

As is known, for a defined functional group, atoms are bound in a certain form. Usually, a set of standard bond parameters is given. With different bonding environments and the effects of some other groups, the values vary, but they all will fall within some acceptable range. For a specified chemical bond, certain bond lengths occur more frequently than others do. If the values of the various bond lengths are represented as a statistical result in a histogram, the peaks in the results indicate what the bond lengths ought to be for various types of such bonds. Normally, the histogram gives a distribution in which most structures have the central tendency around the certain value and the range of the values can be defined from the whole data set. Then, the dispersion of individual observations from the mean value can be inspected to find how the crystal structure represents the dispersion value. Thus, these systematic studies will allow chemists to investigate how the structure affects the distribution.

In the analysis of the retrieved data for systematic studies, the comparisons of many three dimensional structures for the specified parameter can be carried out, which could reveal some important structural features and common properties related to the structures.

The average geometrical parameters for various atomic pairs and functional group are very helpful for chemist in a general sense and also invaluable in model building. Before the CSD appeared, there were some tables of these standard values published based on the structural data obtained before 1960 (Sutton, 1958, 1965). As a greater number and more accurately determined structural data were collected in the CSD, by examining those stored structural data, more complete tables of mean bond lengths for a wide variety of atom pairs and functional groups, which now represent standard values for these bonds, have been given (Allen, Kennard, Watson, Brammer, Orpen and Taylor, 1987; Orpen, Brammer, Allen, Kennard, Watson and Taylor, 1989).

The tables were built on several values. Through the CSD search on the defined bond types, the mean value d of distances between atoms in organic and organometallic compounds and complexes can be calculated as:

$$d = \sum_{i=1}^{n} d_i / n$$

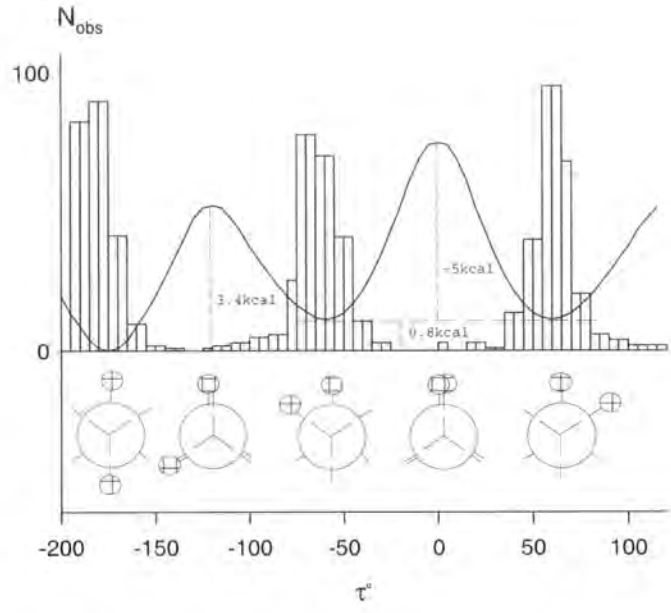
where d_i represents the ith bond length of the observations and n is the total number of observations. In addition, other values, the sample median m, the sample standard deviation σ, the lower quartile for the sample q_l and the upper quartile q_u are given together (Allen *et al*, 1987) to show the statistical results from the retrieval. Now, all these average geometries which include element H, B, C, N, O, F, Si, P, S, Cl, As, Se, Br, Te, I and 682 different bond types in organic compounds and, for metal-ligand distances between the *d*- and *f*-block metals Sc-Zn, Y-Cu, La-Hg, Ce-Lu, Th-U and atoms H, B, C, N, O, F, Si, P, S, Cl, As, Se, Br, Te and I of ligands in organometallic complexes, have been updated in *International Tables for Crystallography*, Volume C (1992).

Obviously, it is important to have a correct definition for the bond type of interest. Thus, the effect of different chemical environment to each bond could be compared. For a certain bond type, the bond length or angle, due to, e.g. substitution and re-hybridization on a defined atom, may deviate from the standard value. This can be seen easily in any individual crystal structure. The systematic analysis of many related structures can reveal how these substitutions and rehybridizations affect these geometric variations. Some examples can be found in the studies of substitutent-induced ring deformations in benzene (Domenicano, Vaciago and Coulson, 1975; Domenicano and Vaciago, 1979) and the effect of small-ring fusion on the geometry of benzene etc. (Allen, 1981). In the former example, the correlation of change of the specified angle versus electronegativity of substituted groups was given and showed the angular deformations of benzene dependent upon the electron-withdrawing or -donating ability of groups. And in the latter, a variety of mean geometry for monocycloalkeno-benzenes with different ring sizes (3-6) were investigated. The conclusion that small-ring fusion to benzene might cause partial double-bond fixation in the aromatic ring, was further confirmed in this study and the variation of geometric parameters on the benzene ring with different size ring were illustrated. The information derived from such studies can be correlated with theoretical calculations, reactivity, spectral properties, and other physical phenomena (Allen, Kennard and Taylor, 1983).

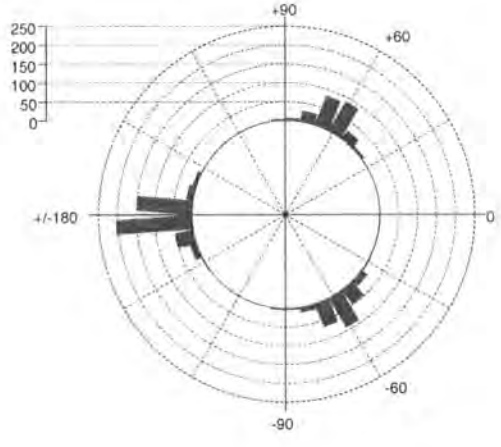
2.3.2 Relationships Between Conformation and Stability

With a minimum potential energy, the spatial arrangements of atoms in a molecule give a stable conformation. In solid-state conformation, such arrangements are directly dependent on the specific chemical and crystallographic environment. From the crystallographic experiment, each structure represents a local minimum energy state and the information about the preferred conformation of molecule can be obtained. Thus, in a systematic study for an interesting substructure or fragment, by defining some structural parameters that can represent the molecular conformation, the various possible conformations for different molecules can be examined from the correlation of these parameters. Average or typical geometries for each fragment or molecule are derived. Meanwhile, clear correlation between geometric parameters and some chemical, physical or biological properties will provide a sound basis of the structure correlation principle.

Bond lengths and angles are considered to have only minor changes within different conformation. Torsion angles are usually used to describe the conformations, because they measure rotation about bonds. In a simple system, sometimes a single torsion angle distribution can express a clear conformational preference. A well-known example is 1,2-disubstituted ethanes. The substituent groups here include halogen atoms and *S*, *Si* etc. The single torsion angle is defined as τ about the C-C bond. As τ runs from 0° to $\pm 180^{\circ}$, corresponding to a complete revolution of one substituent group with respect to the other. Obviously, when $\tau=\pm 180^{\circ}$, i.e. staggered conformation, it is always significantly more stable than the eclipsed conformation ($\tau=0^{\circ}$), because in the former, the atoms are as far away from one another as possible and offer the least repulsive interaction. In the plot of one-dimensional potential energy versus torsion angle [Figure 2.4(a)], the minimum energy occurs at this special angle value (Dunitz, 1983). Between these two extremes, there are other two possible conformation at $\tau=60^{\circ}$ and $\tau=-60^{\circ}$ respectively, which also have low energies. The polar histogram [Figure 2.4(b)] resulted from a CSD search and shows a clear preference for τ to be 180° or close to $\pm 60^{\circ}$ which is in agreement with the distribution of potential energy. Since the potential energy at $\tau=\pm 180^{\circ}$ is even lower than those at $\tau=\pm 60^{\circ}$, the relative population of the three peaks would indicate that staggered conformation is more preferred for the real crystalline state of molecules. This is further confirmation to the conformational knowledge base



(a)



(b)

Figure 2.4 **(a)** Potential energy versus conformation torsion angle τ°(Glusker *et al*, 1984)for 1,2-substituented ethane (Substituent groups include halogen, S and Si atoms) and **(b)** frequency of appearance N_{obs} in the CSD versus this torsion angle.

and is useful in the study of their reactivity.

The other system in which there is larger influence of conformational factors, is cyclic molecules. It becomes necessary to use more parameters to define the geometry of the chemical fragment. For an *n*-membered ring, the *n* intra-annular torsion angles are needed. The various conformations that a cyclic molecule may adopt depend mainly on changes in these torsion angles. Such a systematic study is actually a multivariate analysis. If I_r experiment results could be found from the CSD and I_c parameters are used to described the geometry of each observation. These $I_r \times I_c$ elements can be directly generated by the CSD system through defining suitable parameters in the search query of QUEST3D. The matrix M(I_r,I_c) consisted by I_r row and I_c columns may be directly used for the geometry analysis. Also some new relevant variables could be derived from these $I_r \times I_c$ elements. The problem is that with increase of the number *n*, the dimension of the matrix M becomes large. This definitely brings some difficulty to the system. Unlike the above example in ethane, where only one parameter was defined and all the conformations are dependent on this significant value, the meaning of each of multivariable is usually not known. Therefore, the methods used in the simple system through selection of individual parameters or pairs of parameters are not suitable in this kind of system. But some multivariate analysis techniques can be used for the M-matrix.

Covariance and Correlation

One of the methods to analyze multivariate systems in the CSD is to measure the dependence between two parameters. This can be achieved through the calculation of their covariance *C(x,y)* and correlation R(x,y). The covariance is defined as:

$$C(x, y) = \sum_{i=1}^{n} [(x_i - \bar{x}_a)(y_i - \bar{y}_a)] / n$$

in which, x_i and y_i are the *i*th observation values of two variables, and $\bar{x}_a$ and $\bar{y}_a$ are their mean values, respectively. When *C(x,y)* is normalized, the correlation coefficient *R* is given as:

$$R = \frac{C(x,y)}{\frac{1}{n}\left\{\left[\sum_{i=1}^{n}(x_i - \bar{x}_a)^2\right]\left[\sum_{i=1}^{n}(y_i - \bar{y}_a)^2\right]\right\}^{1/2}}$$

The correlation coefficient is normally valued between -1 and +1, which reflects how two parameters are correlated. The closer to |1| is the absolute coefficient, the more correlated they are to each other. The sign gives negative or positive correlation, respectively. This technique is more commonly used than covariance, because the latter is sometimes difficult to interpret, since it depends upon the units in which the two variables are expressed (Allen *et al*, 1994) Both of them are now available in the CSD (VISTA 2.0).

(b) Principal Component Analysis, Factor Analysis and Cluster Analysis

All of these three techniques are powerful tools in multivariate analysis. The full details will be given in later chapters and sections. The greatest advantage in using these methods is that the dimensionality of variables may be reduced dramatically. Geometric preference could be identified by the correlation of pairs of significant derived parameters.

Several techniques have been used cooperatively in the analysis of cyclic conformations. For six-membered carbocycles, the major existing shapes to describe the conformations are boat and chair. Considering symmetry factors, PCA was applied on six intra-annular torsion angles and the first three principal components accounted for approximately the total variance (99.9%) (Allen, Doyle and Taylor, 1991a). Based on these three PC's scatterplots, major different conformations can be clearly located according to the PC's coordinates. Further, by using cluster analysis through the calculation of dissimilarity coefficients (Everitt,1980), a unique and asymmetric set of conformational clusters was obtained. Interpretation of these results showed that more detailed conformations between the two major ones, such as, half-chair, twist-boat etc., could be observed. Some modified techniques and practical applications (Allen, Doyle

and Taylor, 1991b,c) provided more detailed clusters in the identification of the conformations. Introducing the spherical coordinates (Cremer and Pople, 1975) and combining to apply with the statistical techniques (Allen and Taylor, 1991, Allen, Taylor and Auf der Heyde, 1991) made a more lucid interpretation of the PCA results in which the chemical meanings of the PC's and some torsion angle combinations become obvious.

Similarly, these techniques stated above can be extended to apply to the medium rings of size seven and eight atoms. With the increase in the size, the ring bonds become more flexible and more conformational shapes could be adopted. In the consideration of the symmetry in this system, enumeration and permutations of the torsion angles were carefully taken into account, which is also a key step for this system in order to carry out the multivariate analysis.

For cycloheptane, there exist four kinds of shapes to describe the conformations, Twist-Chair(TC), Chair(C), Twist-Boat(TB) and Boat(B). They fall into two symmetries, C_2 for TC and TB, C_s for C and B. There were several coordinate systems additional to Cremer & Pople's puckering (CP) coordinates for n-membered rings mentioned above to define the conformations of cycloheptane. The most representative is BPRS four coordinate set (Bocian, Pickett, Rounds & Strauss, 1975), which was given according to the D_{7h} symmetry of cycloheptane. The details about these coordinates are not given here, the importance is that through some transformations (Bocian and Strauss, 1977a) it is possible to map the conformation of a given cycloheptane, by plotting pairs of the given coordinates and the inter-conversion between the shapes could be characterized (Bocian and Strauss, 1977b). By using the CSD data for cycloheptane compounds, a set of more complete conformational mappings was obtained based on these coordinate systems (Allen, Howard and Pitchford, 1993). Meanwhile, symmetry deformation coordinates, dissimilarity calculations, PCA and cluster analysis were all used to identify the preferred conformations in the data sets. An interesting result is that PCA is closely related to the CP analysis. Moreover, when the rings contain hetero-atoms or have substitutions (Allen, Howard, Pitchford, Vinter, 1994; Allen, Garner, Howard and Pitchford, 1994), the symmetry becomes lower and the permutation number is reduced. The results derived from applying the above techniques have related well to conformational energy differences obtained from theoretical calculations. Expanded applications to cyclooctane

and related rings (Allen, Howard and Pitchford, 1996) also mapped and classified the observed conformations from the CSD.

Systematic conformational analyses of both free and metal-coordinated unsaturated 12- (Raithby, Shields and Allen, 1997a) and 15-membered (Raithby, Shields and Allen, 1997b) oxa and thia macrocycles were also undertaken by using the CSD. The conformations existing in these species were classified by a symmetry-modified Jarvis-Patrick clustering technique and the classifications were visualized in conformational space by PCA plots. The possible geometry for the metal-coordination according to the size of metal ion and donor atoms were predicted and given, by comparisons of cluster populations and relative molecular mechanics energies,

In summary, the investigation of three-dimensional ring system structures, displayed the intramolecular or intermolecular geometric properties by using cluster analysis of defined torsion angles (Allen & Taylor, 1991). Combining the techniques of principal component analyses and cluster analysis, further studies on conformational mapping and classification were carried out. This permitted the results to be more readily comprehensible (Allen, Doyle and Auf der Heyde, 1991) and provided some useful points for an analysis of a higher dimensional data set (Allen, Howard and Pitchford, 1993). As an example, the combined multiple analyses identified the major conformational forms of cycloheptane as chair, twist-chair, boat and twist-boat, and supplied two independent pseudorotation pathways in the defined four-dimensional conformational space. More complex analyses are needed for a particular study to derive information on the relative stability of various possible conformations.

2.3.3 Reaction Pathways

In the study of organic reaction mechanisms, theoretical calculations related to the minimal energy states are usually used. In an ideal case, in terms of the structural geometry of the compounds, a variety of geometries representing each step along a reaction pathway are expected to be obtained. The important clue to understand how a reactant converts to a product is one or more transition states, which are usually drawn as saddle points in the energy plots. Crystalline structures represent stable atomic arrangements. Carefully chosen crystal structures with geometries very similar to those of the transition states and the investigation of the change in "active" bond order, may

provide support and evidence for this study. Once "a correlation can be found between two or more independent parameters describing the structure of a given structural fragment in a variety of environments, then the correlation function maps a minimum energy path in the corresponding parameter space" (Murray-Rust, Bürgi, and Dunitz, 1975). In the examples of the ring system conformational study of the last section, it has been seen that some interconversions between conformations could be expressed from correlation plots of the defined parameters. The facilities provided by the CSD made it possible to draw structural correlations in more direct ways.

The structure correlation method has been used to map reaction pathways for variety of different reactions. The most widely used correlation in the studies of ligand elimination from the tetrahedral L_3MY species (Murray-Rust, Bürgi and Dunitz, 1978) or nucleophilic addition to a carbonyl (Cieplak, 1994) was Pauling's relation between bond length and bond number, n_i (Pauling, 1947),

$$\Delta d_i = -c \log n_i$$

where Δd_i is given as a difference of the specific bond length d_i, e.g. M-Y, and the standard value. The bond number is related to the displacement D of central atom M from the plane of the three basal ligands L. The definition of variables is shown in Figure 2.5. In the case of $R_2C{=}O$ fragments, Figure 2.5 is still used for the expression but carbon atom replaces M, other substitutions and the oxygen atom are located at L. The adding or eliminating atom is at Y.

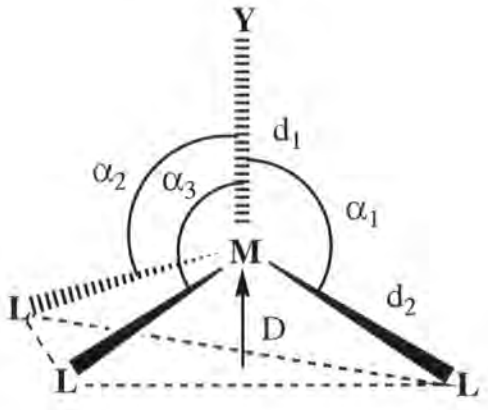


Figure 2.5 Definition of geometric parameters for L_3MY species related to Pauling's formula.

From such a structural correlation, reaction pathway from sp^2 to sp^3 transformations for nucleophilic addition to carbonyl was characterized. The elimination could be addressed as a geometric change from a regular tetrahedron ($n_2 = 1$) to trigonal planar molecule ($n_2 = 0$). Conversely, a similar principle could be used to seek a reaction pathway from three-coordination metal complexes plus one extra ligand to form four-coordination complexes (Pitchford, 1994). In this way, Cu, Ag and Au three-coordination complexes from the CSD were investigated. Suppose the fourth ligand attacks on the metal centre from an optimal geometric direction, i.e. that of orthogonal to the trigonal plane, the interaction distances d_1 of possible ligand atoms with the metal atoms were retrieved and plotted versus the distances D of the metal out of the trigonal plane. The results showed that the distance d_1 is inversely related to the out of plane distance D, and are consistent with the fitting curve from Pauling's formula.

The longest interaction represents the initial attack position, which corresponds to the D value close to zero, and an effectively trigonal planar shape, while the shortest interaction reflects the completed reaction with a large D, which is closer to a tetrahedron. The other example is the investigation of a series of lactam derivatives (Allen, Kennard and Taylor, 1983) from four structures in the CSD. With the change of the halogen atom in the silyl group, the geometry at silicon follows an inversion pathway, which varies starting from a distorted tetrahedron (GEGDOL) with substitution of *F* atom, *Si-O* is in the contact distance and not really bonded, through the *Cl* and *Br* substitutions (FOGBIM10, FUFXAF), *Si* atom passes through the equatorial plane of the trigonal bipyramid (both *Si-O* and *Si-X* bonded). Finally, it is the *I*

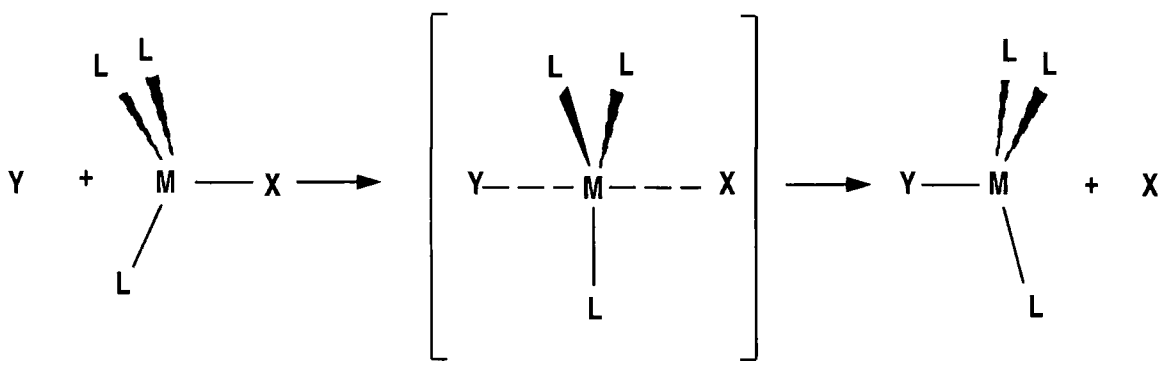
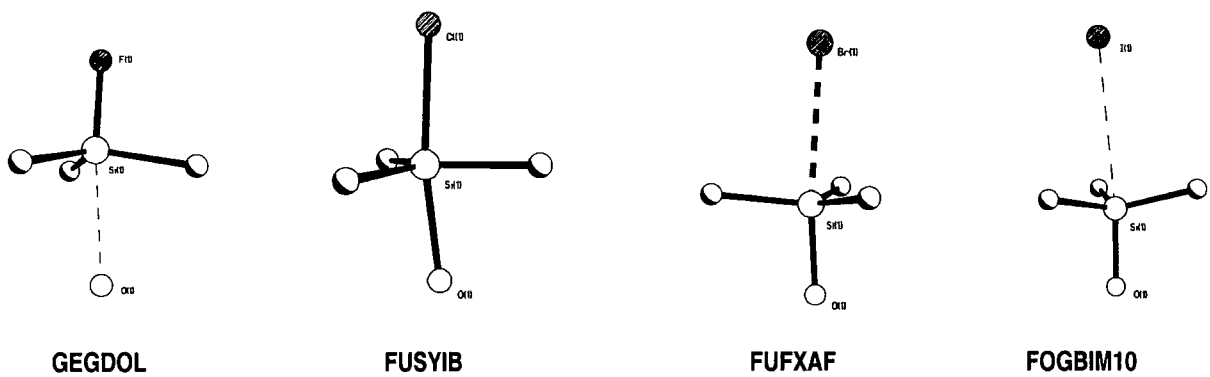


Figure 2.6 Bimolecular nucleophilic substitution reaction for a four-coordinate atom.

derivative (FUDYIB) with an inverted tetrahedron compared with the *F* compound, *Si-O* bonded and *Si-I* in contact distance. This can be related to the associative nucleophilic substitution mechanism for a four-coordinate atom (shown in Figure 2.6). The geometric variations of the real structures in this group are shown in Figure 2.7. The *Si-X* bond from *F* to *I* derivative could be considered as the leaving group *X* is broken, and a new bond, *Si-O*, is formed with the nucleophile (*O*). The intermediate or the transition state, may be seen as trigonal bipyramid (FOGBIM10, FUFXAF) with the leaving group X, and nucleophile *O* in axial positions. The final product of this reaction, is that the leaving group leaves from one side and the nucleophile adds to the other side of a



GEGDOL FUSYIB FUFXAF FOGBIM10

Figure 2.7 Geometry conversions from F to I substitution in lactam derivatives.

trigonal group, so that an inversion of configuration occurs. Similar mechanisms on other substituent groups can be observed in different examples (Zobetz, 1988, 1990).

If these are concerned with S_{N2} reactions, the correlation of some specified bond lengths and angles and their bond number *n* in the investigations of a series of compounds, particularly, *Al, S, P, Sn, Si* and *Ge* tetrahedral derivatives (Murray-Rust, Bürgi and Dunitz, 1975) represented a mechanism of S_{N1} reactions. As a leaving group dissociated, the intermediate was envisioned as planar trigonal, so that a nucleophilic group could attack from either side. Thus, configuration isomers, inversion and retention, were obtained with respect to the reactions. Several other analyses of

chemical reaction pathways from the observed structural geometry have also been studied (Auf der Heyde and Nassimbeni, 1984; Ferretti, 1986).

2.3.4 Intermolecular Interactions in Small Molecule Systems

An important application in database systematic studies is to examine intermolecular interactions. Molecules or ions packing in a crystal provide a regular order such that the total free energy of the system is a minimum. The intermolecular forces, which usually include van der Waals interactions (Glusker *et al*, 1994), hydrogen bonding and so on, also determine this packing. For an ionic compound, it is packed to ensure that the charges should be balanced on its own and also in the whole crystal. Non ionic compounds will adopt more varied of ways to pack together according to their bond properties, but hydrogen bonding is a preferred way if the functional groups are capable of accepting or donating hydrogen bonds. For example, in the $[HMBPT^*]_2[Si_8O_{18}(OH)_2]\cdot 41H_2O$ structure (Harris, Howard, Samadi-Maybodi, Yao and Smith, 1995), silicates are held together to form an octameric cubic cage $[Si_8O_{18}(OH)_2]^{6-}$ and linked into a framework by hydrogen bonding to water molecules. Two triply charged cations $[C_{18}H_{30}N_3]^{3+}$ balance the cage silicate anions to stabilize a host/guest system, (shown in Figure 2.8).

The CSD provides facilities to search for non-bonded interactions. Many of the published papers on this subject are concerned with the geometrical properties of hydrogen bonding, since hydrogen bonding plays a crucial role in packing molecules in extended crystal structures. While X-ray and especially neutron diffraction provide the best way to determine experimental hydrogen-atom positions, the CSD contains a lot of such data, and systematic analyses by retrieving and using these data, will give more directly the overall pattern than by using other methods, such as quantum chemical calculations.

In crystal structures, the distances $X\cdots A$ and $H\cdots A$ recognizing the hydrogen bond in the solid state are given (X the donor and A the acceptor atom). From the C-H···O contact distances and related angles, a statistical study in the CSD (Taylor and Kennard,

---

* HMBPT is 2,3,4,5,6,7,8,9-octahydro-2,2,5,5,8,8-hexamethyl-1H benzo(1,2c:3,4c':5,6c") tripyrrolium. $(C_{18}H_{30}N_3)^{3+}$

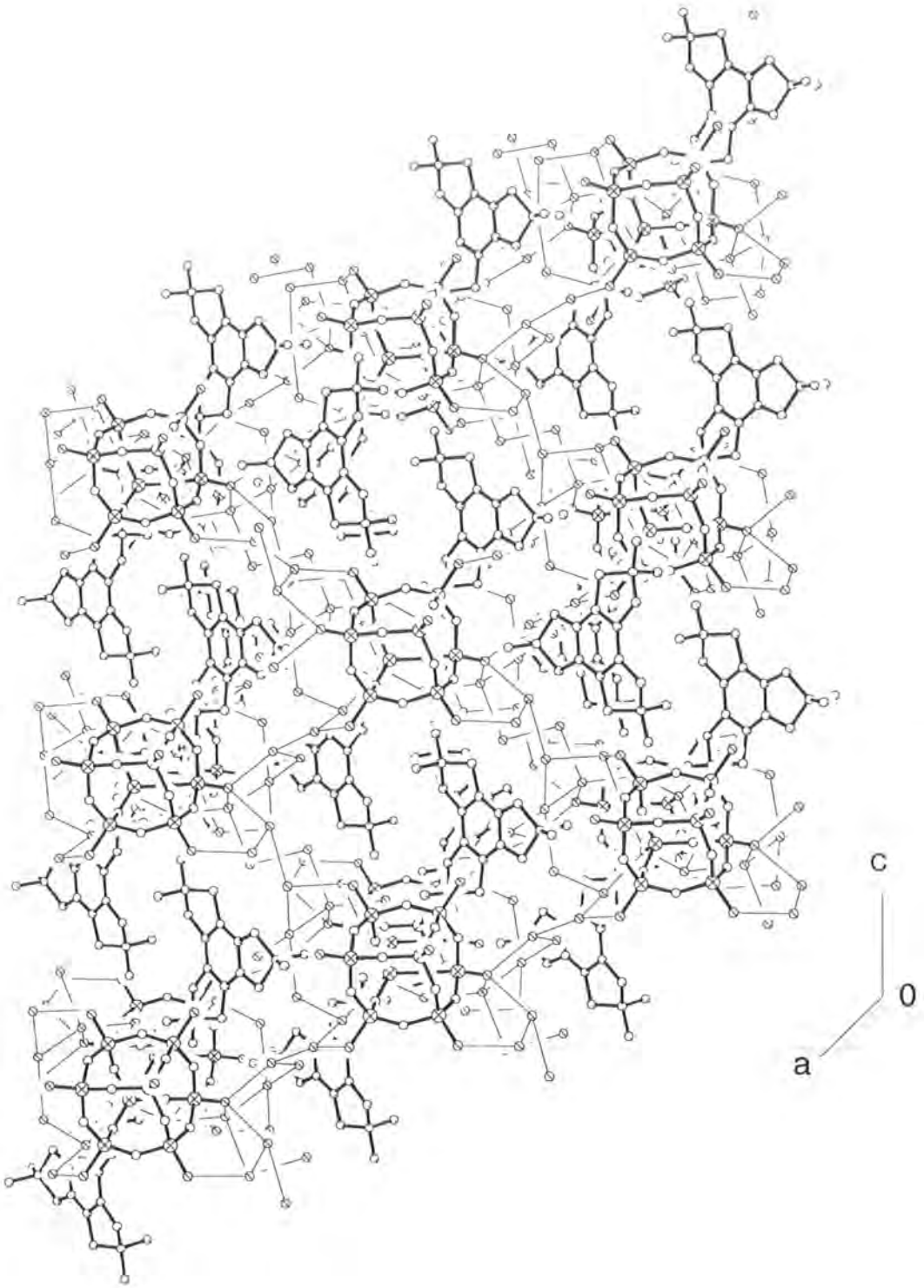
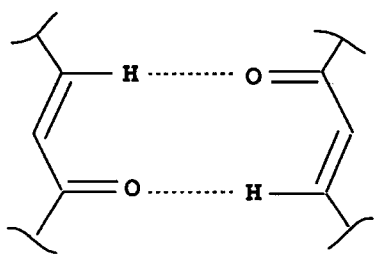


Figure 2.8 The network structure of compound $[HMBPT]_2[Si_8O_{18}(OH)_2]\cdot 41H_2O$

1982) showed clear results about the interaction distances and preferred orientations. With different donor and acceptor atoms and hydrogen bonding environments, hydrogen bonds illustrate different chemical bonding properties. Searches on particular molecules or fragments of this subject and use of different definitions on the parameters and combinations of some other methods for this study, various geometric properties and patterns of hydrogen bonds, such as, the effects of lone pair directionality (Glusker & Murray-Rust,1986) and orbital hybridization (Taylor and Kennard, 1984) to the hydrogen bond, could be characterized.

The C–H···O hydrogen bond system had been well established from crystal structures and also from spectroscopic data (Desiraju, 1991). The proton donors are present in the three different hybridization states C(sp^1)–H, C(sp^2)–H and C(sp^3)–H. Since the hydrogen atoms are connected to different acidic carbon atoms, the hydrogen bond characters could be different. The CSD investigation on the mean C···O distances of these different species showed that shorter C–H···O contacts are obtained for more acidic C–H groups (Desiraju, 1990). In C(sp^1)–H···O and C(sp^2)–H···O interactions, a common hydrogen bond pattern can be often viewed, which is that molecules easily form networks by twofold screw symmetry (Bernstein, Etter and Leiserowitz, 1994), for example, α, β unsaturated ketone functional group.



Sulfur *S* is one of the normal acceptor atoms in forming hydrogen bonds. The hydrogen-bond acceptor ability of sulfur in divalent *Y-S-X* and *C=S* systems (Allen, Bird, Rowland and Raithby, 1997a,b) have been studied using the CSD. The geometries were defined by the parameters shown in Figure 2.8 (a) (b) for *S*(sp^2) (*C=S*) and *S*(sp^3) (*Y-S-X*), respectively. The CSD provided all the related structures based on this definition. The results indicated that with different kind of substitutions on *C=S*, the

hydrogen bond lengths span a wide range. By comparison with the similar bond in $>C=O\cdots H\text{-}O$ and $>C=O\cdots H\text{-}N$, it was shown that much weaker hydrogen bond existed in sulfur systems. According to the definition of geometrical parameters, the preferred directions of H-approach to S were clearly illustrated through the polar histograms of pairs of parameters. Meanwhile, *ab initio* calculations of residual atomic charges and electrostatic potentials were also in agreement with the crystallographic observations.

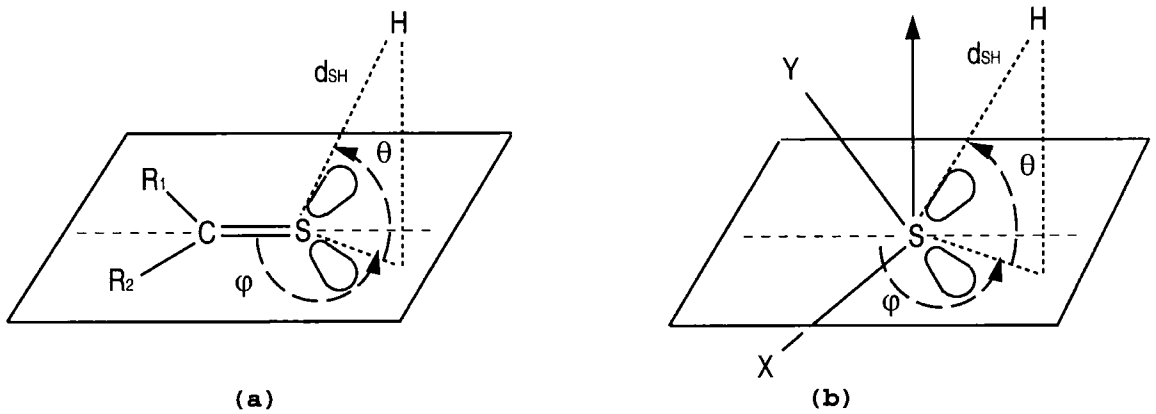


Figure 2.9 Definitions of geometric parameters for **(a)** S(sp^2) and **(b)** S(sp^3)

The direction from which hydrogen bonds approach sp^2 and sp^3 hybridized oxygen atoms have been investigated in a similar manner (Lommerse, Price and Taylor, 1997). The alkyl hydroxyl groups were chosen as hydrogen donor and oxygen atoms in carbonyl, ether, and ester groups as acceptors. The combinations of *ab initio* calculations for the electrostatic energy and contributions of the intermolecular energy terms in the hydrogen bonded model system have provided clear interpretations for the hydrogen bond characteristics of strength, geometry and directionality toward an acceptor molecule in different bonding environments. The work also showed the effects of a lone pair existing for acceptor oxygen atoms on the geometry of the formation of hydrogen bonds.

Hydrogen bonds can be also observed in O,N–H$\cdots\pi$-acceptor systems (Viswamitra, Radhakrishnan, Bandekar and Desiraju, 1993). Acetylenic, olefinic and aromatic groups are all able to form hydrogen bonds. A crystal structural database survey and energy

calculations have characterized the H···C≡C (midpoint) and phenyl (centroid) hydrogen bonds (Steiner, 1995a,b).

Three-membered rings can be as C-H donor to the hydrogen-bond. A CSD study (Allen, Lommerse, Hoy, Howard and Desiraju, 1996) on the C-H protons of cyclopropane and derivatives forming C-H···O hydrogen bonds showed a bond strength ordering: C(sp^1)-H···O > C(ring)-H···O similar or equal to C(sp^2)-H···O > C(sp^3)-H···O, agreed with the existing knowledge. The two interaction modes from edge or centre of the ring when the ring acts as the receptor of hydrogen bond were investigated.

The hydrogen bond between N and H can be seen from a study on the hydrogen-bond acceptor properties of *N* with geometry of the N(sp^2) → N(sp^3) transition (Allen, Bird, Rowland, Harris and Schwalbe, 1995). The substructures were defined as R_1(X=)C-NR_2R_3, (R_2,R_3 = C(sp^3) or H). The N atom undergoes a transition from sp^2 to sp^3 with gradual lone-pair formation on N. With the definition of the specific parameters, the change could be viewed from the variation of the relevant N···H geometry.

The systematic analysis of other aspects of intermolecular distances of some non-hydrogen bonds (Nyburg and Faerman, 1985) and molecular packing (Taylor, Mullaley and Mullier, 1990) were undertaken using the CSD. One of the interesting topics is that the analysis of geometry of intermolecular interaction between halogens and oxygen or nitrogen (Lommerse, Stone, Taylor and Allen, 1996). The CSD survey showed that C-X···O,N interaction distances are within the van der Waals radii sum and occur preferentially along the directions of the extended C-X bond axis. While IMPT (the InterMolecular Perturbation Theory; Hayes and Stone, 1984) calculations showed that these interactions mainly act by electrostatic effects but also that polarization, charge-transfer, and dispersion contributions are all important.

In summary, these studies can show the orientation preferences around certain functional groups and then the particular chemical and biological activity might be understood by the identification of the preferred intermolecular contacts.

2.3.5 Structural Correlation to Analyze and Predict Biological Activity

Through the investigations of structural features of different active molecules, models of three-dimensional features required for a particular biological activity can be proposed.

One of the challenges for chemists and biochemists is to design totally new structurally novel compounds based on three-dimensional criteria. It has been proved that research on the computer design or recognition of molecules predicted to have a specified property may be a useful way for pursuing this purpose. Creating and searching databases of three-dimensional structures of molecules is one part of such research.

Study of protein structures is a direct way for exploring biological activity. As X-ray diffraction techniques are applied in determination of protein crystal structures, interests in the exploitation of three-dimensional protein structures have increased greatly. An important application of protein structures is to understand the geometry and strengths of the non-bonded interactions between ligands and receptors, consequently providing further information for drug design. Because of the complexity of protein structures, it is hard to predict the geometric property of non-bonded interaction from theoretical methods. Experimental information is still the main source for the molecular model. However, the number of known protein structures that contain a coordinated small molecule is quite limited. Not enough high-resolution data are available and there is little structural diversity in the ligands. Therefore, most studies to date have considered small molecules in the CSD with the interaction of ligands in a protein-like environment as an alternative (Klebe, 1990). This readily gives comparison of the geometrical preferences of the trial fragment in small molecule crystal structures with those in protein environments. In consequence, the geometrical property of intermolecular contacts, including many interactions that are difficult to model or predict theoretically could be derived (Poulos, Finzel and Howard, 1987).

In the example given in last section, the hydrogen bond O-H···O systems have similar biological environments to real protein-ligand interactions as the acceptor groups are frequently part of functional groups in protein ligands. The general modes of binding drugs to receptors can involve the geometric conformational studies and a deduced conformation of the active species. The importance is to find a suitable system as a

starting point for identifying the binding mode and thereafter the essential aspects that determine its binding affinity. Investigation of conformations of known ligand molecules can give some ideas about the basic prerequisites for receptor binding when the protein structure details are not known.

In some particular cases, especially for the nonpolar ligand retinol (Klebe, 1994), the conformations in the crystal structures are indeed very similar to that adopted at the binding site. But, some small molecule ligands with several rotatable bonds are more flexible and can have a large variety of different low-energy conformations. It is hard to say which conformation would be adopted at a binding site, without the evidence from the protein-ligand structure data. Klebe's (1995) study on the comparisons of thirty-three compounds present both in the CSD and PDB (Protein Data Bank) with global energy minimum calculations by molecular mechanics indicated that flexible molecules are deformed when binding to protein and this happened as general phenomenon. The degree of deformation is dependent upon the number of rotators in the ligands. In this way, if the target protein structures are available, the comparisons of the conformations of ligand molecules both in crystal packing and the environment of binding site certainly provide more reliable information (Böehm and Klebe, 1996) for the protein-ligand interactions, and thereafter will stimulate the discovery of novel molecules to develop new alternative binding skeletons.

Hydrogen bonds still play an important role in the protein-ligand interaction because various amino acid residues usually interact with functional groups of some molecule ligands in the hydrogen bond forms. The multiple hydrogen bonds features of water molecules connected with protein-ligand interactions (Poornima and Dean, 1995) were described by analyses of 19 high-resolution crystal structure of protein-ligand complexes. The water networks formed by these receptor/ligand interactions were considered to be providing the stability of the protein-ligand complex and are important for any site-directed drug design strategies. It is believed that with the increase in the numbers of structural data of experimental protein-ligand complexes, the studies on this area will be further developed.

References

ACLMEN, ANYCLA, BERXAX, BOLLOD, BORZOX, BRUCLA, BUMREG, CDPPMM, CFAPCH, CLACET01, CUCAPI, FABHOF, 12 hits (Refcode) for the functional group, chloride substituted acetic acid in the CSD.

Allen, F. H., *Acta Cryst.* B**37**, 900-906, (1979).

Allen, F. H., Bird, C. M., Rowland, R. S., Harris, S. E. and Schwalbe, C. H., *Acta Cryst.*, B**51**, 1068-1081, (1995).

Allen, F. H., Bird, C. M., Rowland, R. S. and Raithby, P. R., *Acta Cryst.*, B**53**, 680-695, (1997a).

Allen, F. H., Bird, C. M., Rowland, R. S. and Raithby, P. R., *Acta Cryst.*, B**53**, 696-701, (1997b).

Allen, F. H., Doyle, M. J. and Auf der Heyde, T. P. E., *Acta Cryst.*, B**47**, 412-424, (1991).

Allen, F. H., Doyle, M. J. and Taylor, R., *Acta Cryst.*, B**47**, 29-40, (1991a)

Allen, F. H., Doyle, M. J. and Taylor, R., *Acta Cryst.*, B**47**, 41- 49, (1991b).

Allen, F. H., Doyle, M. J. and Taylor, R., *Acta Cryst.*, B**47**, 50-61, (1991c).

Allen, F. H., Lommerse, J. P. M., Hoy, V. J., Howard, J. A. K. and Desiraju, G. R., *Acta Cryst.*, B**52**, 734-745, (1996).

Allen, F. H., Harris, S. E. and Taylor, R., *J. of Computer-Aided Mol. Design*, **10**(3), 247-254, (1996).

Allen, F. H., Howard, J. A. K. and Pitchford, N. A., *Acta Cryst.*, B**49**, 910-928, (1993).

Allen, F. H., Howard, J. A. K. and Pitchford, N. A., *Acta Cryst.*, B**52,** 882-891, (1996).

Allen, F. H, Howard, J. A. K., Pitchford, N. A. and Vinter, J. G., *Acta Cryst.*, B**50**, 382-395, (1994).

Allen, F. H., Kennard, O. and Taylor, R., *Acc. Chem. Res.*, **16**, 146-153, (1983).

Allen, F. H., Kennard, O. and Watson, D. G., in *Structure Correlation, Eds* by H. B. Bürgi and J. D. Dunitz, VCH, Weinheim, Chapter 3, 71-109, (1994).

Allen, F. H., Kennard, O, and Watson, D. G., Brammer, L., Orpen, A. G. and Taylor, R., *J. Chem. Soc. Perkin II*, S1, (1987).

Allen, F. H. and Pitchford, N. A., *Experimental and Computational Approaches to Structure Based Drug Design*, 71-82, Erice, Italy 9-19 May, (1996).

Allen, F. H., Garner, S. E., Howard, J. A. K. and Pitchford, N. A., *Acta Cryst.*, **B50**, 395-404, (1994).

Allen, F. H. and Taylor, R., *Acta Cryst.*, **B47**, 404-412, (1991).

Auf der Heyde, T. P. E. and Nassimbeni, L. R., *Inorg. Chem.*, **23**, 4525-4532, (1984).

Bernstein, J., Etter, M. C. and Leiserowity, L., in *Structure Correlation, Eds* by H.B. Bürgi and J. D. Dunitz, VCH, Weinheim, Chapter 11, 431-507, (1994).

Bocian, D. F., Pickett, H. M., Rounds, T. C. and Strauss, H. L., *J. Am. Chem. Soc.*, **97**, 687-695, (1975).

Bocian, D. F. and Strauss, H, L., *J. Am. Chem. Soc.*, **99**, 2876-2882, (1977).

Böham, H.-J. and Klebe, G., *Angew. Chem. Int. Ed. Engl.*, **35**, 2588-2614, (1996).

Böham, H.-J., Klebe, G., Lorent, T., Mietzner, T. and Siggel, L., *J. Comp. Chem.*, **11**(9), 1021-1028, (1990).

Cambridge Structural Database System Manual, **2**, (1992).

Cieplak, A. S., in *Structure Correlations*, *Eds.* by H. B. Bürgi and J. D. Dunitz, VCH, Weinheim, Chapter 6, 205-292, (1994).

Cremer, D. and Pople, J. A., *J. Am. Chem. Soc.*, **97**, 1354-1358, (1975).

Cross, R. J., *Chem. Soc. Rev.*, **14**, 197-223, (1985).

Desiraju, G. R., *J. Chem. Soc., Chem. Commun.*, 454, (1990).

Desiraju, G. R., *Acc. Chem. Res.*, **24**, 290-296, (1991).

Domernicano, A. and Vaciago, A., *Acta Cryst.*, **B31,** 221-234, (1975).

Domernicano, A. and Vaciago, A., *Acta Cryst.*, **B35,** 1382-1388, (1979).

Dunitz, J. D., *X-ray Analysis and The Structure of Organic Molecules*, Cornell Uni. Press, New York, USA, (1979).

Everitt, B., *Cluster Analysis, 2nd Ed.*, Halstead Heinemann, London, UK(1980).

Ferretti, V., *J. Am. Chem. Soc.*, **108**, 2420-2424, (1986).

Glusker, J. P. with Lewis, M. and Rossi, M., *Crystal Structure Analysis for Chemists and Biologists*, VCA, New York, USA, (1994).

Glusker, J. P. and Murray-Rust, P., *J. Am. Chem. Soc.*, **106**, 1018, (1984).

Harris, R. K., Howard, J. A. K., Samadi-Maybodi, A., Yao, J. W. and Smith, W., *J. Sol. State Chem.*, **120**(2), 231-237, (1995).

International Tables for Crystallography, Edited by A. J. C. Wilson, Volume C, Kluwer Acad. Publishers, Bristol, UK, 685-791, (1992).

Kanters, J. A. and Roelofsen, G., *Acta Cryst.*, **B32**, 3328-3331, (1976).

Klebe, G., *Struct. Chem.*, **1**, 597, (1990).

Klebe, G., in *Structure Correlation, Eds.* by H. B. Bürgi and J. D. Dunitz, VCH, Weinheim, Chapter 13, 543-603, (1994).

Klebe, G., *Bioorgan. & Med. Chem.*, **3**(4), 411-425, (1995).

Lommerse, J. P. M., Price, S. L. and Taylor, R., *J. Comput. Chem.*, **18**, 757-774, (1997).

Lommerse, J. P. M., Stone, A. J., Taylor, R. and Allen, F. H., *J. Am. Chem. Soc.*, **118**, 3108-3116, (1996).

Murray-Rust, P, Bürgi, H. B. and Dunitz, J. D., *J. Am. Chem. Soc.*, **97**, 921, (1975).

Murray-Rust, P, Bürgi, H. B. and Dunitz, J. D., *Acta Cryst.*, B**34,** 1787-1793, (1978).

Nyburg, S. C. and Faerman, C. H., *Acta Cryst.*, B**41**, 274-279, (1985).

Orpen, A. G., Brammer, L., Allen, F. H., Kennard, O., Watson, D. G. and Taylor, R., *J. Chem. Soc. Dalton*, S1-S83, (1989).

Pauling, L., *J. Am. Chem. Soc.*, **69**, 542, (1947).

Pitchford, N. A., *Ph.D thesis*, University of Durham, UK, (1994).

Poornima, C. S. and Dean, P. M., *J. Compt.-Aided Mol. Desig.*, **9**(6), 500-512, (1995).

Poulos, T. L., Finzel, B. C. and Howard, B. C., *J. Mol. Biol.*, **195**, 687, (1987).

Sutton, L. E., Ed., *Tables of Interatomic Distances and Configuration in Molecules and Ions*; The Chemistry Society: London; Special Publication **11** (1958) and **18** (1965).

Raithby, P. R., Shields, G. P. and Allen, F. H., *Acta Cryst.*, B**53**, 241-251, (1997a).

Raithby, P. R., Shields, G. P. and Allen, F. H., *Acta Cryst.*, B**53**, 476-489, (1997b).

Stainer, T., *J. Chem. Soc. Chem. Commun.*, 95-97, (1995a).

Stainer, T., *J. Chem. Soc. Chem. Commun.*, 1321-1326, (1995b).

Taylor, R. and Kennard, O., *J. Am. Chem. Soc.*, **104**, 5063, (1982).

Taylor, R. and Kennard, O., *Acc. Chem. Res.*, **17**, 320, (1984).

Taylor, R., Mullaley, A. and Mullier, G. W., *Pestic. Sci.*, **29**, 197, (1990).

Umland, P. and Vahrenkamp, H., *Chem. Ber.*, **115**, 3555, (1982).

VISTA - *Version 2.0 User Guide*, Cambridge Structural Database System, Cambridge Crystallographic Data Centre, UK, October, 1995.

Viswamitra, M. A., Radhakrishnan, R. Bandekar, J. and Desiraju, G. R., *J. Am. Chem. Soc.*, **115**, 4868-4869, (1993).

Zobeta, E., *Z. Kristallogr.*, **185**, 518, (1988).

Zobeta, E., *Z. Kristallogr.*, **191**, 45, (1990).

Chapter 3.

A General Method for Identifying and Classifying Metal Coordination Sphere Geometry

3.1 Introduction

In Chapter 2, the applications of the CSD are mainly concerned with organic molecular systems. The CSD also contains a lot of structural data of organometallic compounds and metal complexes. However, systematic studies on these structural parameters, especially involving transition metals, using the CSD's facilities have not been as frequent as studies in organic compounds. One of the major reasons is that the variety of metal atoms in their bonding and in these systems are usually harder to model computationally, than are the more fully understood organic systems. The same metal atom always has a wide variety of geometry, according to different bonding environments. Therefore, the previous form of systematic analysis is not always appropriate. But sometimes, investigations on the geometry of the ligands can still provide direct observations related to the structural properties (Mülker and Mingos, 1995).

A desirable systematic treatment of metal complexes is to classify and identify the geometry for various coordination numbers. As is well-known, transition metals coordinate with ligands to form different polyhedra in which ligands occupy the vertex positions and the metal atom is at the centre. Although it is not yet possible to predict which shapes of transition metal complexes can be, it is clear that among the factors which determine them are (Sharpe, 1981):

(a) the electronic configuration, oxidation state and energetically accessible orbital of the metal;

(b) high-spin or low-spin character, which in turn depends on

(c) the nature of the ligand;

(d) size and steric effects.

For example, four-coordination complexes are often classified as either symmetrical tetrahedral or square planar, but there are many complexes, which are

known to have distorted or intermediate geometry. A survey on such species could provide information which shows how the geometry of one is converted into the other, that is exploring the reaction pathways of coordination complexes as described in last chapter. Although some physical methods and theoretical calculations can often be used to assign structures of complexes, using X-ray diffraction is the unique method to determine such structures containing many atoms. Thus, the structure correlation studies may reveal information that might not have been accessible by any other methods. The studies, by using the CSD, have given a clear classification to molecular geometry of 3- (Bürgi, Murray-Rust, Camalli, Caruso and Venanzi, 1989), 4-(Murray-Rust, Bürgi and Dunitz, 1978) and 5-coordination (Auf der Hyde and Bürgi, 1989a,b,c). With the higher coordination numbers, there are more difficulties in using database facilities. In this present study, the project is concerned with the higher coordination numbers. The details about the studies and corresponding successful results will be presented in the following sections and chapters.

3.2 The General Geometric Descriptions of Metal Coordination Sphere ML_n

For a metal coordination sphere ML_n, some ideal polyhedra are usually used as references to describe the geometry. In a simple coordination complex, this can be obtained by a direct view of the determined structure. An appropriate selection of coordinates for describing the geometry also gives a rather straightforward expression. The most convenient way is to use bond angle relationships, i.e. internal coordinates. For an *n*-coordination complex, there are $n(n-1)/2$ valence angles to define the geometry. Let us look at four-coordination metal complexes again. The geometry of compounds can easily be classified as tetrahedron or square plane according to the definitions of parameters related to the angles between ligands (L) through metal (M), i.e. angles of ligand-metal-ligand, (L_i-M-L_j), θ_{ij}. These angles in an ideal tetrahedron (all 109.5°) are obviously different from those in square plane geometry, (four 90° and two 180°). A simple definition of the mean values of these angles will distinguish tetrahedral from square planar because the value for the former symmetry is 109.5° whereas for the later it is 120°. Some similar definitions for the specific angle sets for these two different geometries can also be easily found.

When the coordination number n is ≤ 6, it is here called a lower coordination number. In all these complexes, the polyhedra used to describe their geometry have the fixed values for the angles between vertex and the central point, that is, corresponding to the bond angles between ligand and metal atoms as given above. In the cases of coordination number > 6 (higher coordination number, here), most corresponding polyhedra appear no fixed characteristic angles. However, these polyhedra are references to compare with the observed structures of metal complexes, these angles are idealized in this study so that they can be used as standard for this purpose. In practice, the definition for such a standard is not unique. Before the further analyses are given, some ideal polyhedra for describing the coordination geometry of coordination number *2–9* will be illustrated in the following sections. It should be indicated that the coordination spheres in this study are only defined as those complexes of the transition metals and the coordination number is only counted as the same number as that of metal-ligand valence bonds. The complexes which the central atom has electron pairs and the electron pairs directly affect the coordination sphere geometry are not included because the interest is this study is in the bound geometry of transition metals coordination.

3.2.1 Two, Three, Four and Five-Coordination

In these coordinations, each ideal polyhedron has significant angles. It is easier to distinguish one from another by comparing either coordination sphere shapes or the angle values θ_{ij}. The simplest case is $n=2$. Only a linear shape can be adopted when two ligands (L) coordinate with a metal atom (M) in data set of this study. It is obvious that an ideal angle θ_{12} is 180° and any real two-coordination complex structures that deviate from this value can be easily found from the comparison of the determined geometric parameters.

In three-coordination, an ideal geometric description is trigonal planar with D_{3h} symmetry. But some structures can be described in a distorted form that could be a T or Y-shaped planar or a pyramid with three equal M–L distances and the metal atom out of the plane, shown as in Figure 3.1. In these cases, the symmetry lowers to C_{2v} and C_{3v}, respectively.

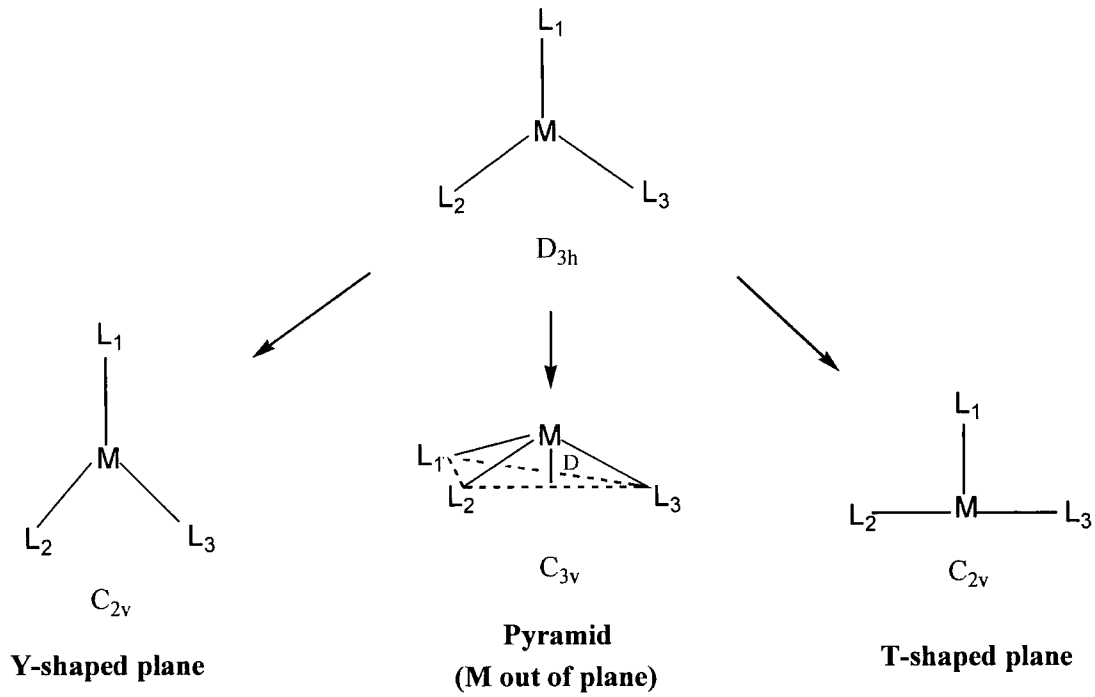


Figure 3.1 Geometrical description in 3-coordination

For the geometric description of ML_4 sphere, there exist two ideal geometries. One is a tetrahedron and the other is a square plane.

For five-coordination, there are also two major polyhedra to describe the geometry of the coordination sphere ML_5. They are trigonal bipyramid (*TBP*) and square-based pyramid (*SQP*). The significant character in the former geometry is that two vertices are located opposite one another along the axial directions and three on the equatorial plane whereas the latter one has four vertices in a plane and one along a perpendicular axis. There are obviously distinct angles to express these two shapes. Any other geometry which has appeared in the real structures of ML_5, can be considered as a distortions from either these two of modes.

Six-coordination in metal complexes is mainly confined to a common geometry, the octahedron, which has regular angular values in the *x,y* and *z* directions. This is a preferred geometry for most metal atoms according to electron configuration and orbits in the metal ion or atom. However, although it is rare, the trigonal prism (*TPR*) also exists in hexacoordination, which has higher internal ligand repulsion energy.

3.2.2 Seven-Coordination

In seven-coordination, there are three polyhedra that are called pentagonal bipyramid (*PBP*), capped octahedron (*COC*) and capped trigonal prism (*CTP*), shown as in Figure 3.2. A pentagonal bipyramid has two types of sites for ligands. Two sites are on axial positions, here numbered as *1* and *2*, and five on an equatorial plane, as *3, 4, 5, 6* and *7*. The angles between the ligands through the metal atom have fixed values. It is the easiest identified shape among the three. The positions *1* and *2* have the loosest coordination environment and an idealized L_1–M–L_2 angle of 180°. The five equatorial ligands are in the same plane and thus the ideal angle between two adjacent ligands through the central atom is 72°. This shape has D_{5h} symmetry and the five-fold axis passes through positions *1* and *2*. Unlike the PBP, the capped octahedron has no fixed values of angles between ligand through metal atom, L_i–M–L_j. The angles are variable in certain range with different ligand–metal distances. The molecule with this shape is usually characterized as C_{3v} symmetry. The three-fold axis is considered to run through from position 3 to the centre of opposite triangular face.

Capped trigonal prism is another polyhedron which L_i–M–L_j angles are also not fixed. Ideal geometry for this polyhedron is described as C_{2v} symmetry. The two-fold

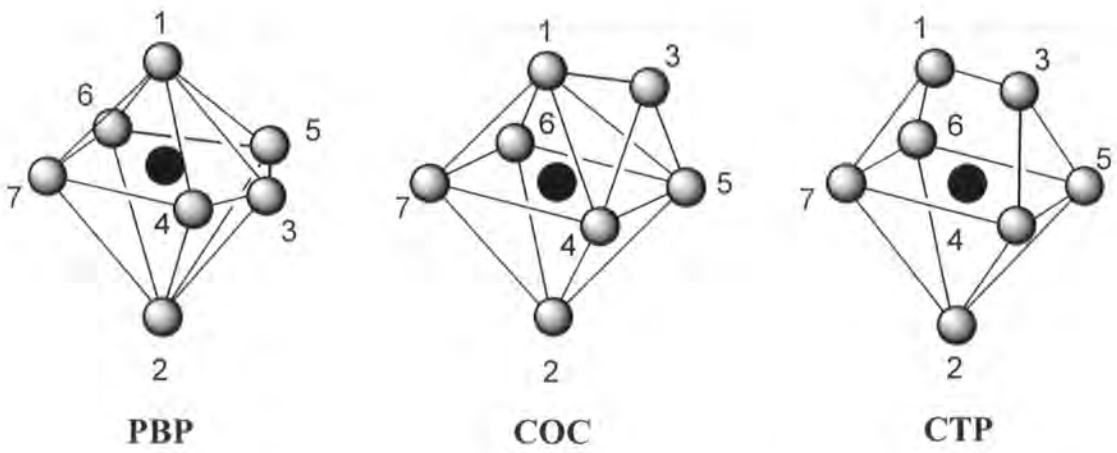


Figure 3.2 Geometrical description in 7-coordination

axis runs through from position 2 to the midpoint of edge 1, 3. Distortion of the equatorial quadrilateral face from the ideal geometry can give C_2 and C_s geometry.

COC can be seen as an octahedron plus an extra atom off one face, while CTP can be seen a trigonal prism plus an extra atom off one rectangular face. Although they appear quite different in Figure 3.2, all of these polyhedra are actually very similar when viewed in a different perspective. Especially, the angles between COC and CTP are indistinguishable sometimes owing to their very similar range of angular change. In addition to the small differences in angles and energy (Thompson and Bartell, 1968) between them, real complexes often adopt geometries which deviate somewhat from the idealized ones. Thus, the description of complex structures in terms of one geometry or another is arbitrary. The same structure could be described in different ways and confusion often happen in literature.

3.2.3 Eight and Nine-Coordination

Geometry of eight-coordination largely falls into two common geometries: dodecahedron (DOD) and square antiprism (SQUP). The cube is another possible geometry but not commonly found in this kind of complex. It is more common in ionic solids with *NaCl* or CaF_2 structures. Occasionally, bicapped trigonal prism and hexagonal bipyramid are also represented in octacoordinate complexes.

The most common polyhedra used in eight-coordination are given in Figure 3.3 accompanied by a cube for the comparison. The SQUP can be seen as a 45° twist of one square face of a cube with respect to another. And in DOD, there are two groups of atoms positioned at *1, 3, 5, 7* and *2, 4, 6, 8*, which form two tetrahedra and interlock each other. The symmetry of this geometry is $D_{2d.}$ A $\overline{4}$ axis passes through *2-4* edge and *6-8* edge and the metal atom locates at the centre.

There are two sets of geometric parameters to describe this different geometry (Kepert *et al*, 1978; Hoard, Hamor and Glick, 1968) and these definitions had been used in some work, but were restricted in their use.

In the case of higher coordination number *9*, the identification and classification between geometries becomes difficult because the angles between ligands through the metal atom are not easily distinguishable. Proper description of geometric changes is

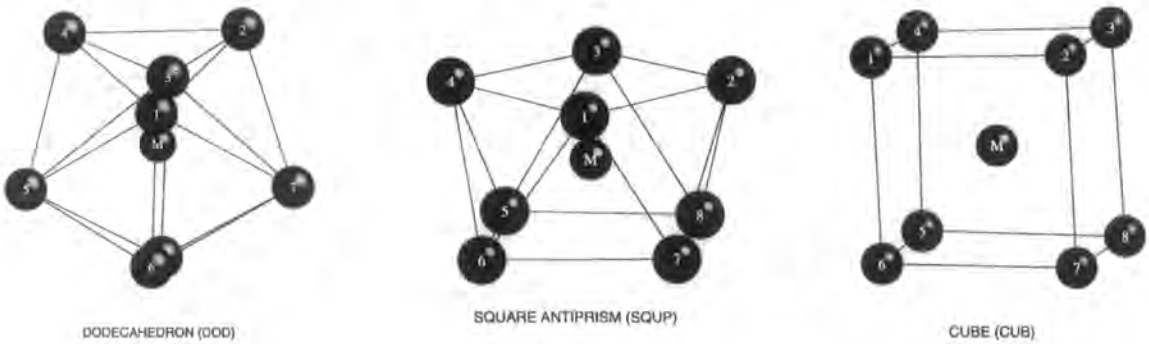


Figure 3.3 Geometrical description in 8-coordination

more complicated and ambiguous. One of the most possible geometry for 9-coordination is tricapped trigonal prism (*TTP*), the other is capped square antiprism (*CSA*), shown in Figure 3.4. In the *TTP*, ligand positions *123* and *456* form the trigonal prism and *789* at the capping positions. These three groups all occupy the vertices of an equilateral triangle, respectively, and these three triangles have mutually parallel planes. While in *CSA*, atoms *1364* and *2759* constitute two parallel squares and the two square planes are twisted about each other by *45*°. In Figure 3.4, the position numbers were assigned in such a way that two geometries can be compared easily, that is, if a view is along the bond connected by position *8* and the central point in *TTP*, it has the same number order as that in the *CSA*. In practice, *CSA* is closely related to the *TTP*. The difference between these two geometries is mainly in some specific angle values that will be given later.

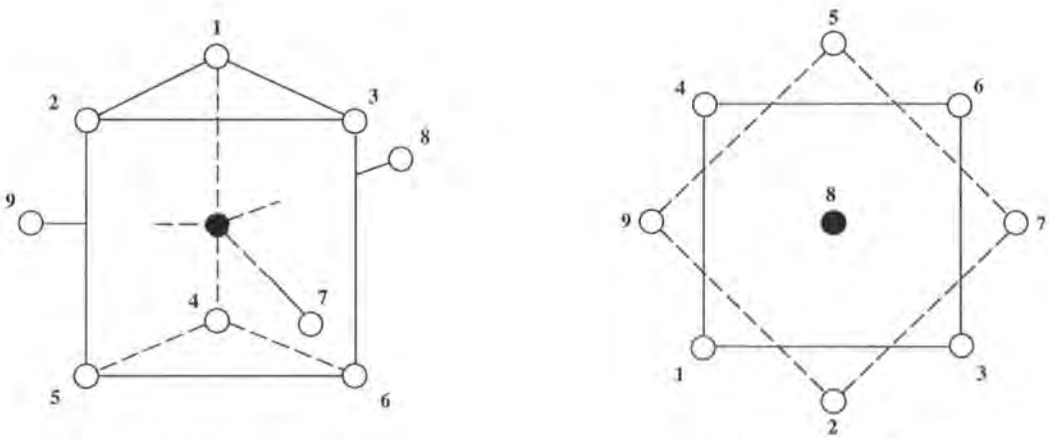


Figure 3.4 Geometrical description in 9-coordination

3.3 Application of Symmetry Deformation Coordinates

3.3.1 Symmetry Considerations

The given coordination polyhedra have in general ideal symmetry and are used as reference configurations to classify or identify the symmetry of real coordination complexes. In fact, many molecules exhibit significant deviations from these symmetries in the crystalline state. Thus, the observed points in the correlation patterns often appear around the expected value and fall in a certain limited range. Large deviations or distortions will be beyond the range. In order to describe the distortion of an observed structure away from a symmetric reference molecule with the same atomic connection or constitution, an expression in terms of a total displacement vector $\boldsymbol{D} = d_j\boldsymbol{p}_j = [d_{j(obs)} - d_{j(ref)}]\boldsymbol{p}_j$ was given (Murray-Rust, Bürgi and Dunitz, 1978), where $d_{j(obs)}$ and $d_{j(ref)}$ are internal parameters of the observed and reference molecules, respectively, and the difference represents the displacement of the observed structure along the displacement coordinate $\boldsymbol{p}_j$.

Some other alternative coordinate systems could be also used in this expression. All of these systems can be derived from the symmetry. Therefore, for any further analysis using these coordinates, the symmetry factors existing in different geometry should be considered first of all. In Table 3.1, the symmetry point groups of various

Table 3.1 Symmetry point groups of various polyhedra

Coordination	Polyhedra	Symmetry	Coordination	Polyhedra	Symmetry
3	Trigonal plane	D_{3h} (C_{3v})	7	Pentagonal bipyramid	D_{5h}
	Y-shaped plane	C_{2v}		Capped octahedron	C_{3v}
	T-shaped plane	C_{2v}		Capped trigonal prism	C_{2v}(C_2, C_s)
4	Tetrahedron	T_d	8	Cube	O_h
	Square plane	C_{2v}		Dodecahedron	D_{2d}
5	Trigonal bipyramid	D_{3h}		Square antiprism	D_{4d}
	Square pyramid	C_{2v}	9	Tricapped trigonal prism	D_{3h}
6	Octahedron	O_h		Capped square antiprism	C_{4v}(C_{2v})
	Trigonal prism	D_{3h}			

polyhedra belonging to *3-9* coordination spheres are presented. The point groups in parentheses represent the possible symmetry when the ideal shape has some distortions.

3.3.2 Symmetry Deformation Coordinates

Symmetry coordinates are a linear combination of the internal coordinates of the molecule and a set of orthogonal and normal basis vectors are formed according to the irreducible representations of the point group of the reference molecule. They are usually applied in geometric analysis in terms of investigation of the displacement vector between the defined coordinates.

As is already mentioned for the example of four-coordination, there are ten internal coordinates for a perfect tetrahedron (T_d) (four for bond length r_1-r_4; six for bond angles θ_{12}, θ_{13}, θ_{14}, θ_{23}, θ_{24} and θ_{34}). Symmetry coordinates may be given as:

$$S_1 = \tfrac{1}{2}(r_1 + r_2 + r_3 + r_4); \qquad S_{2a} = \tfrac{1}{\sqrt{12}}(2\theta_{12} - \theta_{13} - \theta_{14} - \theta_{23} - \theta_{24} + 2\theta_{34});$$

$$S_{2b} = \tfrac{1}{2}(\theta_{13} - \theta_{14} - \theta_{23} + \theta_{24}); \qquad S_{3a} = \tfrac{1}{2}(r_1 + r_2 - r_3 - r_4);$$

$$S_{3b} = \tfrac{1}{2}(r_1 - r_2 + r_3 - r_4); \qquad S_{3c} = \tfrac{1}{2}(r_1 - r_2 - r_3 + r_4);$$

$$S_{4a} = \tfrac{1}{\sqrt{2}}(\theta_{12} - \theta_{34}); \qquad S_{4b} = \tfrac{1}{\sqrt{2}}(\theta_{13} - \theta_{24});$$

$$S_{4c} = \tfrac{1}{\sqrt{2}}(\theta_{14} - \theta_{23}); \qquad S_5 = \tfrac{1}{\sqrt{6}}(\theta_{12} + \theta_{13} + \theta_{14} + \theta_{23} + \theta_{24} + \theta_{34}).$$

If the $\Delta\theta_{ij}$ values (the difference between observed and idealized values) are used instead of the observed internal coordinates only in the formula, the deformation coordinates are given.

By using this coordinate system, the deformations in the observed molecule away from the more symmetrical reference geometry can be described more accurately. It can tell us which types of distortion are expected to be coupled to other types. Especially with an increased number of atoms around a central atom, it is more difficult to use a list of internal coordinates for this description because the possible reference polyhedra are also more complicated and the differences in angular relationships become small. Thus possible overlaps of the angular ranges can occur and assignment can in fact become in-distinguishable. Whereas symmetry coordinates are considered from the symmetry of the molecule itself. "In transforming from d_j to symmetry displacements,

the actual values of $d_j(ref)$ are important only for the totally symmetric displacements since they cancel out for all the others"(Murray-Rust *et al*, 1978). In describing the molecular distortions in 4-coordination complex structures away from a reference polyhedron, tetrahedron (*TD*) and square plane (*SP*), the coordinates were calculated on the basis of the labeling permuted in the 24 different ways for four ligand atoms. The interest was particular in S_{2a} and S_{2b} (Klebe and Weber, 1994), which related to the conversion from *TD* to *SP* geometry. Through the relationship of these two coordinates, the inter-conversion from *TD* to *SP* geometry was mapped. Some other successful applications of these coordinate systems can also be found in the examples given in last chapter, such as conformation of ring systems etc.

3.4 Atomic Permutation

After the idealized polyhedra are given, identification of the geometry of an experimentally determined coordination complex structure is usually carried out by the comparison of the observed structure with those idealized polyhedra as reference geometry. For this comparison, a unique labeling scheme for the ligand atoms sited on vertices of polyhedron is required. Thus, the comparison structure is expressed on the same scale as the reference. For example, for a PBP-like 7-coordinate complex, only when the ligands at the axial positions are defined as number 1 or 2 (see Figure 3.2), deviations from the idealized angles can present their smallest values.

In a statistical study, identification of ligand number involves the permutation isomers, which will give all possibilities of the ligands over the polyhedral sites around the bound centre. Accompanied by symmetry coordinates, each such permutation isomer with the same absolute length of vectors are then classified to the same group. The isomers belonging to the group that has the shortest deformation vector are accepted to yield the final labeling scheme. It is obvious that each such group contains more than one possible labeled fragment which have the same length of deformation vectors but different in relative distribution of ligands to the polyhedra sites. For a simple example, PBP 7-coordination geometry is still chosen here. If L_i represents one of seven ligands and s_i one of seven sites of vertices (Drew, 1977), that is, L_i locates at the site s_i, and if the basal sites (*3, 4, 5, 6, 7*) are fixed, there are still two possibilities for sites *1* and *2* to

have the same values if θ_{12} is used to describe the geometry $\{ \begin{pmatrix} L_1 & L_2 & L_3 & L_4 & L_5 & L_6 & L_7 \\ s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 \end{pmatrix}$(a) and $\begin{pmatrix} L_2 & L_1 & L_3 & L_4 & L_5 & L_6 & L_7 \\ s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 \end{pmatrix}$ (b) $\}$. But for an isomer in which the axial and basal positions *1* and *3* are exchanged $\begin{pmatrix} L_3 & L_2 & L_1 & L_4 & L_5 & L_6 & L_7 \\ s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 \end{pmatrix}$ (c), different values from previous two isomers(a) and (b) will be observed in the comparison of geometric parameters related to the θ_{12} value.

For a set of *n* ligands and sites represented as:

$$\begin{pmatrix} L \\ s \end{pmatrix} = \begin{pmatrix} L_1 & L_2 & L_3 & \cdots & \cdots & \cdots & L_n \\ s_1 & s_2 & s_3 & \cdots & \cdots & \cdots & s_n \end{pmatrix},$$

there exist *n* permutation isomers to distribute labels among the *n* ligands. Thus each seven-coordination complex entry in the data set will require a large number of permutations, *5040(7!),* to assign the labeling scheme. Normally, on the basis of the final labeling scheme, comparisons with the reference polyhedral geometries can be derived and the deformations observed in the coordination complexes can be expressed in terms of symmetry deformation coordinates.

In practice, such isomers with equal deformation vector length can be considered by the order of the point group that represents the reference geometry. The point group D_{5h}, corresponding to the geometry of PBP in 7-coordination, contains 20 symmetry operations (E, $2C_5$, $2C_5^2$, $5C_2$, σ_h, $2S_5$, $2S_5^3$ *and* $5\sigma_v$). This means that for any starting arrangement of labels, 20 isomers which have the same deformation vector lengths will be generated from the operations of all these symmetry elements. Thus there are 252 (*7!/20 =252*) groups of such distortion equivalents. In order to avoid the random distribution in application of statistical technologies by choosing any one of these vectors by chance, Klebe (1994) indicated that "all examples labeled to produce the shortest deformation vector have either to be projected into the same asymmetric unit or symmetry expanded by applying all symmetry operations according to the point group of the reference structure".

In the conformational characterization of ring systems, permutations have also been used. But in the ML_n system, permutation numbers ($n!$) increase more quickly with the coordination number than those in the ring system. Permutation numbers for the ML_n

coordination sphere (n=3-9) corresponding to the ring systems (n=3-9) are listed in Table 3.2 for comparison of the two systems.

Table 3.2 Permutations of atom labels in cyclic systems and ML_n spheres

Cyclic systems			Coordination spheres	
N (atom)	n (torsion)	Permutations	N (angle)	Permutations
3	3	6	3	6
4	4	8	6	24
5	5	10	10	120
6	6	12	15	720
7	7	14	21	5040
8	8	16	28	40320
9	9	18	36	362880

3.5 Multivariate Analysis in the ML_n Coordination System

When using the CSD to carry out a systematic analysis, many hits or samples can be retrieved from the database for a defined parameter to describe the geometry of a molecule. Usually, more than one such parameter is used in the analyses of ML_n system. The defined parameters can have different values for different samples. If the parameter is referred to as a variable, a number of samples for the defined parameters will comprise the multivariate data. The data may be arranged into a matrix with one column for the values of a given variable for all samples (numbered N_s); each row for a sample having the values of N_c variables. This matrix $Q(N_c,N_s)$ could be either the same as the matrix $M(I_r,I_c)$ generated directly from the CSD for the tabulated geometrical parameters, or derived from these parameters which correspond to the selection of those data from the CSD through the defined transformations. This data matrix is the basis for all multivariate data analysis methods.

In last chapter, principal component analysis, factor analysis and cluster analysis and other multivariate analysis techniques were described briefly. They were applied to the systems which have multiple defined parameters to describe the geometries and no

knowledge about which parameter is more important in these descriptions (Murray-Rust & Motherwell, 1978a,b). All these three methods are basis techniques in statistical analysis. Before going further to see how these techniques were and will be applied in the ML_n system, we need to look at these techniques in more details.

3.5.1 Principal Component Analysis (PCA)

The multivariate matrix $Q(N_s,N_c)$ is the starting point for a mathematical description of principal component analysis. PCA can find a new set of uncorrelated variables based on the original set of variables. From the original data set, there are N_c variables, let them be $x_1, x_2, \cdots x_{Nc}$ for each observation. A variance-covariance matrix of the x_{Nc} variables can be derived as an $N_c \times N_c$ matrix, let it be **V**, a vector a_1 is found such that: $a_1^T V a_1$ is maximized with the condition that $a_1^T a_1 = 1.0$, Thus the first principal component is expressed as a linear combination of variables that gives the maximum variation: $p_1 = \sum_{i=1}^{Nc} a_{1i}x_i$; the second *PC's*, $p_2 = \sum_{i=1}^{Nc} a_{2i}x_i$ is found such that $a_2^T V a_2$ has the largest value with the conditions that $a_2^T a_2 = 1.0$ and $a_2^T a_1 = 0.0$. In this way, a complete set of *PC's* can be derived, which are linear combinations of the N_c observed variables so that they have maximum variation and are orthogonal to each other. Thus, a set of correlated variables is transformed to a new set of uncorrelated variables.

The elements of the vectors a_i are known as the principal component loadings which usually represent the relative influence of an original variable on the new component variables, and when it is applied to the systematic studies of molecular geometry, it may be important in the interpretation of the individual defined parameters.

Meanwhile, the vectors $a_1, a_2, \cdots, a_{Nc}$ can be seen as the eigenvectors of the matrix **V** from the derivations of the a_i vectors. Each eigenvector corresponds to eigenvalue, λ_i^2. Therefore, the proportion of variation for the *i*th *PC's* can be obtained by the formula:

$$\lambda_i^2 \Big/ \sum \lambda_i^2$$

For each observation, *PC's* scores are generated. These values can be used to generate the *PC* scatterplots.

In practice, the first few *PC's* may account for most of the total variance, thus reducing the dimensionality of a data set. Once the number of significant PC's is identified, the scatterplots of pairs of PC's can explore the information containing in the original data set instead of using a large number of correlated variables in the original data set.

Alternatively, the correlation matrix can also be used to find the *PC's* instead of the variance-covariance matrix stated above. There are similarity and specifics for both kinds of matrix from mathematical consideration (Chatfield *et al*, 1980), the details are not given here.

Example ML_4

An example of applying PCA to ML_4 coordination sphere geometry using the CSD is given in the following paragraphs.

Investigating on CuL_4 complexes (L = any atom except H, retrieved using the CSD search program QUEST3D), the CSD yields 127 observations. In order to decrease the chelate effect of ligands to the minimum to the coordination sphere geometry, the four ligands are all restricted to be unidentate in the search. The six L–M–L angles θ_{12}, θ_{13}, θ_{14}, θ_{23}, θ_{24}, and θ_{34} (see Figure 3.5) are defined as the basic geometric parameters. With the considerations of the symmetry and label numbering on the CuL_4, a full data matrix of $24 \times 127 = 3432$ fragments, six variable values for each fragment is obtained. PCA is applied to this data set and six *PC's* could be derived, in which the first two *PC's*, PC_1 and PC_2 are degenerate. Each PC_1 and PC_2 gave *47.2%* of the total variance and two together accounted for 94.4% of the total variance in the data set, and a PC_3 accounting for a further 3.5% (Table 3.3). This implies that only the first two *PC's*, PC_1 and PC_2, include most of the information for the CuL_4 geometrical distribution, instead of the considerations from the six valence bond angle values.

The PC_1 vs. PC_2 plot is shown in Figure 3.6, which can be viewed as if it is along the *3*-fold axis of a tetrahedron. The interpretation of the plot simply gives that the central clusters around the origin are readily attributed to tetrahedral (*TD*) complexes, while the other obvious clusters related by the $\bar{3}$ axis (at ~ -4.3, 0.0; 2.0, 3.8; and 2.0, -3.8) represent square planar (*SP*) complexes. Reference back to the CSD shows that the former arises from *Cu(I)* species and the latter from *Cu(II)*.

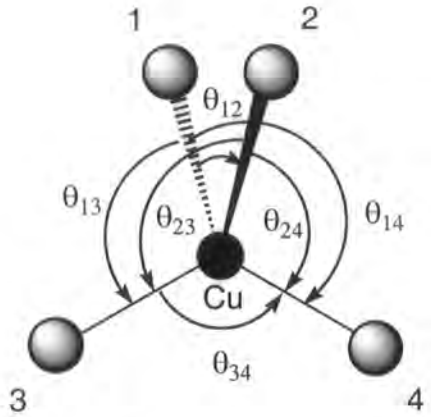


Figure 3.5 Definition of valence angles in CuL_4.

Table 3.3 PCA of CuL_4 sphere

PC's	eigenvalue	variance(%)	Culm. Variance(%)
1	2.832	47.2	47.2
2	2.832	47.2	94.4
3	0.209	3.49	97.9
4	0.042	0.70	98.6
5	0.042	0.70	99.3
6	0.042	0.70	100.0

In addition to the classification of two geometries in the CuL_4 sphere from the *PC*-plot, some other significant features related to the geometric changes in this system can be viewed from the map.One is that there are a number of the experimental observations

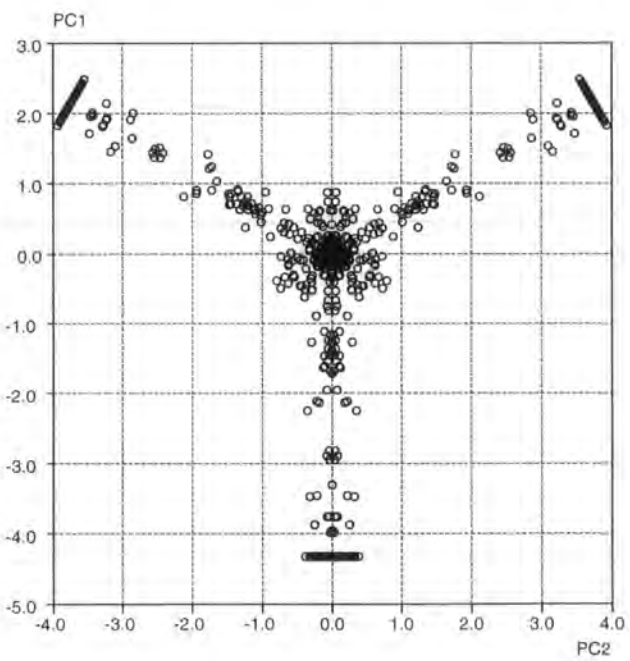


Figure 3.6 PC_1 versus PC_2 from 24-fold expansion of CuL_4(unidentate).

which appear between the two common, and expected geometries, *TD* and *SP*, which will provide important information for the mapping of the geometrical changes that take place along the *TD-SP* inter-conversion pathway. In addition, the three symmetry-related *SP* clusters reveal a range of distortions along an axis perpendicular to the inter-conversion path, while the central *TD* cluster is represented as a relatively compact circle of data points.

3.5.2 Factor Analysis (FA)

Factor analysis is another statistical technique that may be used in systematic analysis of molecular geometric preferences. FA has a similarity to PCA, but there are also several differences between these two methods from mathematical and statistical considerations.

We can still start from the original data set matrix **M**, which has N_c variables, x_1, x_2, $\cdots$, x_{Nc} for each observation. It can be expressed as a linear function of a set of underlying factors F(N_s,N_m), (N_m<N_c), and a residual variant **E**, i.e.

$$\boldsymbol{M} = \boldsymbol{F}\boldsymbol{S}^T + \boldsymbol{E}$$

Thus, each observed variable is:

$$x_i = s_{i1}f_1 + s_{i2}f_2 + \cdots + s_{iNm}f_{Nm} + e_i; \qquad i = 1,2,\cdots N_c .$$

where $\{s_{ij}\}$ $(i = 1,2,\cdots,Nc;\ j=1,2,\cdots N_m)$ are factor loadings and e_i for $i = 1,2,\cdots, N_c$, are independent random variables. From the original data set, a covariance matrix on the N_c variables can be obtained. Through decomposition by eigen analysis of the covariance matrix, the factor scores (the values of the factors for an individual entry variable) and the loading s_{ij} can be derived, in terms of a smaller number N_m, of unobserved variables.

Similarly to the PCA, it also reduces the data dimension considerably since only a few factors are significant in the describing the variances.

3.5.3 Cluster Analysis

Within a data set M(N_s,N_c), similarity exists among some objects, which form different groups such that individuals within a group are closely similar and those in different groups are dissimilar. Cluster analysis aims to classify these individuals to a set of clusters by measuring how "far apart" or how "close" pairs of individuals are. This will also depend upon the form of the N_c variables. The classification can be achieved by the measure of the distance or the similarity between two individuals in turn. The distance matrix can be constructed based on two objects, such as j and k and the elements of distance matrix, d_{jk}, are:

$$d_{jk} = \left\{ \sum_{i=1}^{Nc} D(x_{ji} / s_i, x_{ki} / s_i) \right\}^{\alpha}$$

where x_{ji} and x_{ki} are the *(j,i)*th and *(k,i)*th elements of M, s_i is a standardization for *i*th variable. D is a selected function, it could be a Euclidean distance, Euclidean square distance or absolute distance and so on (Krzanowski, 1988). In Chapter 2, it has been mentioned that modification have been used for this function to classify the geometry in six-membered ring systems. By using a select criterion on the d_{jk} values, a series of overlapping groups from individuals is produced and two clusters are merged to a larger cluster in a hierarchical form until a single cluster is achieved or *vice versa*.

The other type of cluster analysis is a non-hierarchical method, in which a given number of clusters are formed and there is no requirement that the final clusters formed are from the previous clusters. The number of clusters is usually specified prior to the analysis. One of the commonly used methods is K-means clustering (Gordon, 1981).

All of these are based on the measure of similarity of the objects. The classification can be also carried out using discriminate analysis (Chatfield *et al*, 1980).

Both cluster analysis and FA have been combined for use in the geometric analysis of five-coordination d^8 metal complexes (Auf der Heyde and Bürgi, 1989b,c). Analyses were based on the symmetry deformation coordinates, which are represented by two sets of *12*-dimensional deformation spaces for *TPB* (D_{3h}) and *SQP* (C_{2v}), respectively (Auf der Heyde and Bürgi, 1989a). Cluster analysis showed that ideal *TBP* fragments

appeared in the centre of scatterplot diagram represented by two specified coordinate variables corresponding to the D_{3h} symmetry. Three other identical clusters which represented the distorted *SQP's* geometries surrounded the centre. Two types of distorted *SQP* geometries, conventional elevated and flattened *SQP*, were characterized in the centre and the bottom tip of the Y-shaped clustering pattern, respectively, which is one of the scatterplots based on the deformation coordinates derived from the *SQP* (C_{2v}) symmetry. The other two clusters, which belong to permutationally equivalent *TBP's*, were also identified from the diagram. FA generated a group of factors also on the basis of symmetry coordinates. Mappings of the results showed that the distortions of the TBP and *SQP* follow an S_{N2} and Berry coordinate reaction pathway, that is, from tetra- to penta-coordination with attack of the fifth ligand from one of axial directions. Progression is from a flattened *SQP* to an elevated *SQP*, then transformation from the *SQP* to *TBP* and finally elimination of ligand at the opposite axial position to the attack site, according to the relative importance of symmetry coordinates.

A question may be raised from these examples as: Can these techniques be used directly in the higher coordination number system in a similar manner? The facilities provided in the CSD for the application of symmetry-expanded PCA to *127* ML_4 fragments present no real computational problem. However, with the increase of the coordination number, the base of defined geometric parameters becomes larger. For seven-coordination, *21* valence *L–M–L* angles are needed compared with only 6 such angles in four –coordination. In this way, a huge number of multivariate data points will be generated by symmetry expansion for hundreds of the CSD retrieved experimental observations and each observed fragment will have *5040* points from Table 3.2. Such data set matrices definitely exceed the storage capacity of the CSD data analysis program VISTA. So the present database software can not be used simply as in the lower coordination systems. A method to address the problem so that the application of PCA or FA technique can be achieved in 7-coordination will be given in the following sections and related results in later chapters.

3.6 A Discrepancy Index Method for Higher Coordination Number ML_n

From previous sections, it can be seen that statistical techniques, e.g. *PCA*, would appear to be a useful method for detecting the geometrical preferences, deformations and inter-conversion pathways of metal coordination spheres.

However, in a systematic study of *n*-coordination, a search from the *CSD* will always result in a significantly large data set, represented by a multivariate valence angle matrix with row numbers of $(n! \times N_s)$ and column numbers of $k = n(n-1)/2$ from a consideration of the full number of atomic permutational isomers. Such a data set becomes unfeasible for a *2D* substructure search which assigns "any atom type" to ligand atoms, *L*, when the coordination number is greater than five. Thereafter the application of systematic PCA studies to general higher coordination polyhedra is also not feasible. The use of 3D geometrical search constraints (as stated in introduction section on the *CSD*) by either differentiating between examples of standard geometries, (e.g. giving angular ranges for *TBP* and *SQP* in ML_5 species), or restricting ligand atoms to the common types, will naturally reflect the user's preconceptions concerning permissible arrangements for a given archetype. Thus, this will eliminate heavily distorted variants and also examples of other possible archetypal polyhedra. In any general systematic study using the *CSD*, a full data set including every possibility for the given system is usually expected, and the meaningful results are explored as a knowledge acquisition process that does not depend on the preconceptions of the user. Only in this way, can new knowledge be discovered.

3.6.1 $R_{ang}(x)$ Index

By now, the well-known polyhedra with standard geometries for *n*-coordination (*3-9*) have been given, some examples of geometrical descriptions based on these polyhedra for lower coordination complexes (up to 5-coordination) have been demonstrated and some difficulties for such similar treatments in the case of higher coordination have been also stated. In order to avoid these difficulties or sort out the problems, a successful method which is based on the standard geometrical models, especially for 7-coordination, is proposed. Then, a suitable treatment is given so that

the geometrical distortions of the real complexes from these standard models can be viewed.

In this study, exemplified for 7-coordination, a model was proposed based on the three idealized polyhedral forms given in §3.2.2. A simple discrepancy index for all L-M-L angles is set out in the following Equation to measure how far a real observed complex deviates from the idealized model:

$$R_{ang}(x) = 100 \cdot \min\left\{ \frac{\sum_i \left[\theta_{i(obs)} - \theta_{i(std)}\right]^2}{\sum_i \theta_{i(std)}^2} \right\}_{N_p}^{\frac{1}{2}}$$

where $\theta_{(std)}$ represents the standard angle value in an idealized polyhedron, $\theta_{(obs)}$ is the relative angle from an observed complex. R_{ang} values are calculated over all $N_p = n!$ ligand permutation for the different polyhedral forms x and the R_{ang} value is expressed in the convenient form of a percentage. It is obvious that the smaller the R_{ang} value, the closer to the proposed idealized polyhedron is the shape of the observed complex. Therefore, the final decision for the geometry of the complex is one with the smallest R_{ang} value related to the standard model. The development in this system is to generate a more automatically and widely applicable program for use in the higher coordination number systems.

3.6.2 A Program for Calculation of $R_{ang}(x)$ Values

A formula for the calculation of discrepancy index R_{ang} has been given in last section. This is a basic way to identify geometry of an n-coordination sphere. From the criterion value of R_{ang}, a deviation of the complex from the standard geometrical form can be measured numerically. For the different standards (*PBP, CTP, SQP*, etc), here defined as x, a set of discrepancy indices $R_{ang}(x)$ can be obtained. In the construction of a computer program, an important factor is that the atomic permutational symmetry of the ML_n system should be considered so that $R_{ang}(x)$ must be computed separately for all possible set of angles $\theta_{(obs)}$, that is, the $n!$ permutations of $\theta_{i(obs)}$ [$i=n(n-1)/2$]. The *24* permutations for ML_4 are shown in Table 3.4 as an example. In doing so, these valence

angle permutations are derived and applied automatically and the minimum value of $R_{ang}(x)$ is generated.

Normally, for an *n*-coordination (n > 2), there exist two or more commonly observed archetypal polyhedra, e.g. the *PBP* and *CTP* variants of ML_7 geometries, and the discrepancy index of a given experimental geometry from both of these standards, $R_{ang}(TBP)$ and $R_{ang}(CTP)$ can be computed and listed in an output file for further use in systematic analyses.

An output for an example of a 7-coordinate complex (CEHZEU) resulting from the computer program is listed in Table 3.5. The complex is identified by the CSD refcode and with three $R_{ang}(x)$ values (expressed as percentage), the corresponding *21*

Table 3.4 The 24 permutations of atom labels in an ML_4 fragment

Permutation	1	2	3	4	Permutation	1	2	3	4
1	1	2	3	4	13	3	1	2	4
2	1	2	4	3	14	3	1	4	2
3	1	3	2	4	15	3	2	1	4
4	1	3	4	2	16	3	2	4	1
5	1	4	2	3	17	3	4	1	2
6	1	4	3	2	18	3	4	2	1
7	2	1	3	4	19	4	1	2	3
8	2	1	4	3	20	4	1	3	2
9	2	3	1	4	21	4	2	1	3
10	2	3	4	1	22	4	2	3	1
11	2	4	1	3	23	4	3	1	2
12	2	4	3	1	24	4	3	2	1

experimental valence angles are illustrated in a certain order. The program identifies the specific permutation of the observed valence angles that is closest to the fixed values used to define the standard archetype (*x*) and by default, the specific enumeration of the atoms of the experimental CSD structure that gave rise to those angles. In this example, *P2, P1, Cl1, S4, S2, S3, S1* are listed in the required order to enumerate ligand labels *1* to *7* in *PBP* geometry (shown in Figure 3.7). It is obvious that the geometry of this structure is identified as *PBP* since $R_{ang}(PBP)$ has the smallest value among three

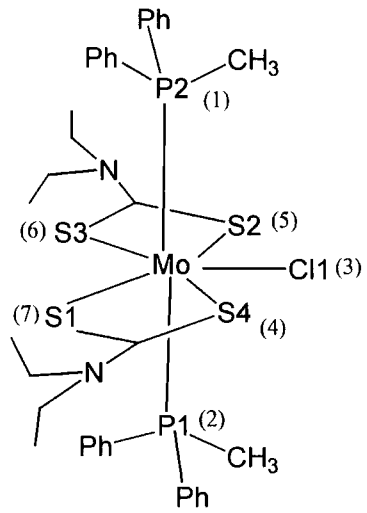


Figure 3.7 Connectivity diagram of the coordination sphere in CEHZEU

Table 3.5 An example of output from the program (ML_7)

```
.....
CEHZEU
The smallest R-index in PBP is %:  3.64;  permutation number is : 838
The Order of Ligands :      P2    P1    Cl1    S4    S2    S3    S1
Angles(°) :    176.5
 86.9     86.6     97.7     88.5     93.2     90.1     90.8     83.5     95.0     88.1
 75.7     76.0    143.1    144.5    151.1    140.5     68.9     68.4    138.8     72.3

The smallest R-index in COC is %: 13.24;
The smallest R-index in CTP is %: 11.38;
The Order of Ligands in CTP:   S3    Cl1    S1    P1    S4    P2    S2
Angles(°) :    143.1
 72.3     95.0    140.5     88.5     68.4    144.5     90.1     75.7     86.9     76.0
 88.1     68.9     93.2    138.8     90.8    176.5     83.5     86.6    151.1     97.7

The Order of Ligands in COC:   S2    S4    S3    S1    P2    Cl1    P1
Angles (°) :  151.1
 68.4    138.8     97.9     76.0     83.5    140.5     68.9     86.6     75.7     90.8
 72.3     88.5    143.1     95.0     93.2    144.5     88.1     86.9    176.5     90.1
.....
```

$R_{ang}(x)$ values from the output.

Finally, the design of the program is expressed in a flow chart shown in Figure 3.8.

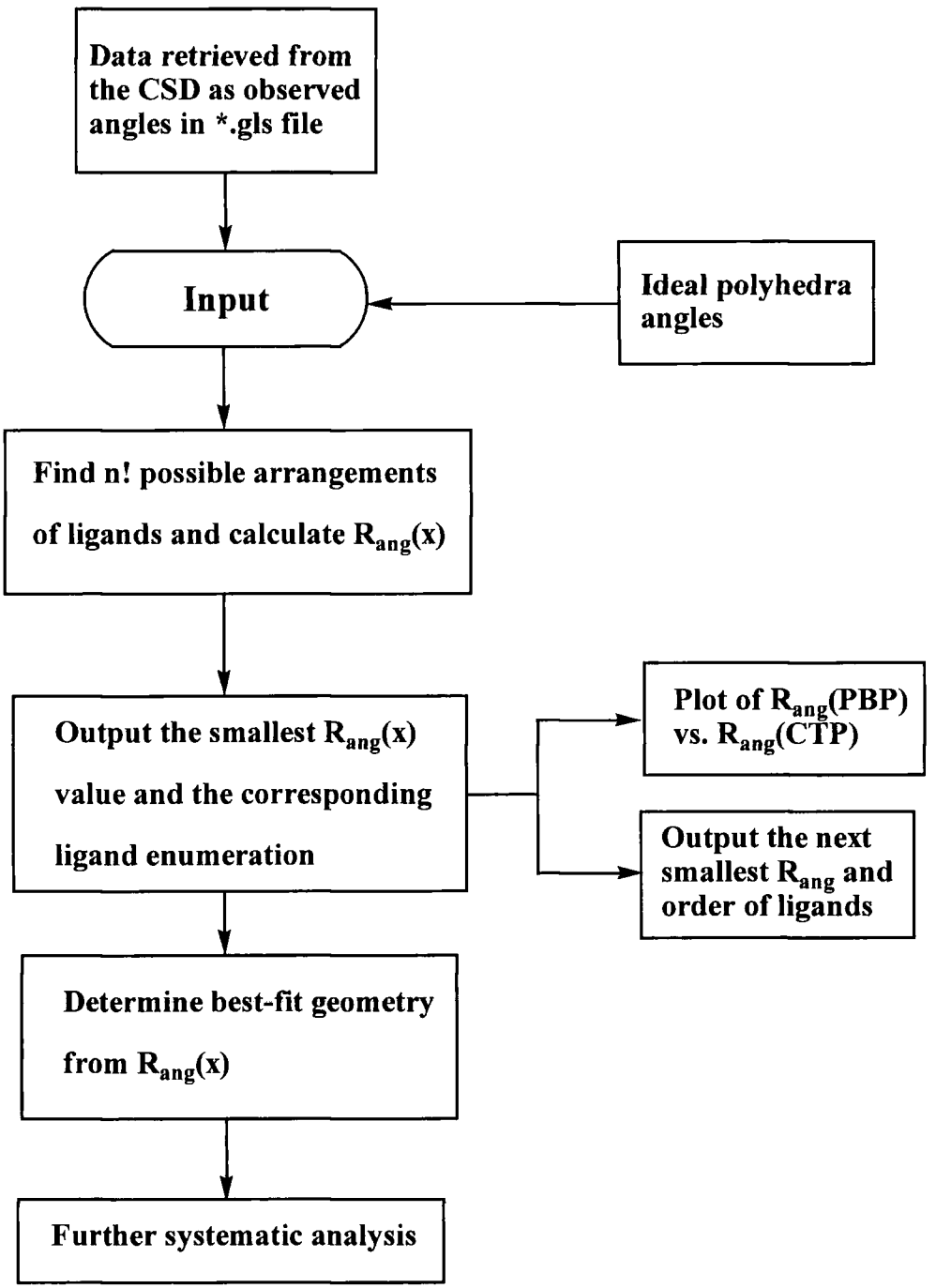


Figure 3.8 Flow chart of the procedure

3.6.3 The Standard Angles for the Idealized Polyhedra in Seven-Coordination

The standard geometries represented by various polyhedra are given previously as is the basis for setting up the program to calculate the discrepancy index $R_{ang}(x)$ values in this study. It is necessary to give a set of standard angle values for each idealized reference polyhedron. In seven-coordination, for the most commonly used archetypes, *PBP*, *COC* and *CTP* (as shown in Figure 3.2), it is not difficult to determine the angle values in the case of *PBP* because the symmetry fixes all the angles between *M–L* bonds. However, attention should be paid to the *COC* and *CTP* modes, since there are no such fixed angles for both shapes. If certain restrictions are specified, a set of limited angle values in these two symmetry can be calculated, which are listed in the columns *4* and *6* of Table 3.6, respectively. They are calculated based on the fact that all M-L distances are equal and two vectors formed by basal sites, *4, 5, 6, 7*, i.e. L_4–M–L_6 and L_5–M–L_7, are perpendicular to each other for *CTP*. In practice, there are no real complexes that can adopt such angles to form a stable coordination. In Drew's work (1977), the angles found in experimental complexes $[W(CO)_4Br_3]^-$ and $[Mo(CNR)_7]^{2+}$ were quoted as reference angles for the *COC* and *CTP* symmetry, respectively. A preliminary comparison of the observed complexes from the *CSD* with these two sets

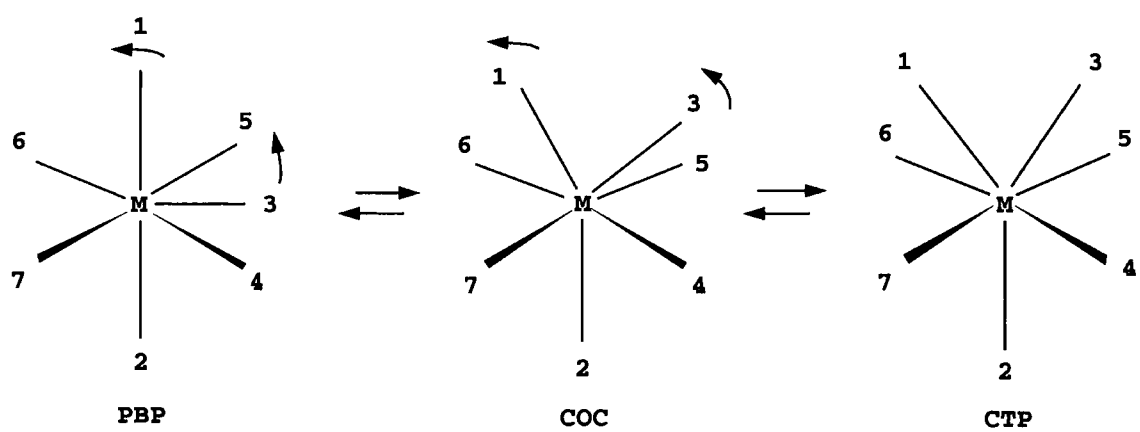


Figure 3.9 The PBP⇔COC⇔CTP interconversion pathway: atomic movements in 7-coordination, atoms 1 and 3 are moved most in this distortion

Table 3.6 Standard angle values for PBP and CTP models and some reference values

Geometry	PBP	COC		CTP		
Angle	STDAG	REFAG	CALAG	REFAG	CALAG	STDAG
L_1-M-L_2	180.0	160.4	180.0	144.2	135.0	141.8
L_1-M-L_3	90.0	74.1	54.7	71.5	90.0	76.4
L_1-M-L_4	90.0	112.8	90.0	118.9	138.6	127.3
L_1-M-L_5	90.0	112.8	90.0	118.9	138.6	127.3
L_1-M-L_6	90.0	76.6	90.0	75.1	75.5	76.4
L_1-M-L_7	90.0	76.6	90.0	75.1	75.5	76.4
L_2-M-L_3	90.0	125.3	125.3	144.2	135.0	141.8
L_2-M-L_4	90.0	76.6	90.0	82.0	90.0	76.4
L_2-M-L_5	90.0	76.6	90.0	82.0	90.0	76.4
L_2-M-L_6	90.0	89.6	90.0	82.0	90.0	76.4
L_2-M-L_7	90.0	89.6	90.0	82.0	90.0	76.4
L_3-M-L_4	72.0	74.1	54.7	75.1	75.5	76.4
L_3-M-L_5	72.0	74.1	54.7	75.1	75.5	76.4
L_3-M-L_6	144.0	125.5	125.3	118.9	138.6	127.3
L_3-M-L_7	144.0	125.5	125.3	118.9	138.6	127.3
L_4-M-L_5	144.0	112.8	90.0	99.9	90.0	87.7
L_4-M-L_6	144.0	160.4	180.0	164.0	180.0	152.7
L_4-M-L_7	72.0	76.6	90.0	78.7	90.0	85.9
L_5-M-L_6	72.0	76.6	90.0	78.7	90.0	85.9
L_5-M-L_7	144.0	160.4	180.0	164.0	180.0	152.7
L_6-M-L_7	72.0	89.6	90.0	99.0	90.0	87.7

STDAG - Standard angles used in the models;
REFAG - Values from literature (Drew, 1977);
CALAG - values calculated from some restrictions.

of angles as the standard angles showed that most complexes which do not belong to the *PBP* symmetry would fit well with these angles and give small R_{ang} values. But for some complexes, the difference between these angles is so small that it is hard to say to which model they belong.

In fact, these three idealized polyhedra can be inter-converted via ligand motion or edge stretching. One of the reaction pathways can be considered as follows and is shown in Figure 3.9. In this inter-conversion, starting from *PBP*, one of five ligands in the basal plane, atom 3 moves out of this plane towards the apical ligand *1*. Meanwhile, ligand *1* moves away from axial position and ligand *4, 5, 6, 7* relocate a little and may not be

coplanar, which transfers the geometry from the *PBP* to the *COC*. Further movement of the ligands *1* to *3* until they occupy sites symmetrically located above a plane formed by the ligands *4, 5, 6, 7,* will finally result in *CTP*. The changes between *COC* and *CTP* are small and *COC* can be seen as an intermediate on the inter-conversion path from *PBP* to *CTP*. Therefore, in this study only two models are used to derive the relative $R_{ang}(x)$. One is *PBP*, the other is *CTP-like* geometry and its angles are obtained by an optimization technique. In this way, although two standard reference geometries are defined, a *COC*-like geometry can still be viewed as an intermediate through the reaction pathway from PBP to CTP and R_{ang} values can be used here to investigate the changes of geometries. This will avoid the confusion in the identification between the *COC* and *CTP* that have appeared in the literature. The details will be given later. Table 3.6 lists the standard values of angles for the *PBP* and *CTP* archetypes used in this study, accompanied with some literature values from Drew's work.

3.6.4 The Standard Angles for the Idealized Polyhedra in Eight-Coordination

The standard angles for eight-coordination are chosen for the three major geometries dodecahedron (*DOD*), square antiprism (*SQUP*) and cube (*CUBE*). Also those angles for first two geometries are variable in certain ranges for different coordinate environments. In particular, the steric constraints of ligands force the angles to distort from each other. For an average effect, a hard sphere model (Gillespie and Hargittai, 1991) is used, in which all ligand atoms (*L*) are considered as equal and unidentate and thus sit at the surface of the sphere. The metal atom (*M*) is at the centre and all *M-L* distances have the same length. Although this is not always a real case for most of the experimental complexes in higher coordination, they represent a set of idealized angles for the defined geometry. While the comparisons of the discrepancy index $R_{ang}(x)$ values on these standard angles for different geometries are carried out on the whole angle set, the deviation of angles from the standard ones due to the chelate groups in the ligand fragments can be averaged.

The derived standard values for three geometries of eight-coordination are listed in the column headed STDAG of Table 3.7. The experimental angle values from three

Table 3.7 Standard angles for geometries in eight-coordination

GEOMETRY	CUBE		SQUP		DOD	
Angle	STDAG	REFAG[1]	STDAG	REFAG[2]	STDAG	REFAG[3]
L_1-M-L_2	70.5	67.6	74.9	71.7	73.7	72.8
L_1-M-L_3	109.4	101.0	118.5	116.4	138.9	139.7
L_1-M-L_4	70.5	62.6	74.9	75.3	73.7	74.4
L_1-M-L_5	109.4	115.1	142.6	144.0	97.1	95.8
L_1-M-L_6	70.5	74.7	74.9	78.5	73.7	75.1
L_1-M-L_7	109.4	109.6	74.9	77.2	97.1	97.9
L_1-M-L_8	180.0	175.0	142.6	139.9	147.4	145.2
L_2-M-L_3	70.5	63.7	74.9	73.5	73.7	74.4
L_2-M-L_4	109.4	101.3	118.5	113.0	73.7	70.3
L_2-M-L_5	180.0	175.3	142.6	142.7	147.4	145.2
L_2-M-L_6	109.4	109.2	142.6	144.0	126.8	130.2
L_2-M-L_7	70.5	75.3	74.9	78.7	73.7	75.1
L_2-M-L_8	109.4	114.7	74.9	77.5	126.8	133.8
L_3-M-L_4	70.5	70.9	74.9	71.9	73.7	72.8
L_3-M-L_5	109.4	111.7	74.9	77.5	97.1	97.9
L_3-M-L_6	180.0	172.9	142.6	139.9	147.4	145.2
L_3-M-L_7	109.4	112.9	142.6	142.1	97.1	95.8
L_3-M-L_8	70.5	76.8	74.9	77.2	73.7	75.1
L_4-M-L_5	70.5	77.3	74.9	78.2	73.7	75.1
L_4-M-L_6	109.4	111.2	74.9	77.2	126.8	133.8
L_4-M-L_7	180.0	172.1	142.6	144.3	147.4	145.2
L_4-M-L_8	109.4	112.4	142.6	142.1	126.8	130.2
L_5-M-L_6	70.5	75.4	74.9	71.7	73.7	74.4
L_5-M-L_7	109.4	106.7	118.5	113.0	138.9	139.7
L_5-M-L_8	70.5	62.3	74.9	73.5	73.7	72.8
L_6-M-L_7	70.5	64.2	74.9	75.3	73.7	72.8
L_6-M-L_8	109.4	108.0	118.5	116.4	73.7	70.3
L_7-M-L_8	70.5	75.4	74.5	71.9	73.7	74.4

STDAG — Standard angles used in the program.
REFAG — Reference angles from observed structures of the CSD.
REFAG[1] — TBPYUR; REFAG[2] — ACACNP; REFAG[3] — LASYAF.

complexes in the *CSD*, which have geometry close to one of the standard geometries, are also given (in the column headed REFAG). All these complexes are coordinated with bidentate ligands, some angles obviously deviate from the standard ones, but the

basic geometry can be easily identified from the $R_{ang}(x)$. Table 3.8 gives the calculation results of $R_{ang}(x)$ values on these three complexes.

Table 3.8 $R_{ang}(x)$ values(%) for three complexes in the CSD as an example.

REFCODE	R_{ang}(CUBE)	R_{ang}(SQUP)	R_{ang}(DOD)
TBPYUR	4.89	21.17	17.48
ACACNP	20.44	2.56	8.04
LASYAF	19.45	8.54	2.42

3.6.5 The Standard Angles for the Idealized Polyhedra in Nine-Coordination

Similarly, in nine-coordination, standard geometries are also based on the hard sphere model, two geometries, tricapped trigonal prism (*TTP*) and capped square antiprism (*CSA*) are given in Table 3.9.

Table 3.9 Standard angles for the geometries of nine-coordination

Geometry	TTP							
$\theta_{12} \Rightarrow \theta_{19}$	70.5	70.5	96.4	141.1	141.1	131.8	70.5	70.5
$\theta_{23} \Rightarrow \theta_{29}$	70.5	141.1	96.4	141.1	70.5	131.8	70.5	
$\theta_{34} \Rightarrow \theta_{39}$	141.1	141.1	96.4	70.5	70.5	131.8		
$\theta_{45} \Rightarrow \theta_{49}$	70.5	70.5	131.8	70.5	70.5			
$\theta_{56} \Rightarrow \theta_{59}$	70.5	70.5	131.8	70.5				
$\theta_{67} \Rightarrow \theta_{69}$	70.5	70.5	131.8					
$\theta_{78} \Rightarrow \theta_{79}$	120.0	120.0						
θ_{89}	120.0							

Geometry	CSA							
$\theta_{12} \Rightarrow \theta_{19}$	70.1	83.3	83.3	137.6	140.1	137.6	70.1	70.1
$\theta_{23} \Rightarrow \theta_{29}$	70.1	137.6	108.6	137.6	70.1	125.7	70.1	
$\theta_{34} \Rightarrow \theta_{39}$	140.1	137.6	83.3	70.1	70.1	137.6		
$\theta_{45} \Rightarrow \theta_{49}$	70.1	83.3	137.6	70.1	70.1			
$\theta_{56} \Rightarrow \theta_{59}$	70.1	70.1	125.7	70.1				
$\theta_{67} \Rightarrow \theta_{69}$	70.1	70.1	137.6					
$\theta_{78} \Rightarrow \theta_{79}$	125.7	108.6						
θ_{89}	125.7							

It can be seen that the differences in angles between these two standard models become smaller, and thereafter it is more difficult to identify precisely the geometry from individual angles. Also it is obvious that R_{ang} values will be closer for both standard geometries, but an identification is still expected, since there are larger differences on some specified angles than others and some extra criteria will be used to assist in identifying these two different geometries.

References

Auf der Hyde, T. and Bürgi, H. B., *Inorg. Chem.*, **28**, 3960-3969, (1989a).

Auf der Hyde, T. and Bürgi, H. B., *Inorg. Chem.*, **28**, 3970-3981, (1989b).

Auf der Hyde, T. and Bürgi, H. B., *Inorg. Chem.*, **28**, 3982-3989, (1989c).

Bürgi, H. B., Murray-Rust, J., Camalli, M., Caruso, F. and Venanzi, L. M., *Helv. Chim. Acta,* **72**, 1293-1300, (1989).

Chatfield, C. and Collins, A. J., *Introduction to Multivariate Analysis*, Chapman and Hall, London, UK, (1980).

Drew, M. G. B., *Prog. Inorg. Chem.*, **23**, 1, (1977).

Gillespie, R. J. and Hargittai, I., *The VSEPR Model of Molecular Geometry*, Allyn and Bacon, Massachusetts, USA, (1991).

Murray-Rust, P., Bürgi, H. B. and Dunitz, J., *Acta Cryst.*, B**34**, 1787-1793, (1978).

Murray-Rust, P. and Motherwell, S., *Acta Cryst.*, B**34**, 2518-2526, (1978a).

Murray-Rust, P. and Motherwell, S., *Acta Cryst.*, B**34**, 2534-2546, (1978b).

Mülker, T. E. and Mingos, D. M. P., *Transition Met. Chem.*, **20**, 533-539, (1995).

Hoard, J. L., Hamor, T. A. and Glick, M. D., *J. Am. Chem. Soc.*, **90**, 3177, (1968).

Kepert, D. L., Raston, C. L., White, A. H. and Winter, G., *Aust. J. Chem.*, **31**, 757, (1978).

Kepert, D. L., *Prog. Inorg. Chem.*, **25**, 41, (1979).

Klebe, G., *J. of Mol. Struc. (Theochem.)*, **308**, 53-89, (1994).

Klebe, G. and Weber, F., *Acta Cryst.*, B**50**, 50-59, (1994).

Sharpe, A. G., *Inorganic Chemistry*, Longman, London, UK, (1981).

Thompson, H. B. and Bartell, L.S., *Inorg. Chem.*, **7**, 488, (1968).

Chapter 4.

Systematic Studies of Geometry of the Metal Coordination Sphere ML_n Using Crystallographic Data

4.1 Introduction

The details of the procedure used in this systematic geometry study have been given in Chapter 3. The results from the use of this simple and general method on the known structures of *n*-coordination (*n* = *3-9*) complexes retrieved from the CSD are presented in this Chapter. These studies reveal their geometrical preferences, the deformations from the given idealized polyhedral geometries and the interconversions between polyhedral forms (Howard, Copley, Yao and Allen, 1998). The work not only provides a simpler way to express geometrical characteristics of coordination sphere in a systematic analysis, but also addresses the problems of dimensionality and of permutational complexity for higher coordination numbers (*n* =*7-9*). The results to be detailed in this Chapter will be emphasized for these higher coordination species, which can be considered as an update to general systematic studies in coordination chemistry. Also, those results for the lower coordination complexes are illustrated here to compare with some previous work (Murray-Rust, 1982, Taylor & Allen, 1994; Klebe & Weber, 1994). Hopefully, all these successful results which show geometrical insights for the coordination sphere, ML_n, will provide useful information towards knowledge engineering in coordination chemistry.

4.2 Transition Metal Seven-Coordination

In coordination chemistry, the configuration of *d* electrons in the transition metal elements, or the symmetry of *d* orbital with interaction of ligand orbitals, is usually used to predict the properties of coordination compounds. From this knowledge, some theory has been well established in the explanation and interpretation of how a coordination complex is formed, according to the number of *d* electrons and their arrangement in the relevant orbitals. This can be applied effectively to octahedral and

tetrahedral or square planar coordination geometries, corresponding to coordination numbers 6 and 4 because, for these two coordination numbers, preferred configuration depends only on the symmetry (Cotton & Wilkinson, 1988). In other coordination numbers, the optimum configuration is not determined by symmetry alone. But similarly, *d* orbitals can be also considered as split into a lower set and an upper set. Figure 4.1 illustrates *d* orbitals in 7 coordination for three different geometrical shapes (Speer *et al*, 1968). Compared with the orbitals in octahedra, it is obvious that the energy level difference between orbitals in a specific shape and between different shapes become smaller. t_{2g} and e_g are further split into e''_1, e_2' and a_1' in PBP, b_2, a^+, b_1 and a^-, a_2 in CTP. Thus, it is more difficult to predict which geometrical form is more preferred simply from the electron configuration.

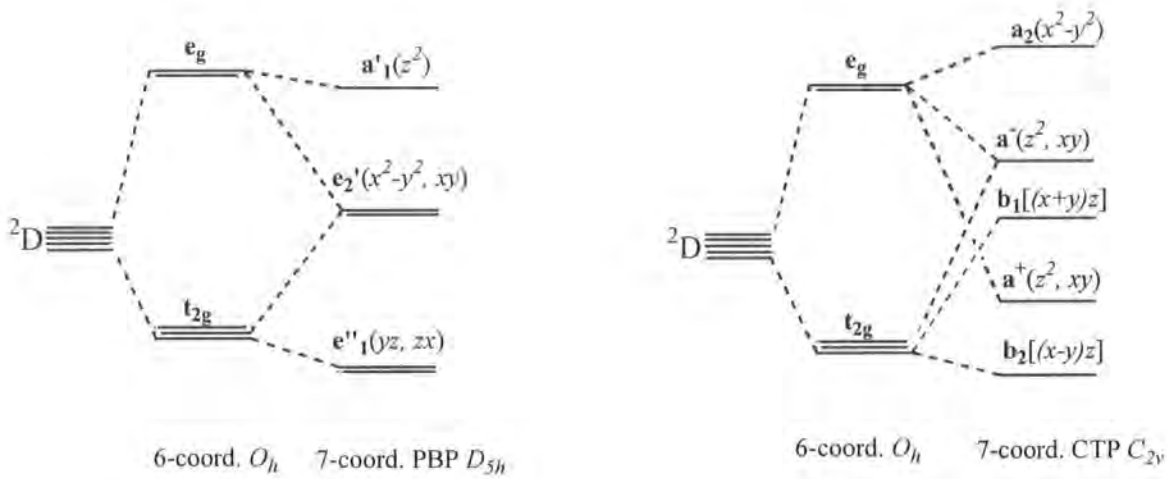


Figure 4.1 Energy level diagram for the d orbitals in 7-coordination field compared with an octahedral field in 6-coordination.

However, the metal oxidation states and the stereochemical requirements of the chelating ligands are probably the decisive factors for a metal to form a complex in this coordination number. There are no any reliable arguments in predicting the geometry of a particular complex.

4.2.1 Database Retrieval

Seven coordination is another common coordination number in transition metal complexes. The structural data of seven-coordination transition metal complexes were retrieved from the CSD (updated version 1995) to examine the geometrical forms of real structures. The general chemical search facilities of QUEST were used. The chemical fragment to be examined was defined as a central atom, either all transition metals TR or specific individual metal M, and connected to seven non-metal elements of the p-block. The coordination number at the metal was restricted to be exactly seven by use of the command TOTAL-COORD-NO in 2D-CONSTRAIN. For the dataset in which all seven ligands are unidentate, the further 2D-constrain, NO-CYCLIC-ROUTES and NO-LINKS are applied to the ligand atoms. Finally, for each entry, the geometrical parameters, bond lengths and bond angles were recorded as well as atomic coordinates for the further calculations. In addition, the search was also restricted by use of the CSD secondary search criteria (CSD User Manual, 1994) to ensure the structural data have high accuracy, that is, (a) error-free for all retrieved entries from the CSD; (b) no disorder in the chemical structure; (c) perfect consistancy between their chemical and crystallographic connectivity representations; (d) atomic coordinates available; (e) $R \leq 10\%$ for structures; (f) polymeric structure not formed and (g) "organometallic" compounds according to the CSD chemical class assignment.

The CSD search using these definitions and restrictions revealed 460 "hits" including various transition metals in a wide variety of coordinated environments. But some of these have very strained or bulky ligands, such as, boron cages complexes. Such complexes deviate greatly from any idealized polyhedra due to the steric effect of these ligands and were rejected from the dataset. The final number of fragments used for the further calculation and analysis for various metals is listed in Table 4.1.

It can be seen that seven-coordinate complexes can be found almost in all first transition series and part in early second and third transition series. The most numbers are observed in *Mo(II)*, *W(II)* and *Fe(III)* complexes. The metals in second and third transition series and column 9, 10, 11, i.e. Rh, Ir, Pd, Pt, Ag and Au have not been found in this coordination number from the data search. It can be easily explained from the nature of these metals (Greenwood and Earnshaw, 1994).

The dataset for seven-coordination in this study includes the complexes with seven metal-ligand bonds. Here the bound ligand atoms are located by the definition of

Table 4.1 Search results for seven-coordination in the CSD

M	N_e	M	N_e	M	N_e
Sc	7	Y	13	La	3
Ti	27	Zr	26	Hf	1
V	19	Nb	25	Ta	30
Cr	5	Mo	95	W	64
Mn	27	Tc	9	Re	8
Fe	33	Ru	2	Os	3
Co	26	Rh	0	Ir	0
Ni	4	Pd	0	Pt	0
Cu	15	Ag	0	Au	0
Zn	22	Cd	33	Hg	8

the CSD valence distance, no extra distance or angle restrictions are given so that all possible heptacoordination complexes rather than the author's subjective preference will be included in the CSD data. Normally, a single metal as the coordination centre is considered, the dimer or multi-metal centres are not included in the dataset except those metal atoms which can be well located at vertex position of any geometrical polyhedron. Also the complexes with lone pairs which occupy coordination sites are excluded.

4.2.2 Basic Geometrical Identification and Classification of Seven-Coordination Complexes by $R_{ang}(x)$ Values

The simple procedure which has been detailed in Chapter 3 was used to perform the comparisons of experimental structures with the idealized geometrical forms, using L-M-L valence angles to achieve the classification and identification of geometries of seven-coordination complexes. The criterion is the $R_{ang}(x)$ value. The geometrical description of a complex is identified by this value and structures with small differences from the standard model are classified as the same geometry as the standard one, that is, PBP, COC and CTP.

All the experimental structures are from the CSD search. The files containing bond angles, bond lengths and atomic coordinates from the search are input selectively to the program. The idealized model angles stored in the files are convenient to compile and further develop, as another input. The arrangements of ligands around the metal for each set of structural data input are automatically considered 5040 times in the program so that all the permutated isomers are included. The R_{ang} value is calculated for each possibility, the smallest one is selected and output with the corresponding arrangement of ligands. Based on this permutation scheme, when a pair of axial ligands is located at the same sites as those in PBP model, the angle formed by two axial positions through the central atom is 180°. Therefore exchange of these positions will not affect the $R_{ang}(PBP)$ value. The equivalent isomers with the same smallest R_{ang} value are those related by symmetry elements E, $2C_5$, $2C_5^2$, $5\sigma_v$ and σ_h of the D_{5h} point group in PBP geometry. Operation σ_h will exchange only the axial positions into each other. Another ten symmetry elements will derive ten equivalent isomers by re-arranging the equatorial ligands. In this way, there are in total twenty such equivalent isomers in the case of PBP. For example, HABLEB (Gao, Guery, and Jacoboni, 1993) is a complex $[ZrF_4(H_2O)_2OS(CH_3)_2]\cdot 2H_2O$ with seven unidentate ligands, four fluorine atoms, two water molecules and a dimethylsulfoxide molecule, shown in Figure 4.2.

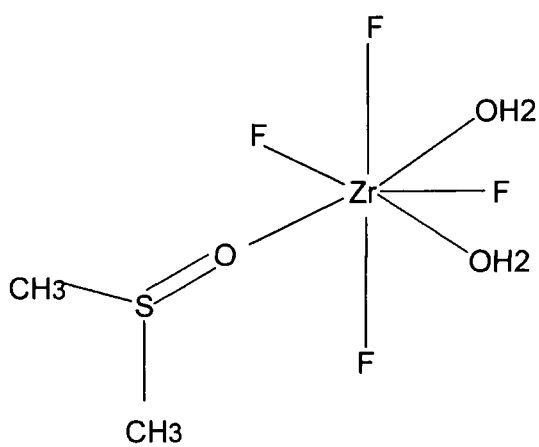


Figure 4.2 Connectivity diagram of coordination sphere in hit HABLEB from the CSD

Table 4.2 Ten equivalent isomers with the smallest R_{ang} value for complex $[ZrF_4(H_2O)_2OS(CH_3)_2]\cdot 2H_2O$ in PBP form

No.	Ligand Sites							Symmetry Operation
	1	2	3	4	5	6	7	
1	F_1	F_4	F_2	O_2	O_3	F_3	O_1	E
2	F_1	F_4	F_2	O_3	O_2	O_1	F_3	σ_v^{I}
3	F_1	F_4	F_3	O_1	O_3	F_2	O_2	σ_v^{ii}
4	F_1	F_4	F_3	O_3	O_1	O_2	F_2	$C_5^{2'}$
5	F_1	F_4	O_1	F_3	O_2	F_2	O_3	$C_5^{2''}$
6	F_1	F_4	O_1	O_2	F_3	O_3	F_2	σ_v^{iii}
7	F_1	F_4	O_2	F_2	O_1	F_3	O_3	σ_v^{iv}
8	F_1	F_4	O_2	O_1	F_2	O_3	F_3	C_5'
9	F_1	F_4	O_3	F_2	F_3	O_1	O_2	C_5''
10	F_1	F_4	O_3	F_3	F_2	O_2	O_1	σ_v^{v}

The calculations with the PBP standard gave the smallest $R_{ang}(PBP)$ value of *2.44%*, in which F_1 and F_4 are located at axial positions. Meanwhile, twenty equivalent isomers with the same R_{ang} value are also derived within the permutation. Table 4.2 lists ten of them, the other ten can be generated by the symmetry operation σ_h, that is, change F_4 to F_1 and vice versa. The first one is assigned to symmetry *E* as a reference, all others are generated by symmetry operations from this reference. In the case of the second model for CTP geometry, the symmetry elements in point group C_{2v} are related. The smallest $R_{ang}(CTP)$ value will appear with those isomers in which ligands *1, 2, 3* are located at the same positions as the standard one. In the idealized model, apical sites *1, 2, 3* and equatorial sites form two planes orthogonal to each other. Normally the equivalent isomers with the smallest $R_{ang}(CTP)$ value are related to symmetry elements *E*, C_2, $\sigma_v(yz)$ and $\sigma_v(xz)$, which have four isomers. But with increasing deviation from this model, the number of such isomers will decrease.

The R_{ang} calculation provides the numerical identification of the geometry of an observed complex. From the R_{ang} value, the basic geometrical shape to which a complex belongs can be determined easily by comparing the derived values from the given archetypes. It is obvious that the smaller value of $R_{ang}(x)$ represents geometry closer to a standard form. From the overall results of this calculation, it can be said that

when the $R_{ang}(x)$ value is less than 10.0%, the observed coordination sphere illustrates a good match with a standard geometry. Therefore, the geometry can be identified as PBP or CTP. However, some complexes distort from these idealized geometries and give large $R_{ang}(x)$ value for both geometrical forms (x). In this case, a clear polyhedral form can not be assigned directly to the coordination sphere, but it may represent an intermediate state of interconversion between two geometrical forms. Detailed investigations on these complexes may map the interconversion pathway or reveal the influence of coordination environments to form the coordination geometry for a particular complex, which will be discussed in the later sections. Nevertheless, the closest geometry can still be found from the $R_{ang}(x)$ values and any particular distortions might be of great interest and importance in our systematic analyses.

Referring the observed molecular geometry to one of idealized geometrical archetypes completes the description of coordination sphere geometry.

The geometry PBP has characteristic angles and the difference in the calculation of the $R_{ang}(x)$ values from the CTP-like model is significant. But as was indicated in Chapter 3, there exists another geometry COC that is less different in the angles from the CTP form. This polyhedron has been considered as the intermediate in the interconversion from the PBP to CTP as shown in Figure 3.9. Although idealized angle values are not defined in the program, the COC can be still observed by viewing this conversion pathway. The scheme is described as follows:

(1) Find the difference of each angle between two observed models, that is,

$$\Delta_i = \theta_i(\mathrm{PBP}) - \theta_i(\mathrm{CTP})$$

(2) Start from one model, PBP and change the angles by a small percentage of Δ_i, such as, 1% in each step, increase the change until the other model is achieved.

(3) Calculate $R_{ang}(i)$ value with respect to each step of angle change.

(4) Plot R_{ang} value against the angle change percent and find the lowest point that represents the smallest $R_{ang}(i)$ value in the plot.

According to the geometries of the two models, Δ_i have different values corresponding to the different angles. For a large Δ_i, the amount changed in each step is also large. The direction of change, that is, addition to or subtraction from the starting

angle values, is dependent upon the sign of Δ_i value. Thus, the conversion of one model (PBP) to the other one (CTP) is obtained through a simulation of the motion of the ligands, by changing the related angles. The R_{ang} value is still a criterion along this interconversion pathway. It is obvious that the smallest R_{ang} value reflects the conformation of a complex compared with the proposed models. At this point, the description of the geometry can be expressed by the percentage of the angle movements from the PBP to CTP.

In the program, the angles are designed to change starting from an idealized PBP model. For a complex with close to PBP geometry, all angles are close to those in a starting model, and then the smallest R_{ang} value should appear at the beginning of the changing steps. As the angles move further towards the CTP model, the deviations between the observed angles and model ones are raised and the R_{ang} value increases steadily. In contrast, for a CTP-like complex, the R_{ang} value decreases as the angles move closer to the final model. Therefore, the smallest R_{ang} value can be obtained in the end of this change process. In the case of the COC geometry, at the beginning of the angle change, it will show a larger R_{ang} value because it has a larger deviation from the PBP. As the model continues to change, the R_{ang} value decreases and the smallest R_{ang} will appear at some point along the transformation from the PBP model to CTP. As the model moves further to the CTP after the smallest R_{ang} value, the angles deviate from the model angles and the R_{ang} value rises again. Since the angles in the COC are closer to the CTP than PBP, the changing curve will not be symmetrical with respect to the smallest R_{ang} value. The smallest value representing the COC geometry will appear closer to the ideal CTP geometry.

Figure 4.3 shows the typical curves that represent the three cases described above. The distribution is plotted as the R_{ang} value versus the angle change that reflects the percentage change away from the PBP towards the CTP model. The $R_{ang}(x)$ calculations gave 2.44% in PBP and 12.6% in CTP for complex HABLEB, respectively. It is clearly classed as PBP geometry and it illustrates a line with positive slope. While results of calculation on complex BAVXUR provided $R_{ang}(x)$ values of *14.22%* and *0.9%* for PBP and CTP. It has a nearly perfect CTP polyhedron, and shows a line with negative slope in the plot. Complex CMPMOC has a geometrical form between these two models, with two R_{ang} values, *16.5%* (PBP) and *7.6%* (CTP), respectively. The reference $R_{ang}(COC)$ value for this complex is 1.8%. It gives the smallest R_{ang} value at ~ 72% of angle change, which represents an intermediate between PBP and CTP. Here

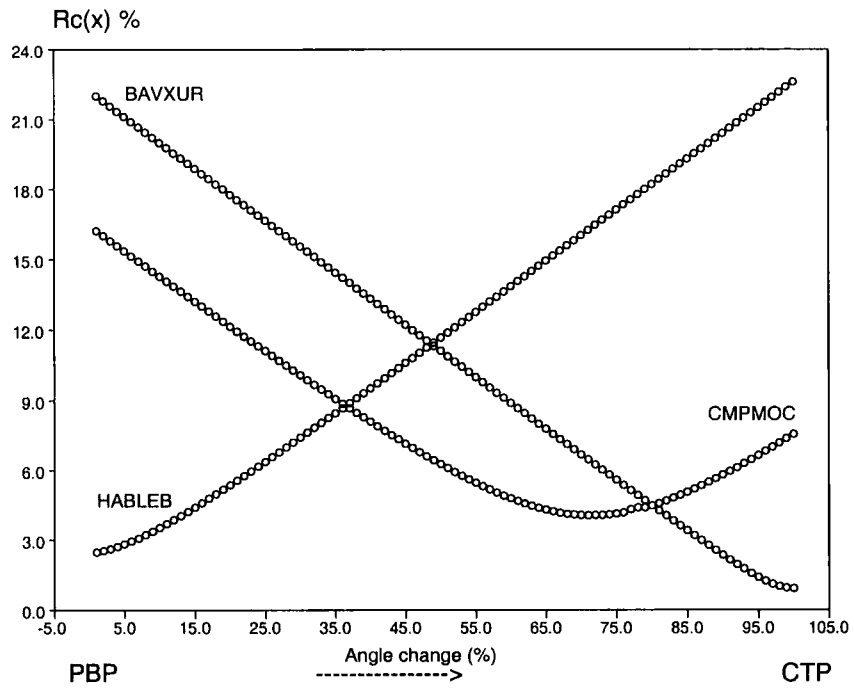


Figure 4.3 R_{ang} value(%) *vs.* angle change(%) from ideal PBP to CTP geometry.

it could be defined as COC geometry according to the proposed interconversion pathway.

The procedure for identifying and classifying the geometry of a coordination sphere in each individual observed structure, that most closely approximates the idealized one, may be consequently summarised as two points: (i) calculate $R_{ang}(x)$ values from all possible $n!$ permutations of the ligands as it appears originally in the angular data matrix and find that permutation which corresponds to the minimum deviation from one of the idealized geometrical forms; (ii) calculate $R_{ang}(i)$ value at each step needed to transfer one geometrical form to the other and the smallest value with angle change shows how the individual observed geometry will distort from the idealized one by the expression of a percentage change between two models.

The procedure was applied to all seven-coordination complexes retrieved from the CSD, it resulted in basic identification of geometry for these known structures by the $R_{ang}(x)$ values. For the structures without any chelate ligands, an $R_{ang}(x) \leq 10\%$, which corresponds to an idealized geometrical form, can be obtained and geometry can be classified directly.

ML_7 with all unidentate ligands

Geometries of seven-coordination compounds in this category show the minimum influence from the ligand conformations. The main factors to make the geometry distort from the idealized one, are weak interactions between ions or molecular packing effects. For example, $[NbF_7]^{2-}$ anion (Brown & Walker, 1966) is distorted from CTP because of $K^+\cdots F$ interactions. Usually, there is no distinct deviation in the M-L bond angles from the standard angles.

The CSD search gave 62 such compounds* and 67 fragments, which contain transition metals *W, Mo, Zr, Ta, Re, Nb, Ti, Y, Cr, La* and *Tc*. In these complexes, only 10 compounds have seven identical unidentate ligands. The remaining complexes all have different unidentate ligands. For many molecules of this type, L···L repulsions are considered as a major factor in the choice of geometry (Drew & Wilkins, 1974). Table 4.3 gives the results of $R_{ang}(x)$ calculations on these complexes together with the smallest R_{ang} (denoted as SR_{ang}) which quantifies in the interconversion between PBP and CTP models (*AC%*).

The results show that PBP geometry is not a favoured form for 7-coordinate unidentate complexes. Only six compounds (FINBAF, FINBEJ10, HABLEB, PEHJAN01, PIVVAR, TIMNIM) adopt this geometry, all have a d^0 configuration. Packing effects due to hydrogen bonds in these complexes seem to dominate the choice of geometry. They all have at least two halogens coordinated directly to the metal, of which two of them enter axial sites of the PBP. This is consistent with the theoretical calculation from the orbitals (Hoffmann, Beier, Muetterties and Rossi, 1977), in which the analysis of ML_7 charge distribution in PBP showed apical ligands distinctly more negative, that is, better σ acceptors should preferentially sit at the axial positions. These

* The CSD 1995 version searched 47 ML_7 unidentate compounds and 48 fragments. As this section was being written , the CSD 1997 version came available, the number given here is the updated data from 1997 version.

Table 4.3 Results of $R_{ang}(x)$ calculation and interconversion model from PBP to CTP for 60 unidentate complexes

Refcod	Compound	R_{ang} (%)			SR_{ang} (%)	AC (%)
		PBP	COC	CTP		
BAVXUR	$W(CN^tBu)_7 \cdot 2PF_6$	14.2	7.32	0.94	0.94	100.0
BAYFUC	$V(CO)_6(AuPPh_3)$	16.7	6.79	9.68	6.87	71.0
BEGSOV	$Cr(CN^tBu)_7 \cdot 2PF_6$	11.8	5.09	6.31	4.45	80.0
BEZFUH	$Mo(CNPh)_7 \cdot 2PF_6$	13.3	4.17	5.57	3.70	83.0
BOGBII	$W(CO)_4(PMe_3)I_2$	14.5	1.84	6.53	3.33	75.0
BTOIRE	$Re(CNPh)_4Br_3$	15.1	1.75	6.66	3.30	66.0
BUTXAP	$W(CO)(CN^tBu)_4I_2 \cdot CHCl_3$	14.8	2.65	6.71	3.18	74.0
CIPWPB	$W(CO)_3(PMe_2Ph)_3I \cdot BPh_4$	14.3	4.99	2.99	2.36	92.0
CIWYAI	$Cr(CO)_4(SnPh_3)_3 \cdot NEt_4$	14.8	4.02	5.72	4.08	64.0
CMPMOC	$Mo(PMeCl_2)_3Cl_4 \cdot CS_2$	16.4	1.79	7.57	4.07	71.0
CMPPTC	$Tc(CO)(PMe_2Ph)_3Cl_3 \cdot EtOH$	15.6	1.05	6.84	2.88	72.0
CUGSUS	$W(PMe_3)_3Cl_4$	14.8	2.79	4.78	1.96	80.0
CUKRIJ	$Ta(PMe_3)_3Cl_4$	14.2	3.96	3.58	2.11	87.0
CUSMIM[1]	$Ta(CO)_3(PMe_3)_3Cl$	11.9	8.10	4.61	4.61	100.0
CUSMIM[2#]	$Ta(CO)_3(PMe_3)_3Cl$	11.5	8.36	5.35	5.34	98.0
CUYZEB	$Ta(CO)_3(PMe_3)_4 \cdot Ta(CO)_5PMe_3$	14.3	4.70	5.74	6.26	73.0
DADVIN	$ZrF_7 \cdot 3[C(NH_2)_3]$	12.0	8.12	3.84	3.82	99.0
DAMTIU	$Mo(CO)_2(OOCMe)_2(PMe_3)_3$	11.8	7.35	5.26	4.96	93.0
DAXHOZ	$Re(CO)_2(PMe_2Ph)_2Br_3$	15.0	2.62	6.58	3.53	66.0
DEVYOS	$Nb(PMe_2Ph)_3Br_4$	12.5	5.59	3.76	3.22	92.0
DUKCUH	$W(CO)_4Br_3 \cdot [HC(N^iPr_2)OEt \cdot CH_2Cl_2$	15.7	0.75	6.66	3.11	74.0
ENFZRB10	$2ZrF_7 \cdot 3NH_3(CH_2)_2NH_3 \cdot 2H_2O$	12.0	7.89	3.75	3.73	98.0
ENFZRB[11#]	$2ZrF_7 \cdot 3NH_3(CH_2)_2NH_3 \cdot 2H_2O$	13.7	6.93	3.78	3.54	94.0
FINBAF	$Zr(NCMe)_4I_3 \cdot Zr(NCMe)_5I$	4.16	15.1	13.2	4.15	1.0
FINBEJ10	$Zr(NCMe)_4Br_3 \cdot Zr(NCMe)Br_5 \cdot MeCN$	4.70	14.3	12.2	4.73	1.0
GIFGAD	$W(CO)_3(NCMe)_2I_2$	15.4	4.50	7.63	3.15	67.0
GIFGAD10	$W(CO)_3(NCMe)_2I_2$	15.4	4.49	7.63	3.16	67.0
GIJBUW	$Nb(PMe_3)_3Cl_4$	14.5	3.96	3.66	2.32	88.0
HABLEB	$Zr(H_2O)_2(OSMe_2)F_4 \cdot 2H_2O$	2.44	14.7	12.6	2.49	1.0
HIBNOV	$Re(PMe_2Ph)_2(CNCy)_2 \cdot C_6H_4Cl_2$	14.3	2.64	6.36	3.15	63.0
IBICMO	$Mo(CN^tBu)_7 \cdot 2PF_6$	14.2	7.34	0.68	0.68	100.0
ICBICW	$W(CO)_2(CN^tBu)_3I_2$	12.5	4.01	6.12	4.07	79.0

Table 4.3 *contd.*

ICBICW02	$W(CO)_2(CN^tBu)_3I_2$	14.6	2.71	5.10	2.91	81.0
ICBICW03	$W(CO)_2(CN^tBu)_3I_2$	14.7	4.41	6.27	4.56	73.0
JANGIO	$Mo(NO)(PMe_3)_3Cl_3$	15.4	2.96	5.19	3.06	81.0
JEGSAP	$Mo(CO)(CN^tBu)(PMe_3)_3Cl_2$	12.6	4.99	5.75	5.46	52.0
KUDCAN	$Mo(CO)_2[P(OMe_3)]_3(SnBuCl_2)Cl$	11.1	7.54	5.42	5.19	93.0
LEPDIT	$W(CO)_4I_3 \cdot W(CO)_3I \cdot C_{20}H_{40}S_4$	14.3	2.06	6.41	3.34	76.0
LEPDOZ	$W(CO)_4I_3 \cdot W(CO)_2(C_{20}H_{40}S_4)I$	16.0	1.31	7.10	3.56	70.0
MPPCMO10	$Mo(PMe_2Ph)_3Cl_4 \cdot EtOH$	15.4	1.62	6.75	2.79	72.0
PEHJAN01	$Y(THF)_5Cl_2 \cdot Y(THF)_2Cl_4$	2.45	14.6	12.5	2.09	6.0
PIBWUS	$Ti(CO)_6[Sn(C_6H_{11})_3] \cdot K(crystand2.2.2)$	12.9	6.96	4.31	4.05	94.0
PIVVAR	$ZrF_7 \cdot (H_3NCH_2CH_2)_2NH_3 \cdot H_2O$	3.26	14.8	12.8	2.00	2.0
PSNCOV	$V(CO)_5(SnPh_3)_2 \cdot NEt_4$	13.2	5.33	4.20	3.29	89.0
SABCUT	$W(CO)_3(SbPh_3)I_3 \cdot PPh_3 \cdot OEt_2$	14.5	3.44	6.66	3.59	75.0
SARNII	$W(CO)_4I_3 \cdot P^tBu_3I$	16.2	1.04	7.36	3.73	72.0
SODZUG	$Nb(CO)_3(Pme_3)_4 \cdot Nb(CO)_5PMe_3$	14.7	4.75	6.03	6.61	61.0
TCDCPM	$Mo(CO)_3(Pet_3)_2Cl_2$	15.3	2.19	6.11	3.09	76.0
TIMNIM	$Y(H_2O)_2(NCMe)_2Cl_3$	4.12	14.6	12.4	3.74	8.0
TUNDIP	$W(CO)_3(PPh_3)(NCMe)I_2$	13.5	4.44	5.27	3.03	81.0
VESSER	$W(CO)_2(PMe_3)_3Br_2$	12.2	5.46	6.28	5.59	50.0
YEVDUY	$W(CO)_3(NCMe)(SbPh_3)Br_2 \cdot CH_2Cl_2$	13.8	3.30	5.89	3.88	71.0
YEVFAG	$W(CO)_3(NCMe)(SbPh_3)Br_2 \cdot CH_2Cl_2$	14.1	2.80	5.89	3.46	72.0
YIDBES	$2WF_7 \cdot W(C_{10}H_8N_2)_2F_4 \cdot NCMe$	12.1	9.25	7.46	7.46	100.0
YIDBES[1#]	$2WF_7 \cdot W(C_{10}H_8N_2)_2F_4 \cdot NCMe$	11.8	8.88	5.49	5.49	100.0
YOCSIS	$W(CO)_4Cl_3 \cdot 18\text{-crown-}6 \cdot H_2O$	15.2	2.06	6.41	3.20	67.0
YOCSOY	$W(CO)_4Cl_3 \cdot C_{12}H_{30}N_2O_5$	15.3	1.86	6.47	3.35	75.0
YOCSOY[1]	$W(CO)_4Cl_3 \cdot C_{12}H_{30}N_2O_5$	15.1	1.73	6.40	3.49	66.0
YUNDUG	$W(CO)_3I_2(\eta^2\text{-dppf})$	15.4	3.16	6.53	3.87	67.0
YUNMUP	$W(CO)_2(PEt_3)_2(CNCy)Br_2$	14.1	3.92	4.28	2.34	83.0
YUNMUP[1#]	$W(CO)_2(PEt_3)_2(CNCy)Br_2$	14.6	3.51	4.98	2.50	81.0
YUVTOY	$W(CO)_3(NCEt)_2I_2$	15.1	3.95	6.84	3.12	67.0
ZABNEV	$W(CO)_3(NCEt)(AsPh_3)I_2$	13.6	4.14	5.91	3.07	60.0

Table 4.3 *contd.*

ZEQQAN	$La(OCHMe_2)_4I_3$	14.4	6.85	8.10	6.03	75.0
ZOBGIG	$MoF_7 \cdot NCMe \cdot NO_2$	12.8	7.11	7.06	5.46	81.0
ZOBGUS	$ZrF_7 \cdot C_{11}H_{24}N$	11.3	6.46	5.79	4.80	86.0
ZOXCUK	$Zr(CO)_5(SnMe_3)_2 \cdot 2(C_{20}H_{40}O_{10}K)$	14.1	6.68	3.65	3.49	96.0

Refcode with superscript number means that there is more than one fragment in the crystal structure.

sites are also the least crowded with respect to other ligands. In CTP form, the capping position (labeled 2 in Figure 3.2) has also the same effect. Thus, if L···L repulsion is a major factor determining the geometry of the coordination sphere, when molecules contain one halide atom (*X*) or a bulky ligand, the *X* atom or large size ligand usually occupies this position to form CTP geometry. These examples can be seen in KUDCAN, PIBWUS, CIPWPB and CUSMIM etc. There are a total of 19 fragments having CTP or close to a CTP polyhedron.

The results obtained from the theoretical treatment of the seven coordination sphere based on the hard sphere model indicated that the COC polyhedron is the energetically preferred geometrical form (Gillespie & Hargittai, 1991) but that the other two forms are only slightly higher in energy. So only small movements of the ligands through low energy barriers are required to convert one geometry into another, (Gillespie, 1992). The $R_{ang}(COC)$ values are also calculated from the reference angles as listed in Table 2.6 for comparison with other two geometries in this procedure. The most common geometrical form should be COC in this data set. 42 fragments have the minimum deviation from this polyhedral form, which has almost 65% population of the whole data set. It can be seen that the experimental complex structures are consistent with the theoretical prediction, when the chelate effect of ligands is considered to be minimum.

Moreover, how these complexes are close to other two geometrical modes as intermediates between interconversion from PBP to CTP can be seen from the values listed in the last two columns. It is clear that these intermediates appear after 50% change from PBP to CTP model and most of them locate at 60% to ~90%. When a R_{ang}

value shows a ≥ 90% change from PBP to CTP, it is considered that the coordination sphere has close to or perfect CTP geometry. Therefore, SR_{ang} and AC values provide an unambiguous confirmation of the assigned coordination polyhedron. Indeed, a random or ambiguous geometry assignment between COC and CTP is postulated because these two forms are quite similar geometrically. For example, $Ti(CO)_6[Sn(C_6H_{10})_3]^+$ (PIBWUS, Ellis & Yuen, 1993) was initially assigned as a COC with the tin atom at the capping position of the octahedron. However, the present calculation shows a good match to the CTP form and the tin atom locates in the least crowned capping site of the rectangular face. Two other coordination spheres, $W(CO)_3(CN^tBu)_4I_2$ (CIPWPB, Drew & Wilkins, 1974) and $Mo(CO)_2[P(OMe)_3]_3(SnBuCl_2)Cl$ (KUDCAN, Miguel, Perez-Martinez, Riera & Garcia-Granda, 1991) were not clearly defined as COC or CTP geometries. In this study, their geometries can all be clearly identified as CTP, which give SR_{ang} values *2.99%* and *5.42%* with AC of *92%* and *93%*, respectively.

Nevertheless, there are some structures that do have a geometry close to two polyhedral forms, especially between COC and CTP. The calculations of $R_{ang}(COC)$ and $R_{ang}(CTP)$ gave only small differences between these two criteria. This can be found in compound $Ta(PMe_3)_3Cl_4$ (CUKRIJ, Cotton, Duraj and Roth, 1984), which has 3.96% [$R_{ang}(COC)$] and 3.58% [$R_{ang}(CTP)$], respectively. It is difficult to say which geometry CUKRIJ is closer to from these two discrepancy values only, but the results of SR_{ang} and AC calculation provide a clear identification. SR_{ang} gives a smallest value of 2.11% at 87% of AC. This means that in the interconversion pathway from PBP to CTP through COC, the geometry of the coordination sphere in this compound is indeed close to COC and CTP but more inclined to the CTP form. Similar examples can be seen in GIJBUW (Cotton, Diebold and Roth, 1985) and TUNDIP (Baker *et al*, 1996). In the latter one, it has less inclination to CTP, which has the smallest $R_{ang}(i)$ value at 81% of AC.

The above results indicate that the $R_{ang}(x)$ value is a key factor in determining which polyhedral model can best be used to describe coordination sphere geometry. Then, it is important is to know what range the $R_{ang}(x)$ value falls into, only then can a complex be described by the geometry from which the relevant $R_{ang}(x)$ value is derived. Through investigation of all the data and combining the results from the $R_{ang}(x)$ and SR_{ang} and AC values, it is concluded that when SR_{ang} appears within the first or last 10% of AC, the geometry can be classified clearly as a PBP or CTP polyhedron. While for the SR_{ang} between 60% and 80% of AC, it could be COC geometry. For the SR_{ang} in

other ranges of AC, the geometry is usually regarded as close to two polyhedra. The *AC* value can show how close the geometry of a coordination sphere is to these polyhedra.

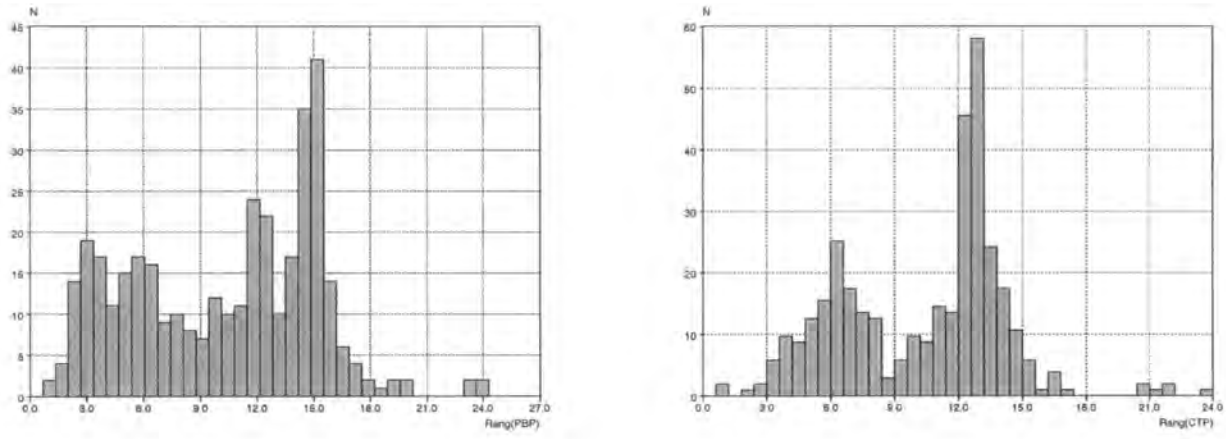
Once a geometrical form has been assigned to a coordination sphere from the relevant $R_{ang}(x)$ value, a question may be raised as to which isomer (i.e. the arrangements of ligands) is the best expression in the description of geometry, with each contributor minimized over the 5040 possible ligand atom permutations? Is there another (or more than one) appropriate mapping of ligand atoms? A test from a calculation output which included the smallest and next smallest $R_{ang}(x)$ values for each complex in both PBP and CTP models showed that the smallest $R_{ang}(x)$ value for the corresponding isomers in the PBP form was identified from others. Similarly, there is a distinguishable $R_{ang}(x)$ value for the smallest and next smallest values in the case of CTP. This implies that the smallest $R_{ang}(x)$ in the procedure provides as a reliable criterion value.

Ligands with Denticity >1, Preliminary Observations

The use of this simple but effective procedure has readily identified the geometry of coordination sphere for all unidentate ligand complexes. Since no chelating effects are included in these species, the R_{ang} values for all possible models are generally distinct from each other and therefore a geometrical form can easily be assigned to each coordination sphere. However, for the larger set of complexes with multidentate ligands, the geometry identification from the calculations will not be as simple. The effect of chelate ligands will become an important factor in the determination of geometry and distortions from idealized geometrical archetypes will often be observed. In these cases, R_{ang} values for any geometry might be large and difficulty in geometry assignment will occur. Therefore, it is necessary to investigate further and classify the ligand type to see how much geometry is affected by different coordination environments in the systematic analysis.

This procedure has been applied initially to all seven-coordination complexes with various types of ligands from the CSD. It is similar to those for complexes with all unidentate ligands, the geometry of most complexes can be identified and classified by the distinct R_{ang} values. Unlike the all unidentate ligands case, the total numbers of complexes for three different geometries are more average. Of the 372 structural fragments in the data set, 126 (with $R_{ang}(PBP) < 10\%$) belong to the PBP geometry.

This proportion (~ 34%) for PBP in PBP, COC and CTP geometries is obviously increased compared with that (< 9%) in all unidentate ligands complexes. This implies that chelate ligands can change the geometrical preferences in coordination sphere, and details about this will be discussed later. For an overview, distributions of R_{ang} values in the PBP and CTP are shown in Figure 4.4 (a) and (b) as histograms, respectively.



(a) (b)

Figure 4.4 R_{ang}-histograms for PBP (a) and CTP (b)

Essentially, these bimodal distributions, with one peak close to the origin and the other at an $R_{ang}(x)$ of 15-20%, indicate the two major components whose relative frequencies can be obtained very simply by inspection, with the two peaks being connected by a series of data points with intermediate R_{ang}-values, indicative of a geometrical interconversion pathway. Details of these will be explored by more complex multivariate analyses.

4.2.3 Comparison with the Potential-Energy surface results

A theoretical calculation based on VSEPR theory (Valence Shell Electron Pair Repulsion) was used to treat the stereochemistry of coordination compounds having several coordination numbers (Kepert, 1987; Gillespie and Hargittai, 1991). In this model, the determining factor of geometrical choice of coordination sphere is that of repelling points on a 'sphere', formed by ligands about the metal atom. Each metal-ligand bond is considered to act as a point. The total repulsion energy U can be expressed as a summation of the repulsion energies of every pair of metal-ligand bonds u_{ij}:

$$U = \sum_{ij} u_{ij}$$

and u_{ij} is inversely proportional to some power m of distance d_{ij} between them:

$$u_{ij} = \frac{a}{d_{ij}^{m}},$$

where a is the proportionality constant. All the bonds are assumed equal (r) and using X to represent the numerical repulsion energy coefficient, thus,

$$U = \sum_{ij} a d_{ij}^{-m} = aXr^{-m}.$$

X is a function of m and the geometry of the coordination polyhedron and it can be seen as the sum of effects of individual ligand repulsion energy coefficients Y_i for each ligand:

$$X = \sum_{i} Y_i .$$

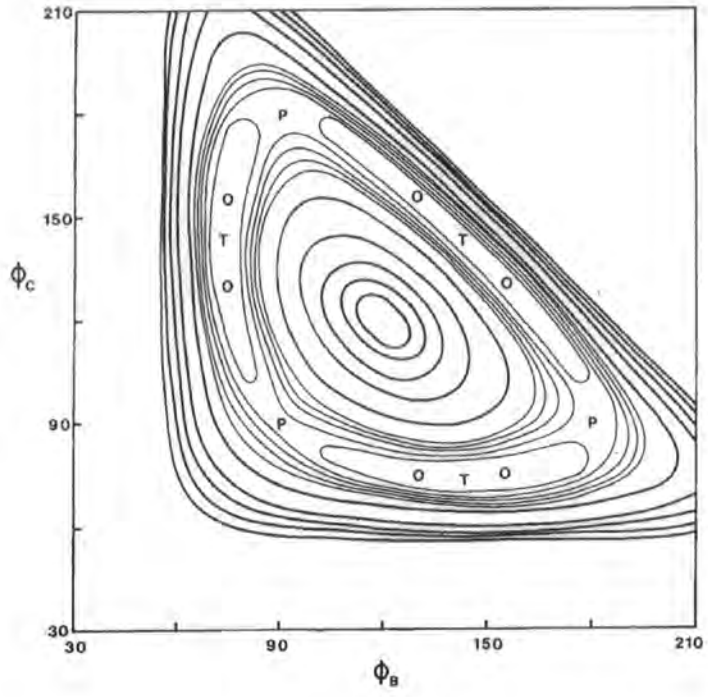
In this way, the minimum values for the repulsion energy coefficient X for the three geometrical polyhedra in seven coordination were given as *3.266 (PBP)*, *3.230 (COC)* and *3.231 (CTP)* (Kepert, 1979). They are very similar. The calculation of the potential energy surface for some specified angles between ligands and the metal atom

resulted in a map projected onto the appropriate angle ϕ_B-ϕ_C* plane [shown here as Figure 4.5(a) for comparison], in which the locations of three geometries *PBP*, *COC* and *CTP* could be clearly indicated as *P*, *O* and *T*, respectively. The "reaction coordinates" were observed from the energy contour lines. It can be seen that the COC form is an intermediate between the PBP and CTP. It is obvious that the COC form in this energy plot is much closer in energy and geometry to the latter. This current calculation on R_{ang} values also represents the same point as stated in last section.

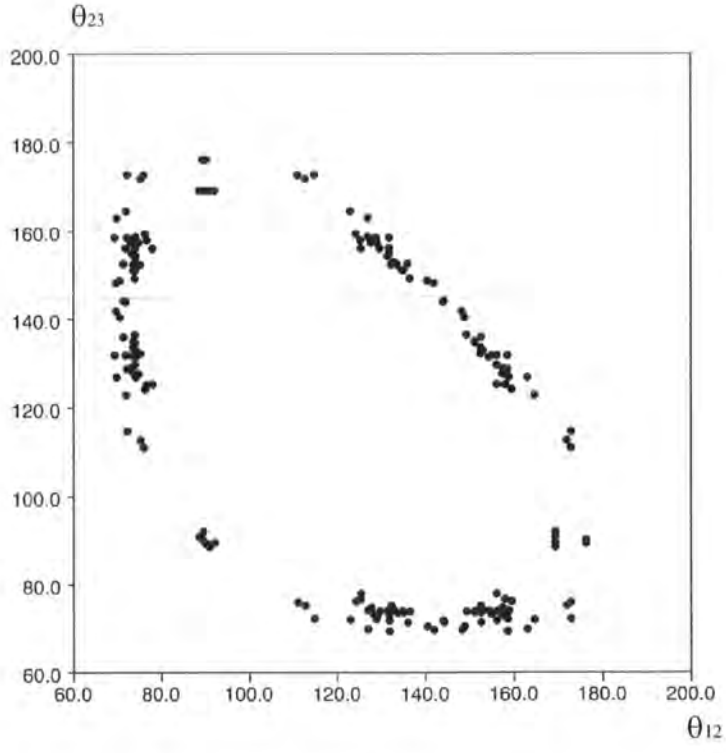
Considering the symmetry, a mirror plane is retained with this transformation from PBP to CTP. All the angles are relative to this plane and this plane can be, in fact, viewed equally from Figure 3.2, which is made up by positions *1*, *3* and *2* (equivalent to A, B and C in Kepert's model). Therefore, the corresponding definitions for ϕ_B and ϕ_C in this labeling scheme are θ_{23} and θ_{12}, respectively.

The results for the geometry of observed coordination compounds based on the calculation of R_{ang} values, present a perfect matching to stereochemical prediction by the potential energy surface. All the observed coordination spheres that have been identified as different geometrical forms appear along the energy minimum of a "moat". For the comparison with the potential energy surface plots, the ligand atoms at position *1*, *2* and *3* are also chosen as the starting atom in turn, so that each pair of such angles appears more than once on the corresponding θ_{23}-θ_{12} plots. Thus, very similar plots to that derived from the potential energy surface method are obtained. Figure 4.5 shows the plots of θ_{23} against θ_{12} for (b) compounds with all unidentate ligands, (c) *Fe* complexes, with all kinds of ligands and (d) all seven coordinate complexes in the data set. As in the potential energy surface plot, PBP [P in (a)] occurs around three sites on the plots at: θ_{12} = 90°, θ_{23} = 90°; θ_{12} = 180°, θ_{23} = 90°; and θ_{12} = 90°, θ_{23} = 180°. CTP [C in (a)] lies around θ_{12} = 144.2°, θ_{23} = 71.5°; θ_{12} = 144.2°, θ_{23} = 114.2°; and θ_{12} = 71.5°, θ_{23} = 144.2°. COC [O in (a)] is considered as falling between these areas. The locations of PBP, CTP and COC could be indicated clearly and the interconversion between these geometries is observed from the energy contour lines. COC has a higher population in all unidentate ligand compounds as given in the last section and also shows here in (b). This implies that COC seems be energetically the most favorable form when the steric effects of the ligands are kept to the minimum. Whereas *Fe* (a mixture of unidentate and multidentate) complexes have mostly either in PBP or CTP

* The ϕ angular coordinates are defined as the angles between the bonds M-L_{ax} and other M-L bonds

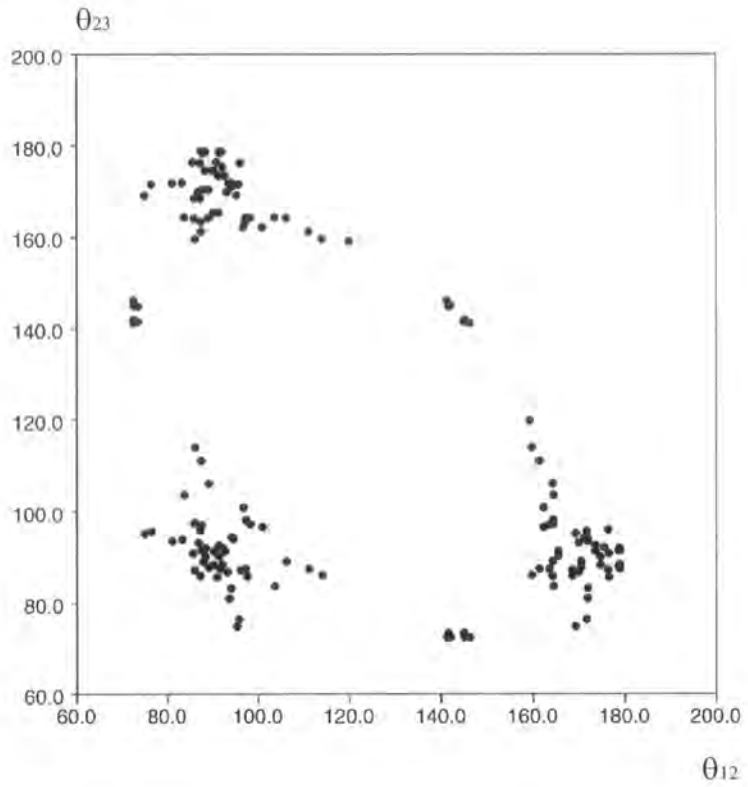


(a) Projection of the potential Energy surface for ML_7 onto the ϕ_B-ϕ_C plane (°). This is reproduced from Kepert, 1979 and it is given here to compare with this work.

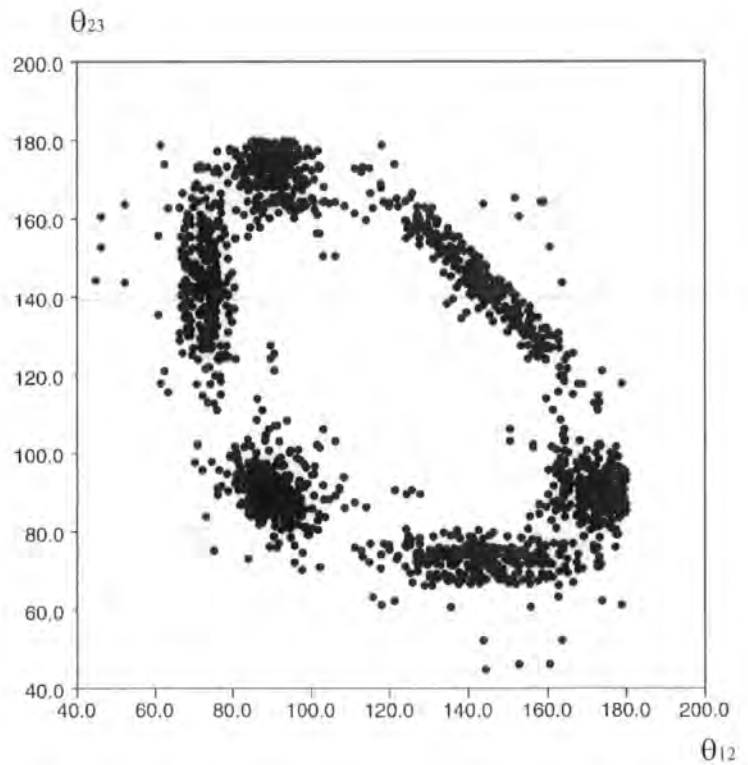


(b) θ_{23} *vs.* θ_{12} for all unidentate ligand molecules

Figure 4.5 comparison of plot of potential energy surface(a) with the plots in complex structures (b) (c) (d).



(c) θ_{23} vs. θ_{12} for *Fe* complexes



(d) θ_{23} vs. θ_{12} for all seven coordination fragments

Figure 4.5 *contd.*

geometry. From a total of 28 molecules, 25 belong to PBP and 3 to CTP. No clear COC form is observed in these complexes. Plot (c) shows that all the points locate around these two clusters. Finally, plot (d) illustrates a more complete geometry represented for all three polyhedral forms and intermediates between these forms. In addition, some extreme examples, probably with a larger chelate effect, are outside the minimum energy zone.

The potential energy surface obtained from this point-charge model provided a prediction of three geometries applied in the $ML_{7\ or\ n}$ coordination sphere. Although a simple formula could be used in this theoretical calculation, it does depend upon some initial assumptions. Firstly, repulsion is determined by the power number m and the m is not known. The descriptions of the three geometries are somehow dependent on the value of m. While the three polyhedra have closely similar energies, therefore the m should be carefully estimated. Secondly, an important assumption in this model is that all M-L bonds have to be equal in length so that the individual repulsion energy coefficients Y_i could be derived and thereafter the total repulsion energy coefficients X. In practice, it is difficult to find seven equal M-L bonds in a real compound. When some $M\text{-}L_i$ is distorted significantly from others, the predictions become less clear. This will need a further classification according to the ligand types. In fact, in such a classification it is difficult to include all possible cases and becomes very complicated with different coordinate sites.

A comparison of results showed that the distributions of geometry from available crystallographic data, using the discrepancy index R_{ang} value method are the same as predictions from the theoretical model. This means that the method given here also provides an accurate identification of geometry for the seven-coordination sphere and it is obviously even simpler than a point-charge model. In this calculation for all seven-coordination complexes [shown in Figure 4.5(d)], no special treatments of the ligand type are needed, and those structures with large distortions from the standard polyhedra can be located easily on the plots. Therefore, further analysis of the reasons why a structure deviates from standard geometry can be carried out based on these R_{ang} values.

4.2.4 Vertex Index $VR_{ang}(x)$

It has been seen that the deviations of the observed structure from an ideal geometry can be reflected by a relevant R_{ang} value. Influencing factors to the stereochemistry of complexes also involve the nature of ligands in addition to the metal atoms. As already shown, complexes with all unidentate ligands coordinated to the metal atom can more easily have closer geometry to the idealized models. But with bi- or multi-dentate ligands, because of the stereochemical requirements of these chelating ligands, the distortion from the ideal geometries will be more easily observed in the compounds. This will directly affect the calculation of $R_{ang}(x)$ therefore sometimes no significant index value from a symmetrical polyhedron is presented.

In order to investigate the influence of ligands on the formation of different geometrical coordination spheres, a vertex VR_{ang} factor, which is again a discrepancy index between observed and idealized values, is used to examine how each ligand atom deviates from its idealized position. This value is actually similar to the R_{ang} values used before. The difference is that each VR_{ang} represents distortions of individual ligands, instead of the whole complex as in R_{ang}. Therefore, for each vertex VR_{ang} factor, only angles related to this vertex are compared with the proposed idealized ones. Similarly, the more a ligand deviates from the idealized vertex position, the larger is VR_{ang}. Thus, a VR_{ang} can be defined as:

$$V_i R_{ang} = 100 \cdot \left\{ \frac{\sum_{j=1}^{7} \left[\theta_{ij\,(obs)} - \theta_{ij\,(std)}\right]^2}{\sum_{j=1}^{7} \theta_{ij\,(std)}^2} \right\}^{\frac{1}{2}} \qquad (i \neq j)$$

There are always six ligands adjacent to the specified ligand *(i)*, six relative observed angles $\theta_{ij(obs)}$ are used to compare with the idealized values $\theta_{ij(std)}$. Once the scheme is applied to the data set, the effects of each ligand on the geometry of coordination sphere are clearly numerated from the VR_{ang} values. The results of some iron seven-coordination complexes are presented here as examples. Their R_{ang} and VR_{ang} values are listed in Table 4.4 and the structures are illustrated in Figure 4.6 (a).

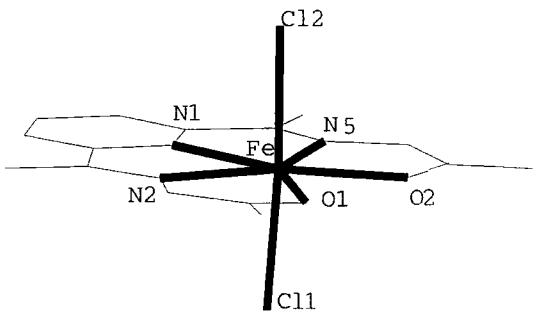
Table 4.4 VR_{ang} indexes for some selected compounds

REFCODE	R_{ang}(%)			V_iR_{ang} (%)						
	PBP	COC	CTP	1	2	3	4	5	6	7
APSCFE	2.90	15.1	13.1	Cl_2	Cl_1	O_2	O_1	N_5	N_2	N_1
				2.89	3.73	2.36	3.11	2.90	2.58	2.44
GEZMED	12.1	14.5	13.2	O_8	O_7	O_5	O_4	Cl_1	O_3	O_1
				4.64	4.96	16.3	11.4	11.4	17.8	10.3
JOCTOK	6.10	12.6	10.7	O_5	O_3	N_3	N_2	N_1	O_1	O_7
				8.50	7.89	2.54	5.99	7.47	4.69	2.46
LIEDFE	5.18	14.5	12.0	O_4	O_1	O_6	O_3	O_5	N_2	N_1
				7.62	8.00	1.02	4.37	2.33	4.53	4.42

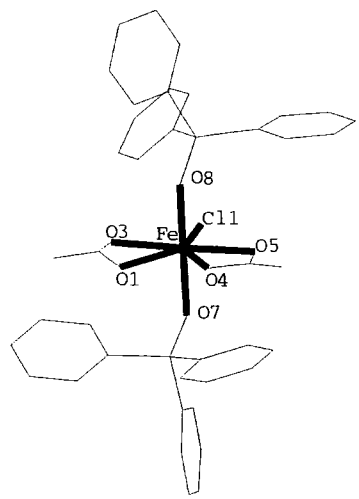
Complex APSCFE $[FeCl_2(DAPS)]Cl{\cdot}H_2O$ (Palenik, Wester, Rychlewska & Palenik, 1976) has a nearly perfect PBP geometrical coordination sphere and all the valence bond angles L-M-L are very close to the standard values. The relative $R_{ang}(PBP)$ gives 2.90% and $R_{ang}(CTP)$ 13.7% and this clearly allows the FeL_7 to be assigned as a PBP polyhedron. The five ligand atoms in 2,6-diacetylpyridine-bis(semicarbazone) (DAPS) are coordinated with *Fe(III)* and all located on the equatorial plane. Since the conformation of the molecule DAPS is suitable for forming a pentagon with *Fe(III)* atom and the size of *Fe(III)* can be fitted into the pentadentate hole, two chlorine atoms are located at the axial sites. Each ligand atom is nearly at an ideal vertex site of PBP polyhedron, therefore, the VR_{ang} indices are also small (all of them are < 4.0%).

For complex GEZMED $[Fe(NO_3)_2(OPPh_3)_2Cl]$ (Tomi, Wah & Postel, 1988), it is difficult to see which polyhedral form is suitable for the description of geometry from the $R_{ang}(x)$ values, where all three values are all greater than 10%. But the smallest one appears on the PBP. The VR_{ang} values indicate that the coordination sphere is a distorted PBP because of the bidentate nitrate ligands. The calculation on this PBP R_{ang} value showed that the corresponding oxygen atoms, O_1, O_3, O_4 and O_5, of two nitrate groups occupy sites in the equatorial plane. The chelated angles (*50.3°* and *56.1°*) between two

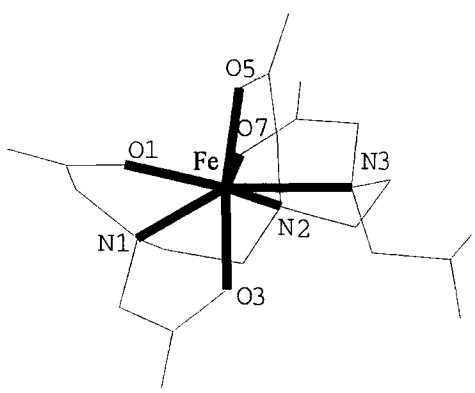
oxygen and the *Fe* atoms (*O-Fe-O'*) are considerably less than the standard angle of 72.0° due to the strain in the bite angle. However two other unidentate ligands (O_7 and O_8) from triphenyl phosphine oxide locate at the uncrowded axial positions in the PBP with angle, O_7-*Fe*-O_8, of *170.5°*. The oxygen atoms in the nitrate groups are obviously a dominant factor to make the geometry of the complex distort from a PBP polyhedron.



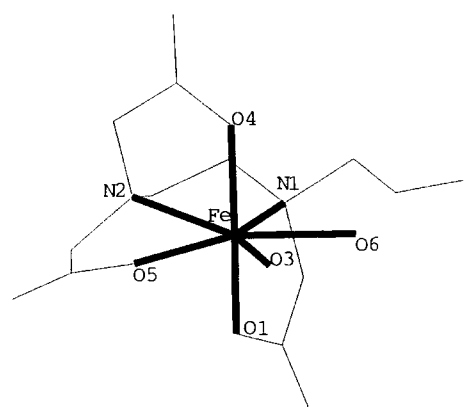
APSCFE



GEZMED



JOCTOK



LIEDFE

Figure 4.6 (a) Structures from the CSD for Fe^{3+} complexes.

The diagram only shows the coordination spheres. Chemical formulae are: **APSCFE**: $[Fe(DAPS)Cl_2]Cl \cdot 2H_2O$, **GEZMED**: $[Fe(NO_3)_2(OPPh_3)_2Cl]$, **JOCTOK**: $(NEt_4)_2[Fe(DTPA)]$, **LIEDFE**: $Li[Fe(OH2)(ETA)] \cdot 2H_2O$, ETA=1,2-Ethylenediamimetetra-acetato.

The corresponding VR_{ang} indexes reflect this ligand influence. The two VR_{ang} indexes for the axial atoms are small, whereas those for the ligands on the equatorial plane are very large, especially for the oxygen atoms, O_3 and O_5, of nitrates.

Two other structures in PBP geometry with Refcode JOCTOK and LIEDFE (Finnen, Pinkerton, Dunham, Sands & Funk, 1991; Lind, Hamor & Hoard, 1964) are coordinated by multi-dentate ligands. The distortions brought about by the chelate ring are such that the ligands at axial positions (*1* and *2*) lean away from the standard line to

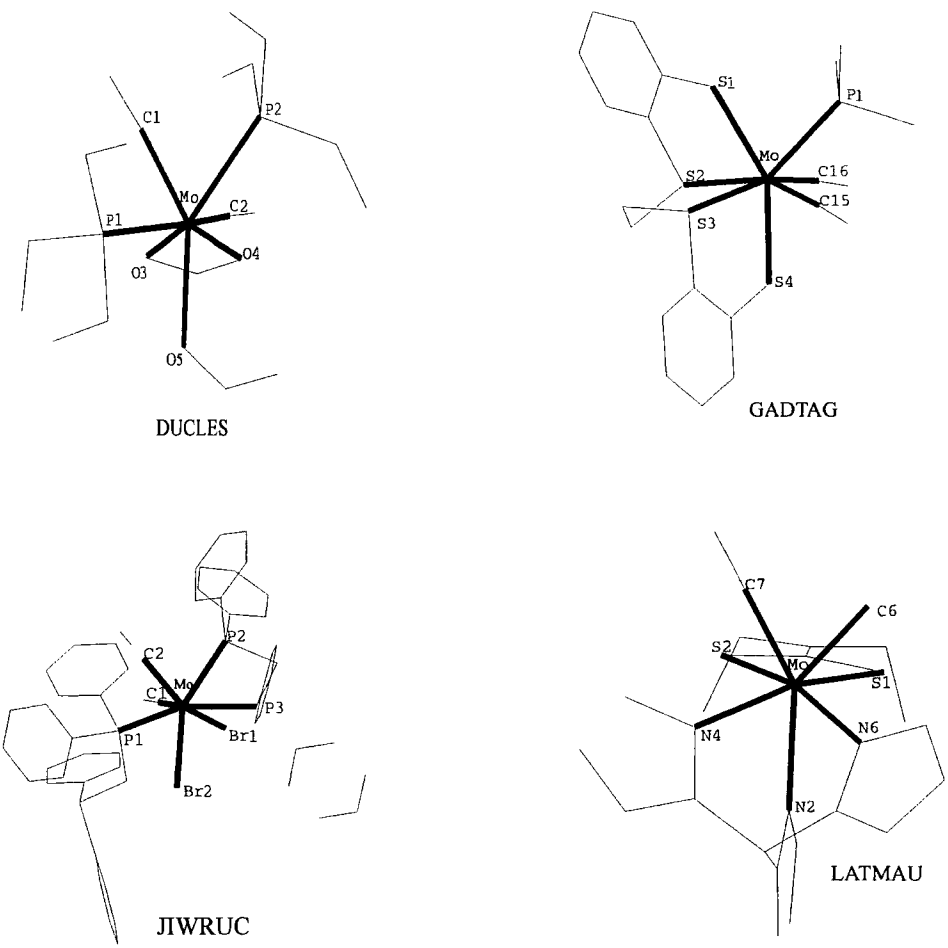


Figure 4.6(b) Structures from the CSD for Mo^{2+} complexes

Chemical formulae, **DUCLES**: $Mo(CO)_2(PEt_3)_2(OOCH)(OCHO)$;
GADTAG: $Mo(CO)_2(S_4C_{14}H_{12})(PMe_3)$, $S_4C_{14}H_{12}$=2,3:8,9-dibenzo-1,4,7,10-tetra-Thiadecane; **JIWRUC**: $Mo(CO)_2(\eta^2\text{-dppm})(\eta^1\text{-dppm})Br_2$;
LATMAU: $Mo(CO)_2(S_2CNEt_2)(HBpz_3)$, $HBpz_3$=Hydrogen tris(1-pyrazolyl)borato.

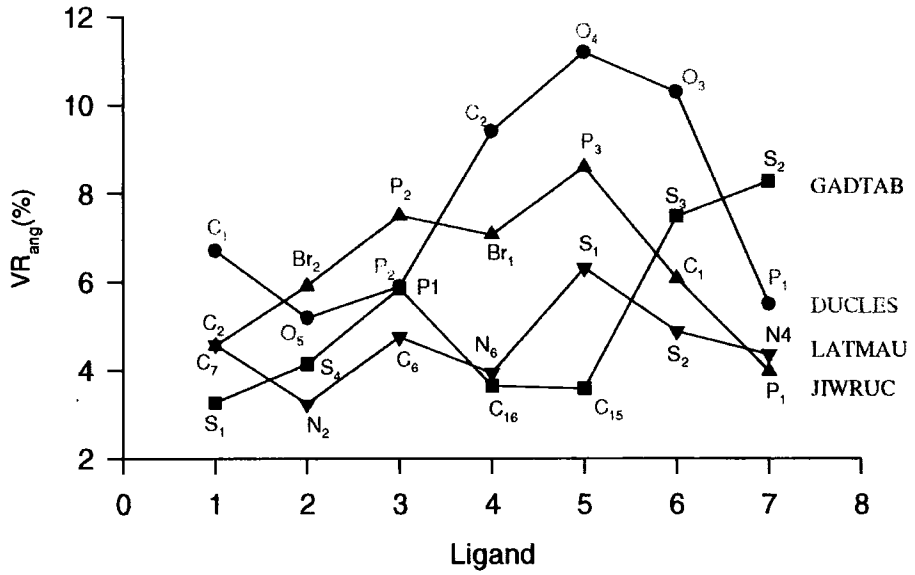


Figure 4.7 VR_{ang} (%) on each ligand

each other. In both molecules, the angles O_{ax}-Fe-O_{ax} are 168.6° and 165.5°, respectively, which deviate considerably from the standard value. All these can be observed from VR_{ang} values in Table 4.4. The VR_{ang} indexes for the axial ligands are clearly larger than others.

In the *Mo* coordination spheres, the VR_{ang} values for some complexes [DUCLES, GADTAG, JIWRUC and LATMAU, as showed in Figure 4.6(b)], with CTP geometry are shown in Figure 4.7. It illustrates the similar influence of each ligand on the coordination sphere geometry. It can be seen that the ligand atoms, O_3 and O_4 in DUCLES (Brower, Winston,Tonker and Templeton, 1986), P_1 and P_2 in JIWRUC (Shiu, Yih, Wang and Liao, 1991), S_1 and S_2 in LATMAU (Shiu, Lee, Wang & Cheng, 1993) on the bidentate ligands, and S_2 and S_3 in GADTAG (Shellmann, Keller, Moll, Campana & Haase, 1988) on the tetradentate ligand all have larger distortions from the standard positions and also affect the neighbour atoms, especially those on the quadraplane.

4.2.5 Multivariate Analysis in ML_7

By now, the geometry of the ML_7 sphere fragment for each extracted structural data on seven-coordinate complexes of various transition metals has been identified in terms of $R_{ang}(x)$ values. These $R_{ang}(x)$ values refer to the three idealized polyhedral forms, PBP, COC and CTP and the closest conformation is obtained over the calculation on 5040 possible permutations of ligand atoms. This allows multivariate analysis to be carried out on the resultant *R*-classified data set and mapping of each geometrical cluster from relative coordinates reveals distortions from the ideal points and thereby the expected coordinates from one geometry to another are mapped out. In Figure 4.5, some maps on the specific pair of angles have been given, which showed clusters belonging to different geometry and those conformations distorted most significantly from these clusters and may provide a reaction pathway for transformation between these polyhedral forms. Multivariate analysis is considered to be a more comprehensive analysis, because more complete geometrical descriptions are involved.

Multivariate analysis will result in the coordination geometries being expressed by means of a set of PC or FA factors for both PBP and CTP conformations. Thus, a single point represents a specific ML_7 fragment in a PC's or factor scatterplot. Of course, the dimensional numbers in two different geometrical spaces (PBP-space and CTP space) are expected to be less than twenty-one (the number of valence angles of an ML_7 fragment). In a PBP-space, an ideal PBP is at the centre and each structure is related to PBP and in a CTP-space, each structure is related to a CTP.

As stated in the previous chapter, the multivariate analysis is greatly dependent upon the symmetry of the geometrical spaces. These are D_{5h} for PBP and C_{2v} for CTP. Any real structure can appear as one of its isomers related by symmetry operations of the point group in both spaces. In order to obtain a complete data expression and represent the symmetries of the geometrical spaces, each observed conformation would be expanded into *20* and *4* isomers according to the symmetry operation of the point groups D_{5h} and C_{2v}, respectively. Therefore, a larger multivariate matrix will be generated.

4.2.5.1 Symmetry in ML_7

The observed molecular geometries for ML_7 spheres are referred to idealized PBPs and CTPs or COCs, of D_{5h} and C_{2v} or C_{3v} symmetry, respectively. For each fragment whose geometry was determined by the $R_{ang}(x)$ value, there are *20* or *4* possible isometric arrangements of the ligands for the PBP and CTP, respectively. These isomers are related to each other by the symmetry operations of the D_{5h} and C_{2v} point groups. The starting point is from the ligand arrangements resulting from the smallest $R_{ang}(x)$ value, which represents a minimum distortion from either PBP or CTP. Since the order in which the *R*-determined ligand lists is consistent with that in the reference geometry, this isomer is defined as number one by the symmetry operation *E*, any other symmetrical isomers being related to this one by a symmetry operation. Figure 4.8(a) and (b) illustrate symmetrical isomers related by corresponding symmetry operations in the two point groups, respectively. Any observed structure can be represented in either of the 4 or 20 asymmetric units for the two different geometrical spaces.

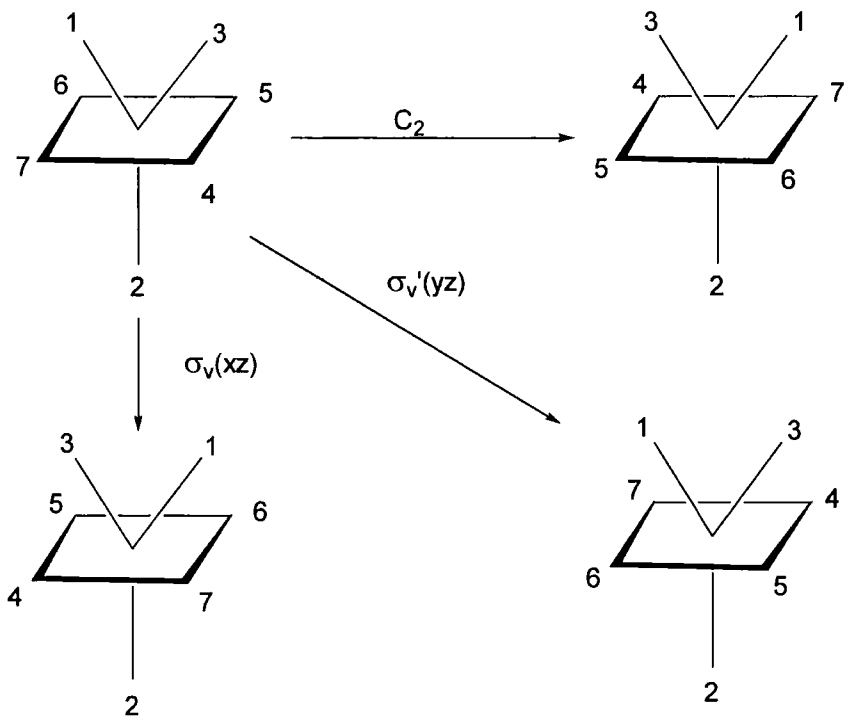


Figure 4.8(a) Symmetry operations in C_{2v}

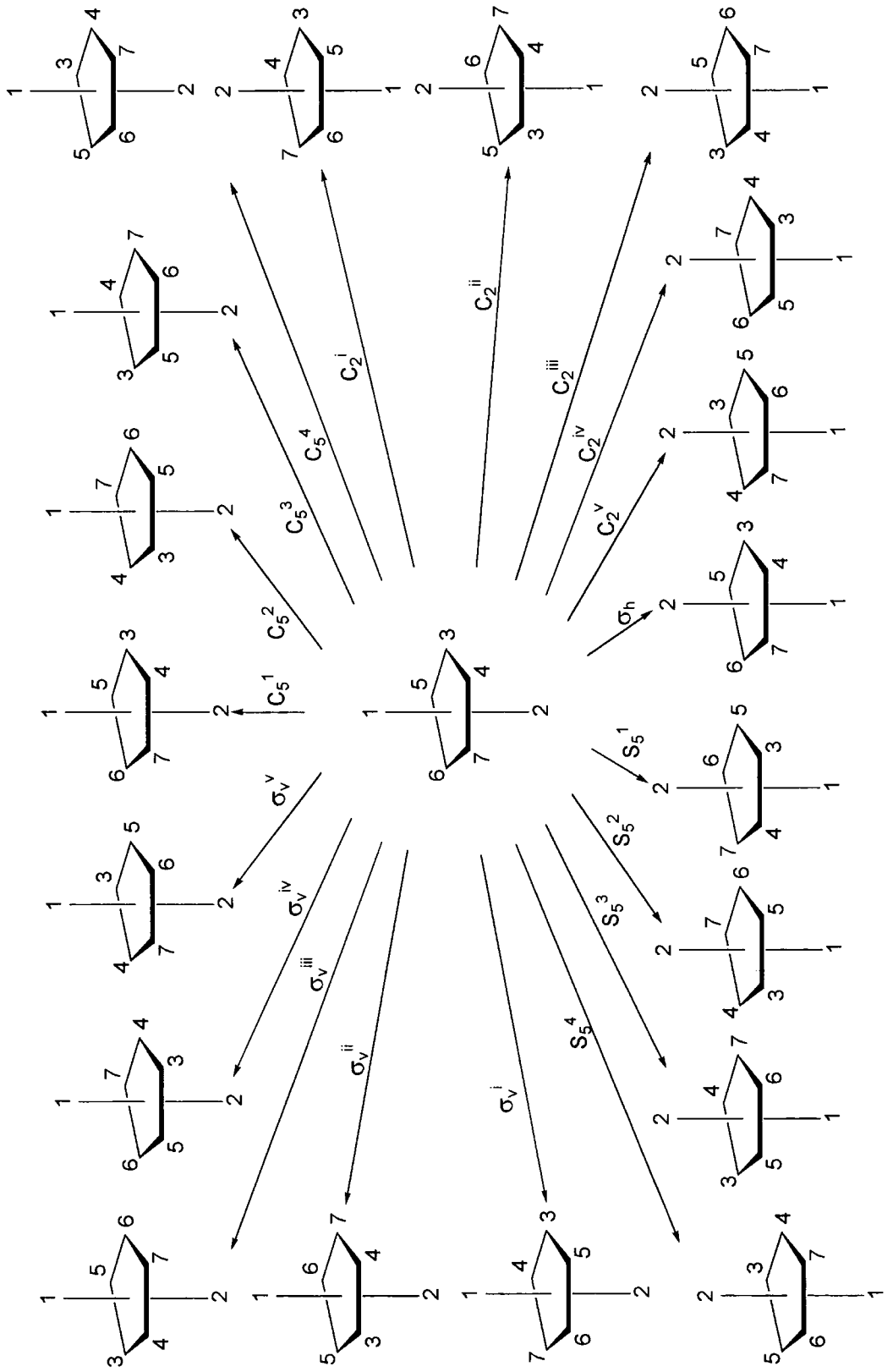


Figure 4.8 (b) Symmetry operations in D_{5h}

For a systematic analysis of ML_n systems, labeling the *n* ligands should ensure that the representative points are all restricted to the same asymmetric unit and not randomly distributed among the available subspaces. Indeed, with the *R*-determined geometry method, the final order of ligands related to a reference geometry brings a corresponding symmetry to the parameter space (Murray-Rust, 1982). This was achieved by *n!* permutation. The complete symmetry expanded data used in statistical analyses will present all possible symmetry in the data distribution. Thus, in the ML_7 case, each observed structure (or data point) is expanded by a factor equal to the order of the point group to which the structure is referred. In the PBP geometrical space, each observed structure is multiplied by 20 and in the CTP geometrical space, each data is expanded to 4. In doing so, an expanded data matrix is obtained for further multivariate analyses. It should be emphasized here that no 'new' data have been added, but that the existing data is expanded according to the related symmetry to fill the appropriate space. All these processes can be accomplished in the FORTRAN program.

4.2.5.2 Symmetry Coordinates

Some advantages of using symmetry coordinates for describing a displacement of an observed structure from the more symmetrical reference structures have been discussed in Chapter 2. Murray-Rust *et al* (1978, 1979) have generalized the approach to describe distortions from a reference structure of any given symmetry.

In the ML_7 system, it is more complex to express the symmetry coordinates with *21*-dimensional basic bond angle coordinates. For an *N* atom system, *3N-6* such coordinates normally result. There should be *18* internal coordinates for an ML_7 sphere (*7* ligands plus a metal atom). If only bond angles are considered, *11* independent symmetry coordinates should be obtained. Before these complete coordinates are given, we firstly visualize the symmetry properties of this system by specifying a simplified lower dimensional space instead of using a *21*-dimensional representation of the angle coordinates.

In the case of seven-coordinated complexes, whatever polyhedral form the structure adopts among the three modes (PBP, COC and CTP), for an ideal geometrical form, it can be seen that the ligands at position *1*, *2* and *3* (see Figure 3.2) are always kept on the same plane and have significant differences in the angles from those ligands

on the equatorial plane. Although the ligand at the position *3* in the polyhedron PBP has the same environment as other four ligands on the equatorial plane, it is perpendicular to the line formed by the positions *1* and *2*. The special value of *90°* is an important criterion distinguishing the PBP from other modes. Therefore, an investigation can put emphasis on these three ligands to see how they can be used to describe distortions from the reference polyhedra of a given symmetry.

The ligands at the positions *1*, *2* and *3* retain a C_{2v} symmetry in these idealized geometrical modes (PBP and CTP) for the defined plane. Thus, a set of symmetry coordinates for this symmetry will be derived. If the two-fold axis is aligned with the *z*-axis (see Figure 4.9), the irreducible components can be written as:

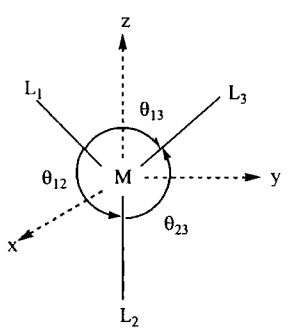
$$\Gamma_\theta = 2A_1 + B_2$$

So, the symmetry coordinates related to the angles are given as follows:

$$s_1 = \frac{1}{\sqrt{3}}(\theta_{13} + \theta_{12} + \theta_{23})$$

$$s_2 = \frac{1}{\sqrt{6}}(2\theta_{13} - \theta_{12} - \theta_{23})$$

$$s_3 = \frac{1}{\sqrt{2}}(\theta_{12} - \theta_{23})$$



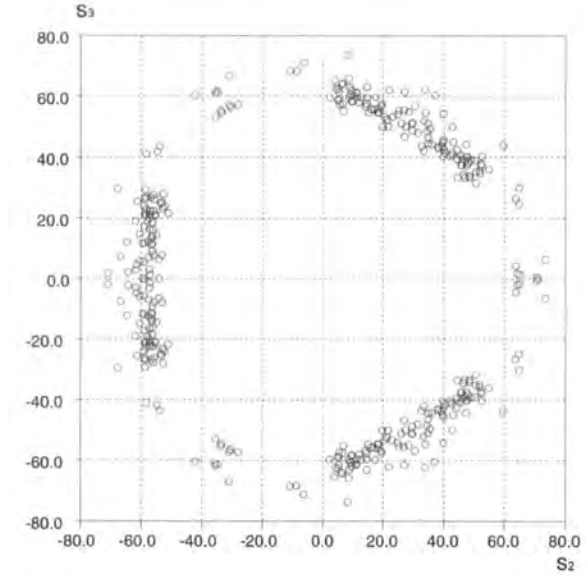
C_{2v}	E	C_2	$\sigma_v(xz)$	$\sigma_v'(yz)$
A_1	1	1	1	1
A_2	1	1	-1	-1
B_1	1	-1	1	-1
B_2	1	-1	-1	1
Γ_θ	3	1	1	3

Figure 4.9 Symmetry in C_{2v}

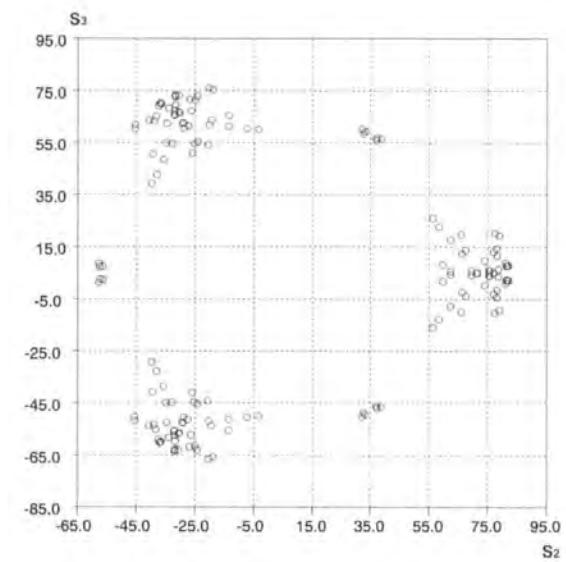
The first two symmetry displacement coordinates s_1 and s_2 transform as A_1, which are linear combinations of the original set of basic vectors derived directly from irreducible representation of A_1, i.e. $s_{1a} = \theta_{13}$ and $s_{2a} = {}^{1}/_{\sqrt{2}}(\theta_{12} + \theta_{23})$. The kernel symmetries are $\mathbf{K}(A_1) = C_{2v}$, $\mathbf{K}(B_2) = C_s(\sigma_v')$. There are no co-kernel symmetries (McDowell, 1965). Since only three atoms are involved in this symmetry coordinate system, there are only 6 different ways of labeling these atoms. Starting with a set of ligand arrangements corresponding to the smallest R_{ang} value, the angles are expanded relative to these three atoms. A high symmetry cluster pattern in which three kinds of geometries can be identified is yielded. In the case of symmetry COC, the corresponding three ligands to those on the same plane only have symmetry C_s. But by using the symmetry coordinates resulting from C_{2v}, this polyhedral form can be still visualized as an intermediate in the transformation between PBP and CTP. Three types of geometries can therefore be characterized in the two-dimensional subspace of the symmetry coordinates.

From the point of view of vibrations, s_2 and s_3, the two coordinates represent the motions of three ligands on the plane. s_2 indicates that for the three angles on the plane, one increases while the other two reduce, or vice versa. This means that the ligands at position 1 and 3 move close to or apart from each other; s_3 gives the changes of the two angles related by a 2-fold axis or mirror plane perpendicular to the plane, one stretches and the other shrinks. This corresponds to distortions of ligand *1* and *3* in the same direction. All these also represent geometrical change in three polyhedral forms. Thus, the s_2 versus s_3 plots show three clusters which belong to the three geometries in a similar pattern to the θ_{23} - θ_{12} plots arising from the discussions of the potential energy surface (see Section 4.2.3).

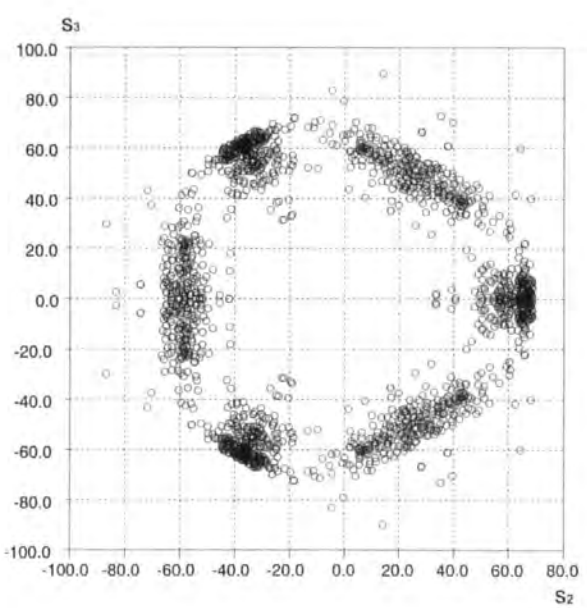
In Figure 4.10(a) for all unidentate ligand complexes, the PBP clusters are located around (s_2 = *71.3*, s_3 = *0.0*; s_2 = *-35.6*, s_3 = *61.8*; s_2 = *-35.6*, s_3 = *61.8*) and the CTP around (s_2 = *-58.9*, s_3 = *0.0*; s_2 = *29.4*, s_3 = *-51.0*; s_2 = *29.4*, s_2 = *-51.0*). The COC form appears between these two forms. Figure 4.10(b) gives the plot for all seven-coordinate *Fe* complexes in the data set, which shows two major clusters in PBP and CTP forms in this data subset. Figure 4.10(c) includes all seven coordinate complexes. Since the symmetry of three ligand atoms is considered, all the points on the plots can be seen as related by a mirror plane that is perpendicular to the paper and parallel to the s_2 axis and



(a)



(b)



(c)

Figure 4.10 s_3 *vs* s_2 for(a) Unidentate, (b) Fe and (c)all seven-coordinate complexes

through the origin of axis s_3. This can also be seen as a two-fold symmetry axis in two-dimensions.

All these coordinates are derived from the three atoms on the specific plane, by which the distortions and symmetry described are all dependent on the movements of these three atoms. In order to investigate all contributions of seven ligands to the distortions of reference geometry, complete symmetry coordinates from all *21* internal bond angles in an ML_7 sphere are needed.

Application of group theoretical methods, as in the C_{2v} case given above, to the ML_7 sphere, firstly in point group D_{5h}, by using the symmetry operations shown in Figure 3.8(a), gives the irreducible representation as:

$\mathbf{D_{5h}}$	E	$2C_5$	$2C_5^2$	$5C_2$	σ_h	$2S_5$	$2S_5^3$	$5\sigma_v$
Γ_θ	21	1	1	3	11	1	1	5

and irreducible components which can be written as:

$$\Gamma_\theta = 4A_1' + 3E_1' + 3E_2' + A_2'' + E_1'' + E_2''$$

The internal coordinates chosen for this problem are the 21 internal angles. Normally, for these 8 atom system (7 ligands plus one metal atom) system, there should be 3×8−6 = 18 genuine symmetry coordinates that can be derived. If seven bond lengths are excluded (fixed), 11 independent coordinates should be obtained. As there are 21 internal coordinates as start, an equal number (*21*) of symmetry coordinates can be constructed according to the regulations given above, as to the derivation of symmetry coordinates. Moreover, six of them are actually not independent among these internal angles. Ten derived coordinates can be considered as redundant. Four of these belong to the species A_1', four to the species E_1' and two to the species E_2'. Thus, the final results for various representations are (including bond lengths r_1, r_2, r_3, r_4, r_5, r_6 and r_7.) are :

$$S_1(A_1^{'}) = \frac{1}{\sqrt{5}}(r_3 + r_4 + r_5 + r_6 + r_7);$$

$$S_2(A_1^{'}) = \frac{1}{\sqrt{2}}(r_1 + r_2)$$

$$S_3(A_2^{''}) = \frac{1}{\sqrt{2}}(r_1 - r_2)$$

$$S_4(A_2^{''}) = \frac{1}{\sqrt{10}}[(\theta_{13} - \theta_{23}) + (\theta_{14} - \theta_{24}) + (\theta_{15} - \theta_{25}) + (\theta_{16} - \theta_{26}) + (\theta_{17} - \theta_{27})]$$

$$S_{5a}(E_1^{'}) = \sqrt{\frac{2}{5}}[r_3 + (r_4 + r_5)\cos\omega + (r_6 + r_7)\cos 2\omega]$$

$$S_{5b}(E_1^{'}) = \sqrt{\frac{2}{5}}[(r_5 - r_4)\sin\omega + (r_6 - r_7)\sin 2\omega]$$

$$S_{6a}(E_1^{'}) = \sqrt{\frac{2}{5}}[\theta_{34} + (\theta_{35} + \theta_{47})\cos\omega + (\theta_{56} + \theta_{67})\cos 2\omega]$$

$$S_{6b}(E_1^{'}) = \sqrt{\frac{2}{5}}[(\theta_{35} - \theta_{47})\sin\omega + (\theta_{56} - \theta_{67})\sin 2\omega]$$

$$S_{7a}(E_1^{'}) = \frac{1}{\sqrt{5}}[(\theta_{13} + \theta_{23}) + (\theta_{14} + \theta_{24} + \theta_{15} + \theta_{25})\cos\omega + (\theta_{16} + \theta_{26} + \theta_{17} + \theta_{27})\cos 2\omega]$$

$$S_{7b}(E_1^{'}) = \frac{1}{\sqrt{5}}[(\theta_{15} + \theta_{25} - \theta_{14} - \theta_{24})\sin\omega + (\theta_{16} + \theta_{26} - \theta_{17} - \theta_{27})\sin 2\omega]$$

$$S_{8a}(E_2^{'}) = \sqrt{\frac{2}{5}}[\theta_{34} + (\theta_{35} + \theta_{47})\cos 2\omega + (\theta_{56} + \theta_{67})\cos\omega]$$

$$S_{8b}(E_2^{'}) = \sqrt{\frac{2}{5}}[(\theta_{35} - \theta_{47})\sin 2\omega - (\theta_{56} - \theta_{67})\sin\omega]$$

$$S_{9a}(E_2^{'}) = \sqrt{\frac{2}{5}}[r_3 + (r_4 + r_5)\cos 2\omega + (r_6 + r_7)\cos\omega]$$

$$S_{9b}(E_2^{'}) = \sqrt{\frac{2}{5}}[(r_5 - r_4)\sin 2\omega - (r_6 - r_7)\sin\omega]$$

$$S_{10a}(E_1^{''}) = \frac{1}{\sqrt{5}}[(\theta_{13} - \theta_{23}) + (\theta_{15} - \theta_{25} + \theta_{14} - \theta_{24})\cos\omega + (\theta_{16} - \theta_{26} + \theta_{17} - \theta_{27})\cos 2\omega]$$

$$S_{10b}(E_1'') = \frac{1}{\sqrt{5}}[(\theta_{15} - \theta_{25} - \theta_{14} + \theta_{24})\sin\omega + (\theta_{16} - \theta_{26} - \theta_{17} + \theta_{27})\sin 2\omega]$$

$$S_{11a}(E_2'') = \frac{1}{\sqrt{5}}[(\theta_{13} - \theta_{23}) + (\theta_{15} - \theta_{25} + \theta_{14} - \theta_{24})\cos 2\omega + (\theta_{16} - \theta_{26} + \theta_{17} - \theta_{27})\cos\omega]$$

$$S_{11b}(E_2'') = \frac{1}{\sqrt{5}}[(\theta_{15} - \theta_{25} - \theta_{14} + \theta_{24})\sin 2\omega - (\theta_{16} - \theta_{26} - \theta_{17} + \theta_{27})\sin\omega]$$

where $\omega = \frac{2\pi}{5}$.

The kernel (***K***) and co-kernel (***CoK***) of these representations are given as (Murray-Rust, Bürgi and Dunitz, 1979).

	Γ_i	***K***	***CoK***
D5h	A_2''	C_{5v}	--
	E_1'	$C_s(\sigma_h)$	C_{2v}
	E_2'	$C_s(\sigma_h)$	C_{2v}
	E_1''	C_1	$C_s(\sigma_v),C_2$
	E_2''	C_1	$C_s(\sigma_v),C_2$

All these coordinates belonging to different representations can adopt such symmetries and will be used later for the PCA analysis.

In the same way, when the system is considered in C_{2v} symmetry for the whole ML_7 sphere, the irreducible representations can be obtained from the symmetry operations shown as in Figure 3.8(b) and are given as:

C_{2v}	E	C_2	$\sigma_v(xz)$	$\sigma_v'(yz)$
Γ_θ	21	3	3	5

with the irreducible components :

$$\Gamma_\theta = 8A_1 + 4A_2 + 4B_1 + 5B_2 .$$

The *11* independent symmetry coordinates from a linear combination of bond angles, and subject to the condition of orthonormality, are listed below: plus 7 symmetry coordinates derived from bond lengths.

$$S_1(A_1) = \frac{1}{\sqrt{3}}(r_1 + r_2 + r_3);$$

$$S_2(A_1) = \frac{1}{\sqrt{6}}(2r_2 - r_1 - r_3)$$

$$S_3(A_1) = \frac{1}{2}(r_4 + r_5 + r_6 + r_7);$$

$$S_4(A_1) = \frac{1}{\sqrt{6}}(2\theta_{13} - \theta_{12} - \theta_{23})$$

$$S_5(A_1) = \frac{1}{2}(\theta_{16} + \theta_{17} + \theta_{34} + \theta_{35});$$

$$S_6(A_1) = \frac{1}{\sqrt{2}}(\theta_{46} + \theta_{57})$$

$$S_7(A_2) = \frac{1}{\sqrt{3}}(r_4 - r_5 + r_6 - r_7);$$

$$S_8(A_2) = \frac{1}{\sqrt{2}}(\theta_{46} - \theta_{57})$$

$$S_9(A_2) = \frac{1}{2}(\theta_{16} - \theta_{17} + \theta_{34} - \theta_{35});$$

$$S_{10}(A_2) = \frac{1}{2}(\theta_{24} - \theta_{25} + \theta_{26} - \theta_{27})$$

$$S_{11}(B_1) = \frac{1}{2}(r_4 - r_5 - r_6 + r_7);$$

$$S_{12}(B_1) = \frac{1}{2}(\theta_{16} - \theta_{17} - \theta_{34} + \theta_{35})$$

$$S_{13}(B_1) = \frac{1}{2}(\theta_{24} - \theta_{25} - \theta_{26} + \theta_{27});$$

$$S_{14}(B_2) = \frac{1}{\sqrt{2}}(r_1 - r_3)$$

$$S_{15}(B_2) = \frac{1}{2}(r_4 + r_5 - r_6 - r_7);$$

$$S_{16}(B_2) = \frac{1}{\sqrt{2}}(\theta_{12} - \theta_{23})$$

$$S_{17}(B_2) = \frac{1}{2}(\theta_{16} + \theta_{17} - \theta_{34} - \theta_{35});$$

$$S_{18}(B_2) = \frac{1}{2}(\theta_{24} + \theta_{25} - \theta_{26} - \theta_{27})$$

It can be seen that the symmetry coordinates corresponding to the three ligands on the specific plane given previously (s_2 and s_3) are also included in this symmetry coordinate set. The presentations of these two coordinates, as shown in Figure 4.10, are indeed a part of this complete coordinate set. With respect to movements of these ligands leading to a different geometry, the displacements along this pair of symmetry

coordinates, which represent distortions from three reference geometries, were clearly visualized from the s_3-s_2 projections.

The kernel symmetry of representations in this point group are:

	Γ_i	K
C_{2v}	A_1	C_{2v}
	A_2	C_2
	B_1	$C_s(\sigma_v)$
	B_2	$C_s(\sigma_v)$

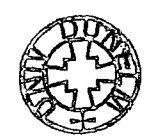
A displacement along any single coordinate produces a configuration with the kernel symmetry. The displacements for some of these symmetry coordinates corresponding to the bond lengths and angles are displayed in Figure 4.11.

The symmetry coordinates related to the bond lengths are classified as two groups that correspond to ligands at apical and equatorial positions, respectively. S_2 and S_3 in D_{5h}, and S_1, S_2 and S_{14} in C_{2v} are linear combinations of the bond lengths connected to the apical positions, r_1 and r_2 in D_{5h}, and plus r_3 in C_{2v}. All of the remainder are combinations of equatorial bond lengths, r_4, r_5, r_6 and r_7.

Table 4.5(a) and (b) list the correlation matrices for those symmetry coordinates of all 7-coordination complexes in the two point group spaces. As was stated previously, symmetry coordinates are constructed according to the irreducible representations of the point group and orthonormality basis. Therefore, symmetry coordinates derived from different symmetry species should ideally be uncorrelated. The correlation between the symmetry coordinates may show whether the derived coordinates do in fact form a basis for the irreducible representation (Auf der Heyde & Bürgi, 1989a).

In this calculation, the symmetry coordinates from different symmetry species are clearly not correlated. The correlated ones all belong to the same symmetry species, which means that symmetry coordinates derived from both point groups in accordance with the conditions for forming the basis of irreducible representation, and thereby all the analyses based on the symmetry coordinates will be reliable. All the correlation coefficients between different symmetry species are zero, or small, and can be ignored here. So these values are omitted from the tables for simplicity.

The correlation between symmetry coordinates can easily be explained by an investigation of the particular distortion that pairs of correlated symmetry coordinates represent (shown in Figure 4.11). In the C_{2v} point group, of more interest, it is the



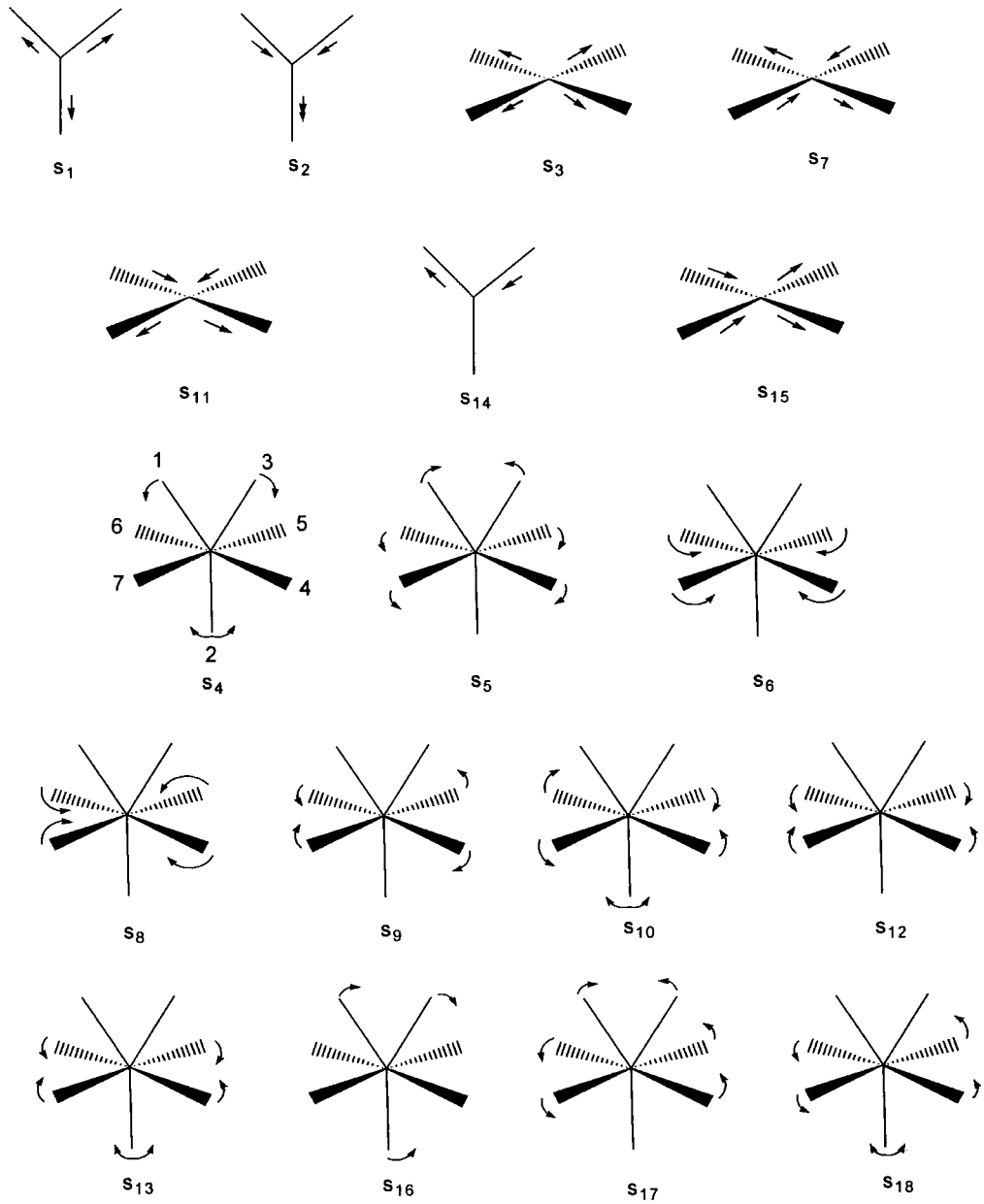


Figure 4.11 Graphical representations of the 18 symmetry coordinates in C_{2v}

correlation of S_{16} and S_{17} ($R = 0.89$), related to the symmetry species B_2. S_{16} represents that angle θ_{13} increases, resulting from ligand *1* moving towards the axial position and *3* down to the quadrilateral plane. S_{17} is indicative of angles above the quadrilateral plane increasing (θ_{34} and θ_{35}) and decreasing (θ_{36} and θ_{37}). These distortions represent the

transformation from C_{2v} to D_{5h} symmetry, which is the reverse pathway of that shown in Figure 3.9.

Table 4.5(a) Correlation matrix for symmetry coordinates in C_{2v}

A_1						
	S_1	S_2	S_3	S_4	S_5	S_6
S_1	1.00					
S_2	-0.14	1.00				
S_3	0.61	-0.23	1.00			
S_4	-0.09	-0.32	0.28	1.00		
S_5	-0.06	-0.12	0.03	0.05		
S_6	0.03	0.40	-0.24	-0.71	-0.56	1.00

B_2					
	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}
S_{14}	1.00				
S_{15}	-0.01	1.00			
S_{16}	-0.50	-0.08	1.00		
S_{17}	-0.55	-0.13	0.89	1.00	1.00
S_{18}	-0.17	-0.04	-0.11	0.03	

A_2				
	S_7	S_8	S_9	S_{10}
S_7	1.00			
S_8	0.01	1.00		
S_9	-0.29	-0.24	1.00	
S_{10}	-0.10	0.45	-0.46	1.00

B_1			
	S_{11}	S_{12}	S_{13}
S_{11}	1.00		
S_{12}	-0.04	1.00	
S_{13}	-0.42	0.13	1.00

(b) Correlation matrix for symmetry coordinates in D_{5h}

E_1'						
	S_{5a}	S_{5b}	S_{6a}	S_{6b}	S_{7a}	S_{7b}
S_{5a}	1.00					
S_{5b}	0.00	1.00				
S_{6a}	0.03	-0.02	1.00			
S_{6b}	0.02	0.03	0.00	1.00		
S_{7a}	-0.30	0.00	-0.51	0.37	1.00	
S_{7b}	0.00	-0.30	0.37	-0.51	0.00	1.00

E_2'				
	S_{8a}	S_{8b}	S_{9a}	S_{9b}
S_{8a}	1.00			
S_{8b}	0.00	1.00		
S_{9a}	-0.09	-0.29	1.00	
S_{9b}	0.29	-0.09	0.00	1.00

A_1'		
	S_1	S_2
S_1	1.00	
S_2	0.49	1.00

A_2''		
	S_3	S_4
S_3	1.00	
S_4	-0.19	1.00

E_1''		
	S_{10a}	S_{10b}
S_{10a}	1.00	
S_{10b}	0.00	1.00

E_2''		
	S_{11a}	S_{11b}
S_{11a}	1.00	
S_{11b}	0.00	1.00

In the case of D_{5h}, a major correlation involves pairs of S_6 and S_7, which each consists of two degenerates, that is, S_{6a} and S_{6b}; S_{7a} and S_{7b}. The distortions represented by these four symmetry coordinates are given in Figure 4.12. S_6 shows the in-plane vibrations of atoms on the equatorial plane, and S_7 vertical to the plane. S_{7a} also

indicates that the axial atoms *1* and *2* move towards the plane. The negative correlation with S_{6a} (R = *-0.51*) implies that with this movement, atom *6* and *7* should not move close to each other, which is represented in S_{6a}.

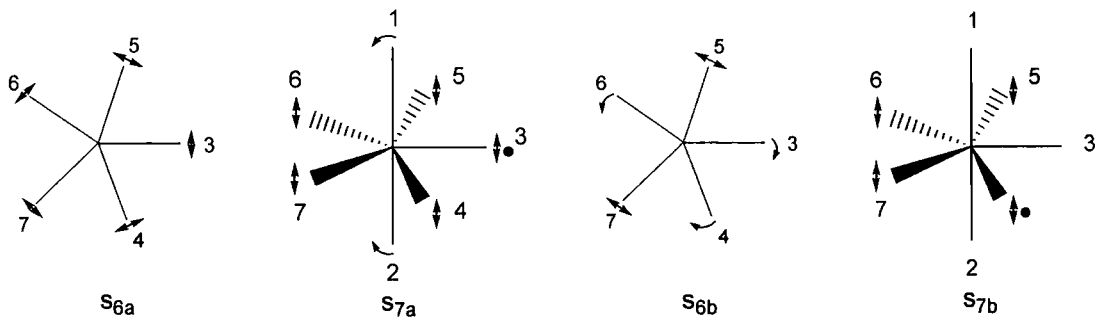


Figure 4.12 Graphical representations of S_6 and S_7 in D_{5h} (← means larger coefficient on corresponding angles)

4.2.5.3 PCA and FA

The principles of PCA and FA and applications in the geometrical systems analysis for specific molecular fragments have been given briefly in the previous chapters. Both techniques have been also performed on the *21* valence angles of the symmetry-expanded data matrix for ML_7 sphere. The results given below show again that these techniques are a powerful tools in the analysis of general geometrical effects for this system.

The present CSD software is not available for doing PCA for the ML_7 dataset due to the overlarge data matrix. The necessary label permutations for this systems analysis have been carried out in the $R_{ang}(x)$ program given before. From the determined geometry, 21 related valence angles are further expended by the symmetry operations and then the program is linked to the PCA or FA subroutine provided by the NAG Fortran library (G03AAF for PCA and G03CAF, G03CCF for FA) (NAG, 1997). Input matrix data is converted into a suitable form for these subroutines and some important parameters for doing these calculations are given in Appendix IV. This process can be expressed simply in a flow chart (Figure 4.13). Finally, the PC or factor scores etc., are returned and recorded in the same format as those required by plotting packages, such

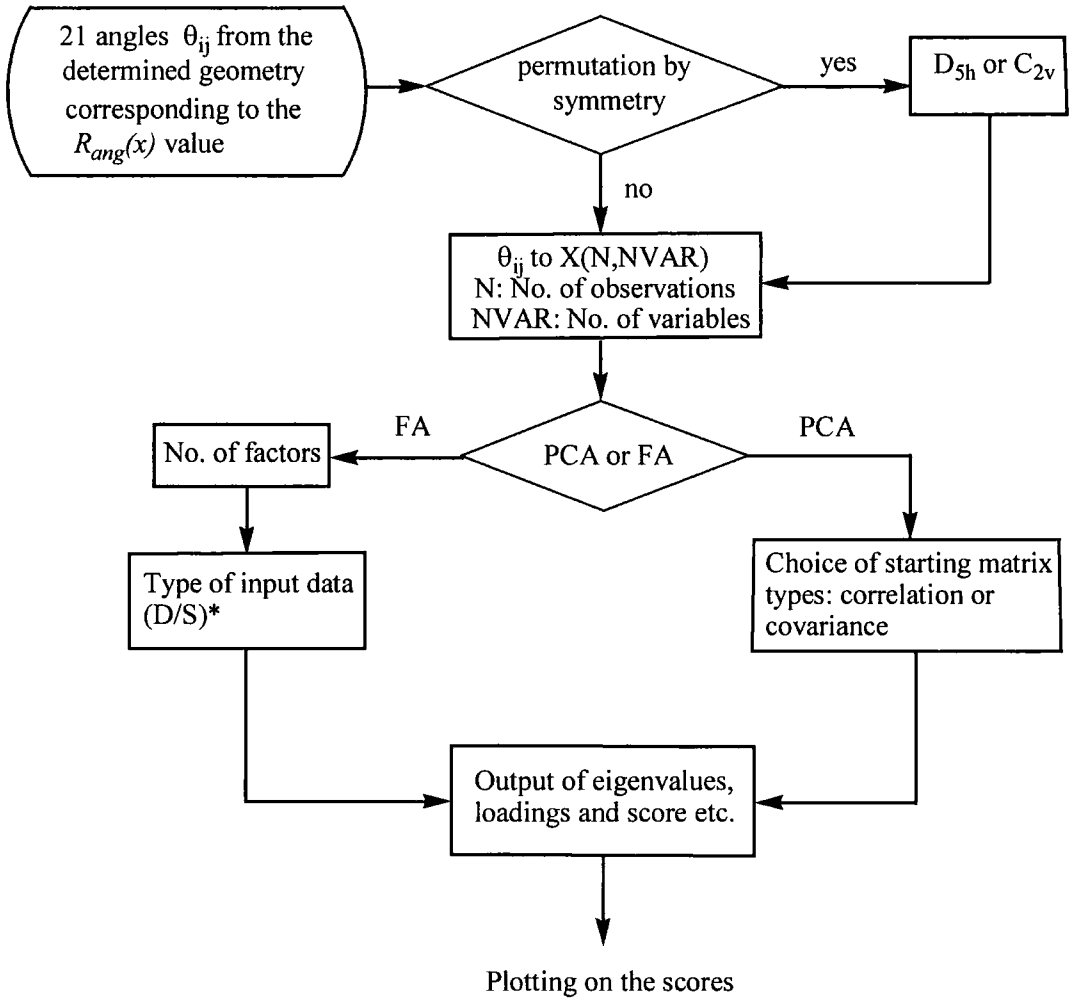


Figure 4.13 Flow chart for adding PCA or FA to the program D/S*: Correlation/Covariance matrix as input data.

as VISTA etc.

Thus, any ML_n data set can be prepared for PCA or FA directly with the $R_{ang}(x)$ program by selecting some key words.

D_{5h} symmetry

Firstly, let us look at PCA results on the 'all-unidentate-ligands' dataset in D_{5h} symmetry. PCA on the covariance matrix resulted in the first five PC's accounting for a total of *89.5%* of the sample variance and five further PC's raise the total of *99.0%* of

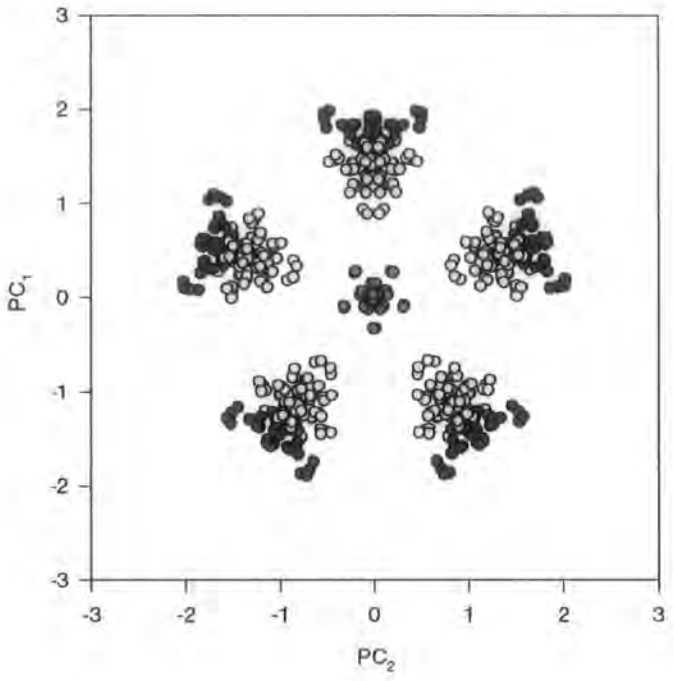
the sample variance. But the major variances appear in the first five PC's. It is obvious that the dimensionality is reduced compared with the *21* valence angles and *18*-dimensional symmetry coordinates systems. Table 4.6 lists the variance accounted for by the first ten PC's.

Table 4.6 PCA of all unidentate ligands ML_7 dataset. Variance (%) accounted for by the first ten PC's

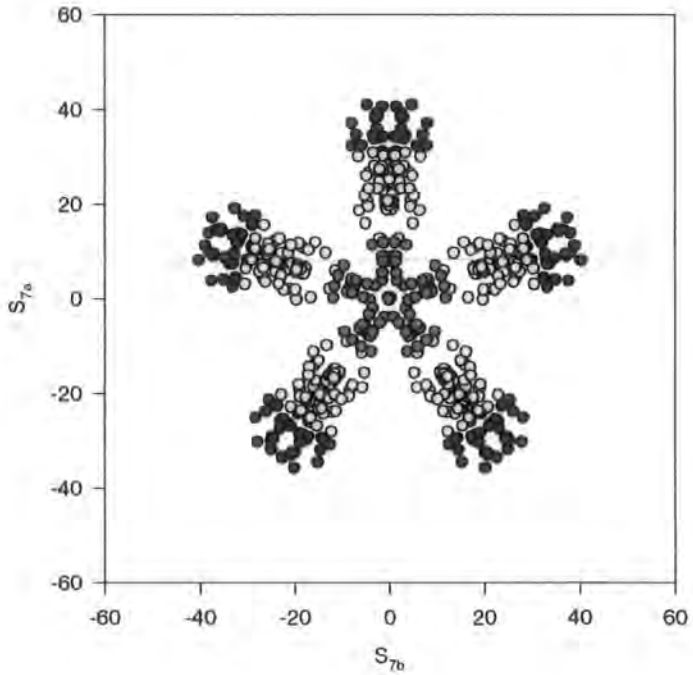
PC's	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7	PC_8	PC_9	PC_{10}
Variance(%)	22.7	22.7	18.3	18.3	7.5	3.6	3.6	0.7	0.7	0.5
Cumul.var(%)	22.7	45.4	63.7	82.0	89.5	93.1	96.7	97.4	98.1	98.6

The first four PC's occur as two degenerate pairs, which have 22.7% and 18.3% angular variances, respectively. From the fifth PC, the variance becomes small. This means that major geometrical effects would be associated with these five PC's. The PC-scattergrams are shown in Figure 4.14(a,c), also accompanied by some symmetry-coordinate scattergrams (b,d). High correlations are observed in these two coordinate systems. PC_1 and PC_2 are related to S_{7a}, S_{7b}, and PC_3, PC_4 are equivalent to S_{11a} and S_{11b}, which belong to representations E_1' and E_2'' of point group D_{5h}. In PC-plots, three geometrical clusters are coloured as red (PBP), yellow (COC) and blue (CTP). PC_1-PC_2 and S_{7a}-S_{7b} plots, are both effectively views along the 1-2 axial-axial vector, PBP appears as the central density at 0.0, 0.0. CTP and COC are the 5-fold growth out of the central density.

The symmetry of the space is *5m* for the degenerate pair of symmetry coordinates $S_{7a}(E_1')$ and $S_{7b}(E_1')$. $\underline{2}$ axis is chosen here as one along r_3 and perpendicular to the $\underline{5}$ axis to preserve the C_{2v} cokernel of these degenerate representations. As was shown in Figure 4.12, the distortions represented by S_{7a} and S_{7b} have the same effects as the transformation from PBP to CTP via COC given in the Figure 3.9. So by deforming the ideal D_{5h} along S_{7a} or PC_1, finally the C_{2v} cokernel configuration becomes dominant. Using the symmetry expanded dataset, the PBP⇒COC ⇒CTP pathway is mapped again in these scatterplots and it corresponds to each of the equatorial atoms being designated as atom 3.

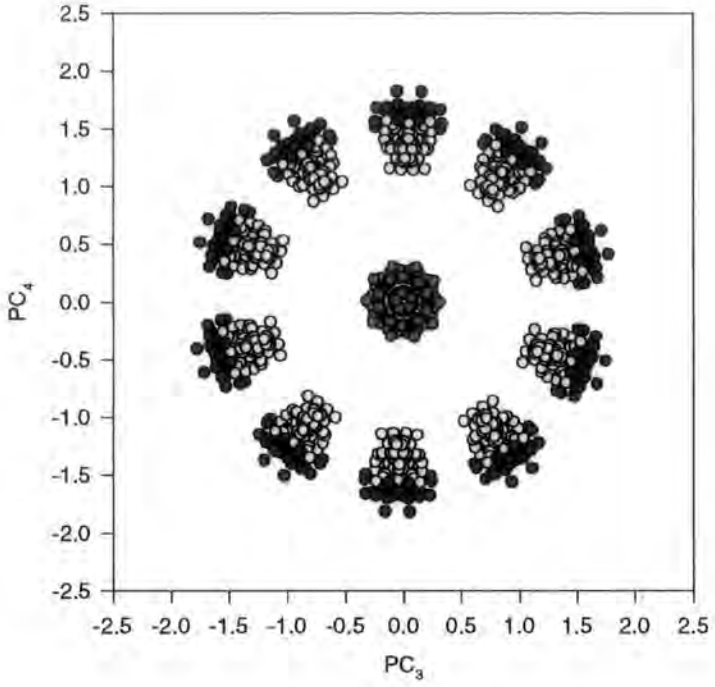


(a)

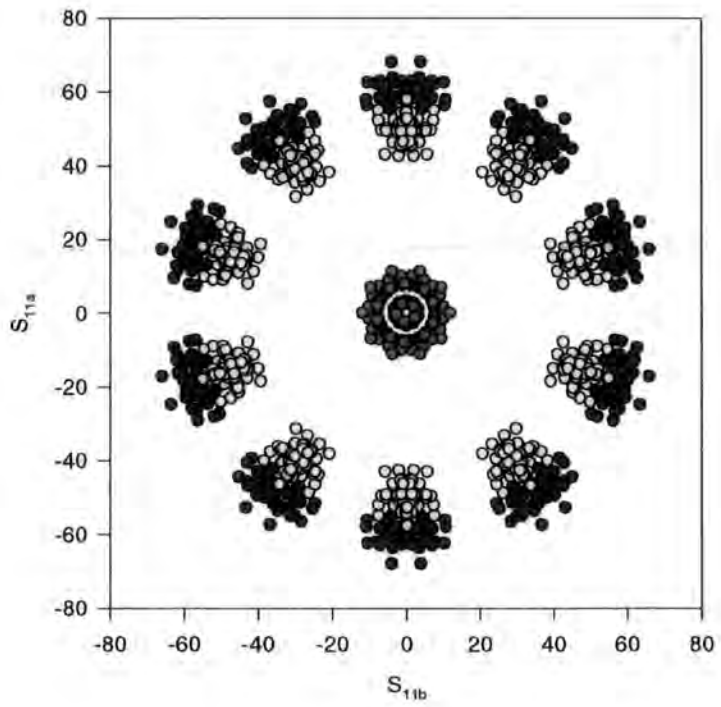


(b)

Figure 4.14 Scatterplots of (a) PC (b) symmetry coordinates for all unidentate ligand complexes. Red: PBP; yellow: COC; blue: CTP.



(c)



(d)

Figure 4.14 *contd.* (c) PC_3-PC_4 (d) S_{11a}-S_{11b}. Red: PBP; yellow: COC; blue: CTP.

The pair of symmetry coordinates $S_{11a}(E_2'')$ and $S_{11b}(E_2'')$ has the symmetry of the space of *10m* (Murry-Rust and Bürgi,1979). The cokernel configuration for this pair is $C_s(\sigma_v)$. They have similar movements to those represented by S_{7a} and S_{7b} for the equatorial atoms but with lower symmetry. The average effect for the distortions on the axial atoms 1 and 2 are cancelled out by the distortions from different directions. The scatterplot of these two coordinates and the corresponding PC plot (PC_3 *vs.* PC_4) shows that *10* symmetry related clusters (CTP/COC) surround the central density cluster (PBP). This is also regarded as a view along the axial-axial vector.

PC_5 accounts for a smaller variance (~7%) but the scatterplot with PC_3 or PC_4 is interesting. There is no corresponding symmetry coordinates to PC_5. It is an axis orthogonal to the PC_3-PC_4 projection. From the PC loadings (Figure 4.15), it can be seen that the major contribution to this PC is θ_{12}, which is an angular value which distinguishes between D_{5h} and C_{2v} geometries. Again, the transformation for the PBP⇒ COC ⇒CTP can be observed along the PC_5 axis. With the angle θ_{12} far away from the ideal value (180°) in D_{5h}, the points go up to the top (C_{2v}), and which are also those of the outer circle in the PC_3-PC_4 plot [Figure 3.14(e)].

The PCA results give the angle loadings on each PC. Figure 4.15 shows the loadings on PC_1-PC_5. Compared with the expressions of symmetry coordinates given above, the correlations between the PC's and symmetry coordinates can easily be seen. The contributions of PC_3 and PC_4 all come from the angles above and below the

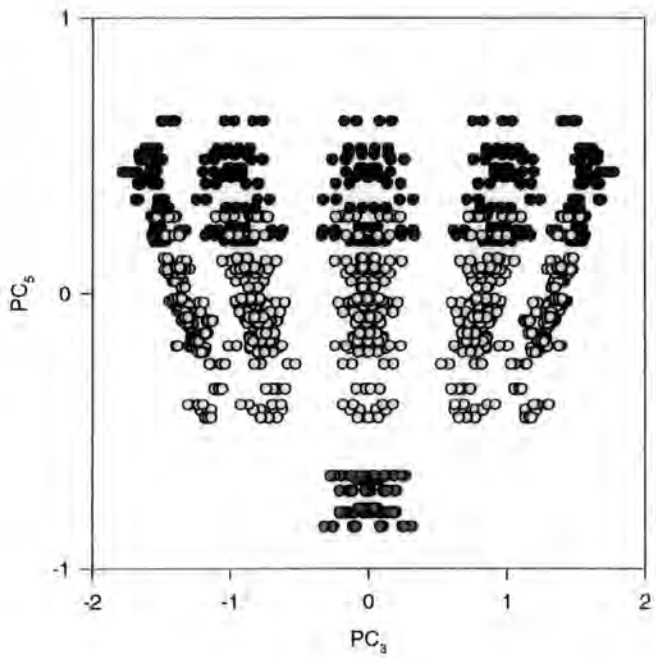


Figure 4.14 *contd.* (e) PC_5 *vs.* PC_3. Red: PBP; yellow: COC; Blue: CTP.

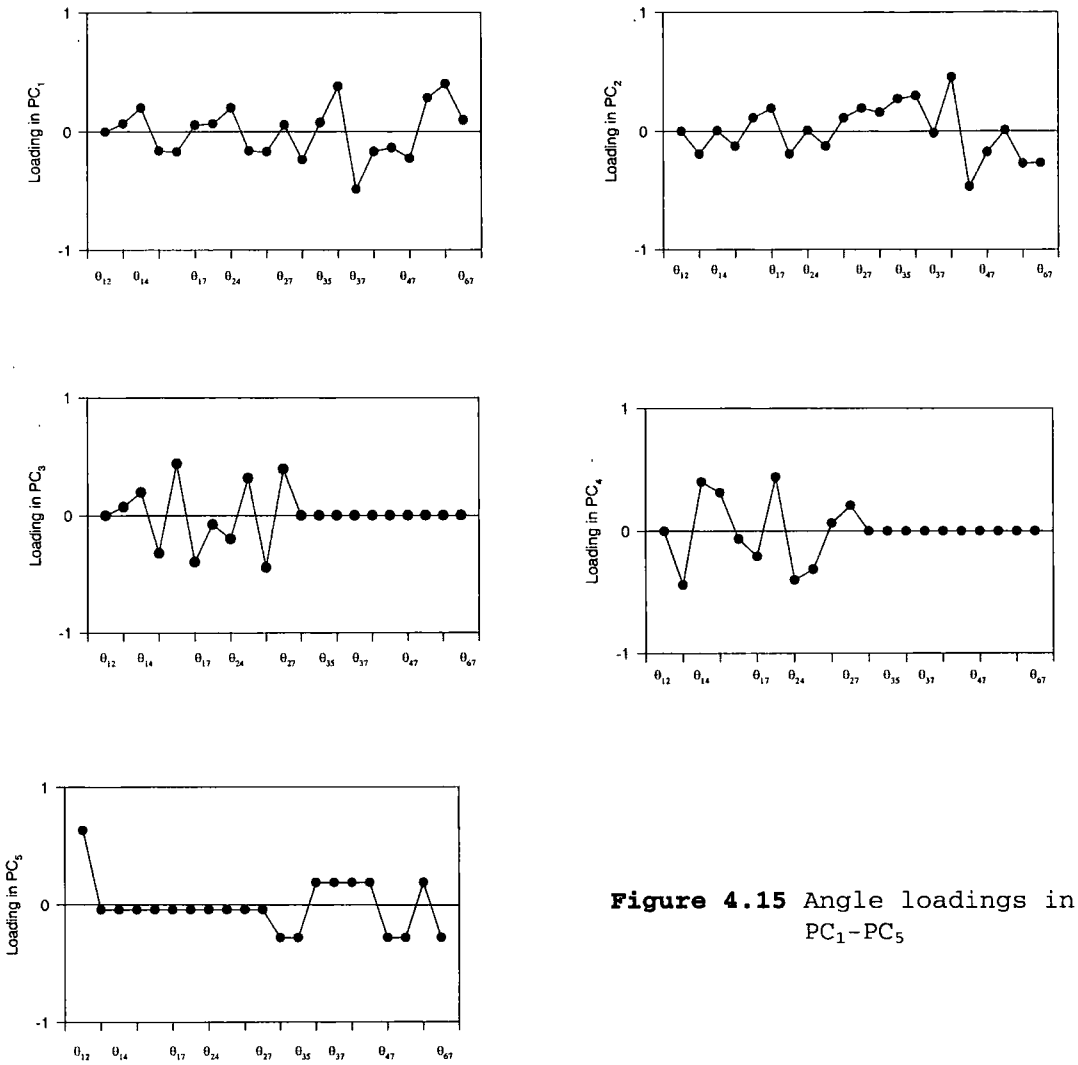
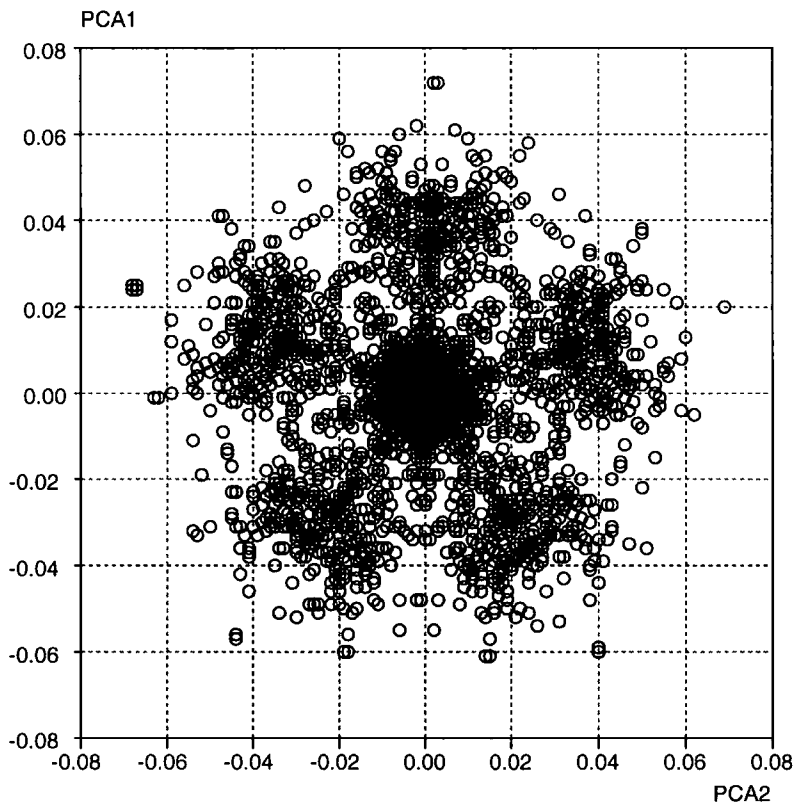


Figure 4.15 Angle loadings in PC_1-PC_5

equatorial plane (i.e. θ_{1j} or θ_{2j}, j = *3-7*). PC_4 is negatively correlated to S_{11a} and PC_3 to S_{11b}. PC_1 and PC_2 are contributed to by almost all the angles. Symmetry coordinates S_7 are related to angles θ_{1j} and θ_{2j} (j = *3-7*, the angles above and below the equatorial plane) while S_6 are only related to the angles in the equatorial plane, θ_{ij}, (i,j = *3-7*). Therefore, PC_1 and PC_2 can be seen as linear combinations of S_6 and S_7.

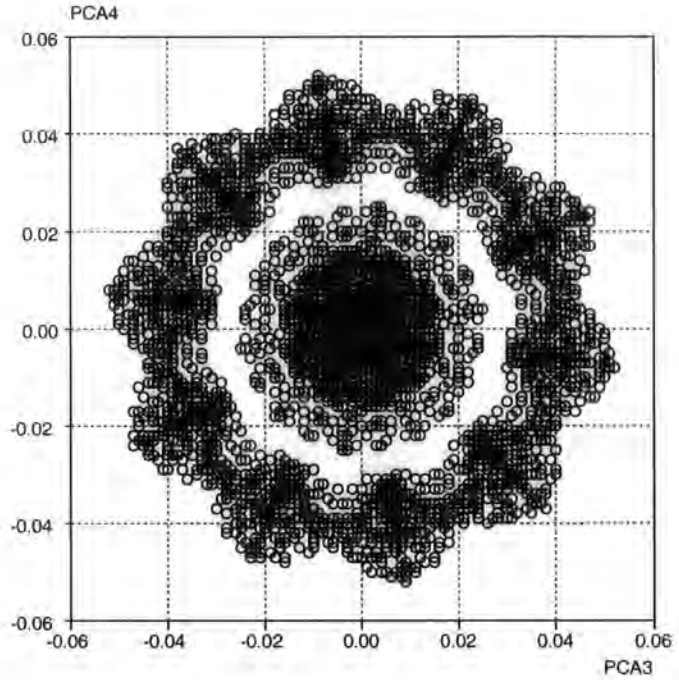
PCA was also performed on the larger dataset of 372 fragments involving all kinds of 7-coordination complexes. The same results as those presented above for all unidentate ligands complexes are obtained. The major difference is that because of complexes with chelate ligands, more distortion or intermediate points can be observed. Figure 4.16 shows these PC-scattergrams.

As a result, a few PC scatterplots involving the five PCs are sufficient to represent the major geometrical clusters and transformation between these geometries visually in angle space. Although the symmetry coordinates can yield similar results, it is complicated to derive these coordinates for high coordination numbers and in this case, 18-dimensional coordinates need considering. Some of them in fact do not provide any useful information in this systematic analysis.

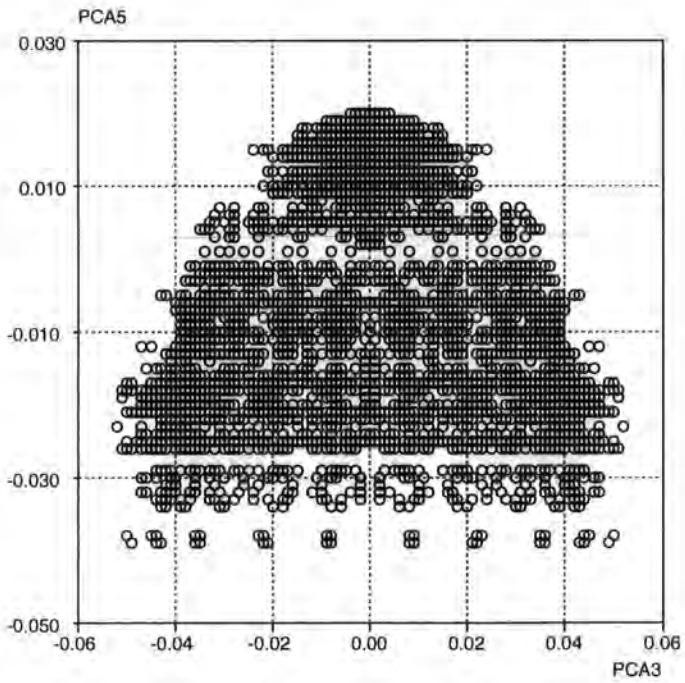


(a)

Figure 4.16 PC scatterplots for all 7-coordinated complexes (a) PC_1 *vs.* PC_2;



(a)



(c)

Figure 4.16 *contd.* (b) PC_3 *vs.* PC_4; (c) PC_5 *vs.* PC_4

C_{2v} symmetry

A similar analysis was performed in C_{2v} symmetry space to compare the results with those observed in D_{5h} symmetry space. This usually reveals how one ideal starting symmetry is distorted towards the other. In this study, attempts were made to fit observed geometries approaching C_{2v} symmetry into a D_{5h} framework (shown above) or *vice versa* (see below).

PCA results in C_{2v} symmetry on the dataset for 7-coordination shows that the first PC, PC_1, accounts for a largest proportion (76.6%) of the total angular variance, and the first ten PC's together account for 99% (Table 4.7).

Table 4.7 PCA of all unidentate ligands ML_7 dataset. Variance (%) accounted for by the first ten PC's (in C_{2v})

PC's	PC_1	PC2	PC_3	PC_4	PC_5	PC_6	PC_7	PC_8	PC_9	PC_{10}
Variance(%)	76.6	8.0	4.7	3.4	2.1	1.2	1.3	0.7	0.6	0.5
Cumul.var(%)	76.6	84.6	89.3	92.7	94.8	96.0	97.3	98.0	98.6	99.1
Symmetry	B_2	A_1	A_2	B_2	A_1	-*	-	-	-	-

* Symmetry is only given to the PC_1-PC_5

It is obvious that dimensionality is largely reduced here. The first few PC's will be used as the axes to express the geometrical preferences, deformations and interconversion pathways between different polyhedra. From the loadings of the PCA results, and comparing these with the symmetry coordinate expressions in point group C_{2v}, PC_1 is related to the coordinates completely belonging to the representation B_2 and the largest coefficient appears as 0.645 for S_{16}. It can be expressed as:

$$PC_1 = 0.645S_{16} + 0.524S_{15}^{'} - 0.507S_{17}^{'} + 0.167S_{17} - 0.139S_{18}$$

where $S_{15}^{'}$ and $S_{17}^{'}$ are two redundant symmetry coordinates in B_2.

The symmetry for the other PCs are obtained in the same way. The large set contributions to PC_2 are S_4 and S_6; S_8 and other redundant coordinates $S_9^{'}$ in A_2 have

large coefficients in PC_3. PC_4 is dominated by S_{18} and PC_5 is mainly contributed to by some other redundant coordinates in the A_1 representation. They can be written as:

$$PC_2 = 0.644S_6 - 0.673S_4 - 0.225S_5 - 0.276S_5' - 0.266S_6'$$

$$PC_3 = 0.757S_9' - 0.746S_8 - 0.385S_{10} - 0.0246S_9$$

$$PC_4 = 0.824S_{18} - 0.317S_{16} - 0.396S_{17}' - 0.185S_{17}$$

$$PC_5 = 0.554S_5' - 0.552S_3' + 0.420S_6 + 0.399S_6' - 0.185S_6'' - 0.126S_5 - 0.032S_1'$$

All the redundant symmetry coordinates included in these expressions are:

$$S_1'(A_1) = \frac{1}{\sqrt{3}}(\theta_{12} + \theta_{13} + \theta_{23}) \qquad S_3'(A_1) = \frac{1}{2}(\theta_{14} + \theta_{15} + \theta_{36} + \theta_{37})$$

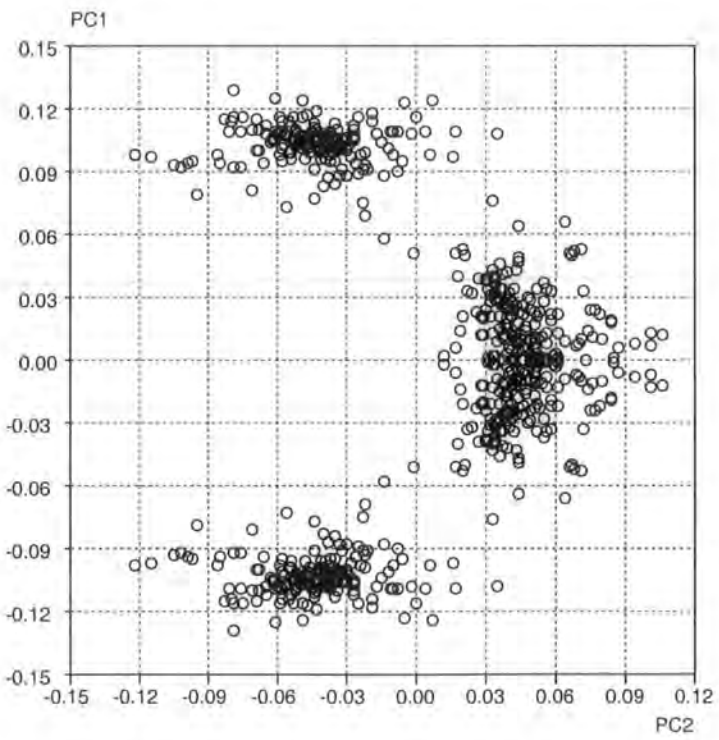
$$S_5'(A_1) = \frac{1}{2}(\theta_{24} + \theta_{25} + \theta_{26} + \theta_{27}) \qquad S_6'(A_1) = \frac{1}{\sqrt{2}}(\theta_{45} + \theta_{67})$$

$$S_6''(A_1) = \frac{1}{\sqrt{2}}(\theta_{47} + \theta_{56}). \qquad S_9'(A_2) = \frac{1}{2}(\theta_{14} - \theta_{15} + \theta_{36} - \theta_{37})$$

$$S_{15}'(B_2) = \frac{1}{\sqrt{2}}(\theta_{45} - \theta_{67}) \qquad S_{17}'(B_2) = \frac{1}{2}(\theta_{14} + \theta_{15} - \theta_{36} - \theta_{37})$$

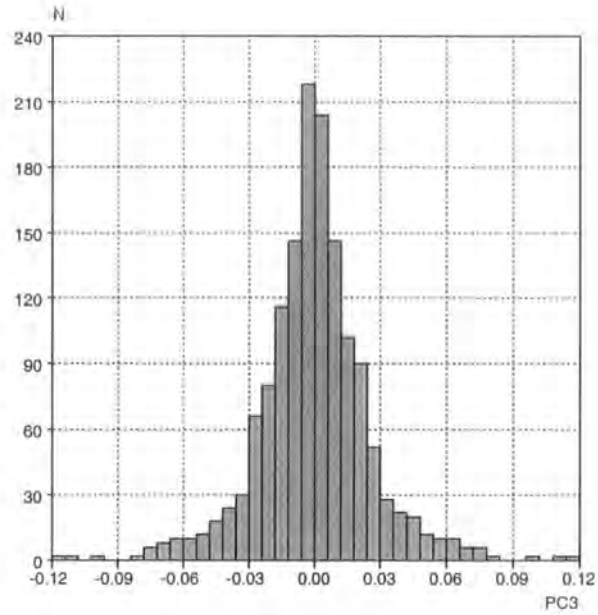
which are considered so, because they are either dependent on those coordinates given in the the last section, such as, $S_6', S_6'', S_9', S_{15}', S_{17}'$, or the positive and negative deviations from the standard values in the expressions will cancel each other for infinitesimal displacements, such as S_1', S_3' and S_5'. The deformations represented by these coordinates can also be found from Figure 4.11. All these linear combinations of symmetry coordinates will give meaningful interpretations to PCs. For example, PC_1 can be seen as a major equivalent to S_{16}, which reflects the distortions of the atoms at positions *1, 2* and *3,* accompanying with the distortions between these apical atoms and equatorial atoms represented by S_{17}' and the angles relocated, θ_{45} and θ_{67}, represented by S_{15}'. These movements are also a tendency to a reverse of the transformation from PBP to CTP form given in Figure 3.9. While PC_2 gives the difference in angular distortions between the equatorial atoms (S_6) and apical atoms (S_4). Figure 4.17 (a) is a

scatterplot of PC_1 against PC_2. It can be seen that PC_1 shows a large difference for different symmetry isomers, especially for PBP complexes and PC_2 gives separate clusters for different geometries. Thus, the scatterplot displays a U shape. The bottom area arises from CTP complexes and the COC form spreads out along the curve. The PBP complexes are located at both ends of the open side, which are related by a mirror plane. Since all the observed geometries are fitted into C_{2v} symmetry in this case, which is a lower symmetry compared with D_{5h}, fewer less points and lower symmetry appeared on the PC-plots. The coordinates (in A_2) in PC_3 cancell out by each other, so that all the values appear close to zero. A histogram of this component is shown in Figure 4.17(b). This will not identify different geometries but only represent deviations from any ideal polyhedra. A scatterplot of PC's with PC_1 [Figure 4.17(c)] gives another clear identification of clusters for PBP (two sides along PC_1) and CTP/COC in the centre.

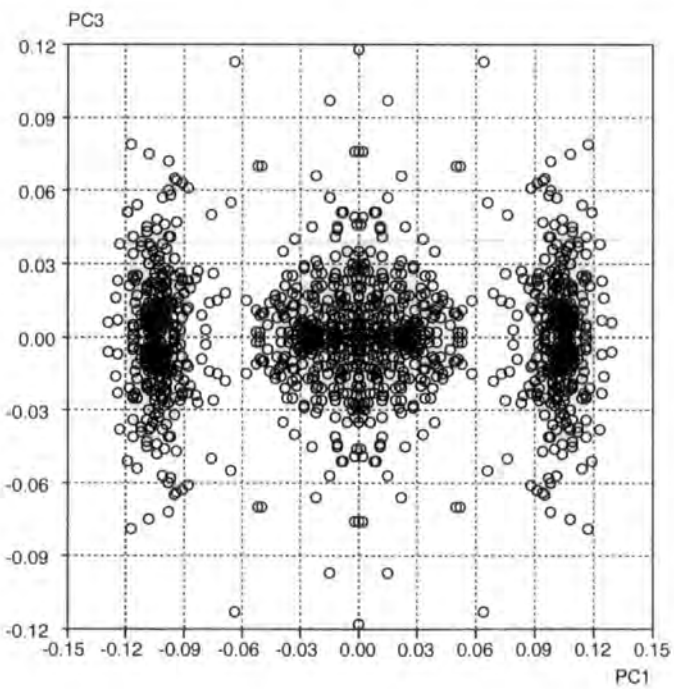


(a)

Figure 4.17 scatterplots of PCs in C_{2v} symmetry(a) PC_1 vs. PC_2



(b)



(c)

Figure 4.17 *Contd.* (b) Histogram of PC_3
(c) PC_3 vs. PC_1

Factor Analysis

Factor analysis is also applied to the same dataset. The results show no significant differences from those using PCA. The importance is the determination of the number of factors (m) that are suitable for this data set. In fact, m is often unknown and it is not easy to select the "correct" value of m (Chatfield and Collins, 1980). Different values have been tried sequentially. It is found that $m = 5$ in D_{5h} symmetry and $m = 10$ in C_{2v} symmetry provide satisfactory results. In PCA, the components are derived as unique and not varied as the number of components included in the analysis. Unlike PCA, factors may change completely as m changes.

The correlation matrix of the data set of *21* basic angles works with FA in this study, and the maximum-likelihood estimation (Basilevsky, 1994) is used to estimate the factor loadings. This provides an advantage in that the scaling problem is free in a mathematical sense, while scaling problems may appear in principal component analysis because the principal components of a set of variables depend critically upon the scales used to measure the variables.

Although it is usually considered that the factors in an m-factor model will not have loadings similar to the component correlations of the first m principal components, there is good agreement in the analysis of this dataset. Figure 4.18 gives the factor loadings of *21* angles in the first five factors in D_{5h} symmetry. For comparison with the results for PCA loadings shown in Figure 4.15, the PC loadings are also plotted with the relative FAC loadings. They show very similar features for FAC_3 and PC_4, FAC_4 and PC_3, FAC_5 and PC_5, except for larger absolute values of the coefficients for most angles in the FACs. Larger differences can be observed in the pairs FAC_2 and PC_1, and FAC_2 and PC_1. These two factors have changed the sign of angles in the loadings and enhanced the effects of some angles, such as θ_{45}, θ_{46} in FAC_1 and θ_{36}, θ_{37} in FAC_2.

Plotting the data in the same way as the PCA results, Figure 4.19 illustrates the scattergrams of pairs of factors for all 7-coordination complexes in two symmetries D_{5h} [(a) and (b)] and C_{2v} [Figure 4.19 (c)-(f)]. They present very similar identifiable clusters for three geometrical forms to those plots generated by PCA. Thus interpretations of these relationships may be obtained in the same manner as for PCA. An interesting feature of the FA results is focused on the scatterplot of FAC_7 against FAC_3 [Figure 4.19(e) and (f)] in C_{2v} symmetry. It not only provides clear clusters for PBP, COC/CTP complexes, as in other plots, but also identifies some special complexes with large

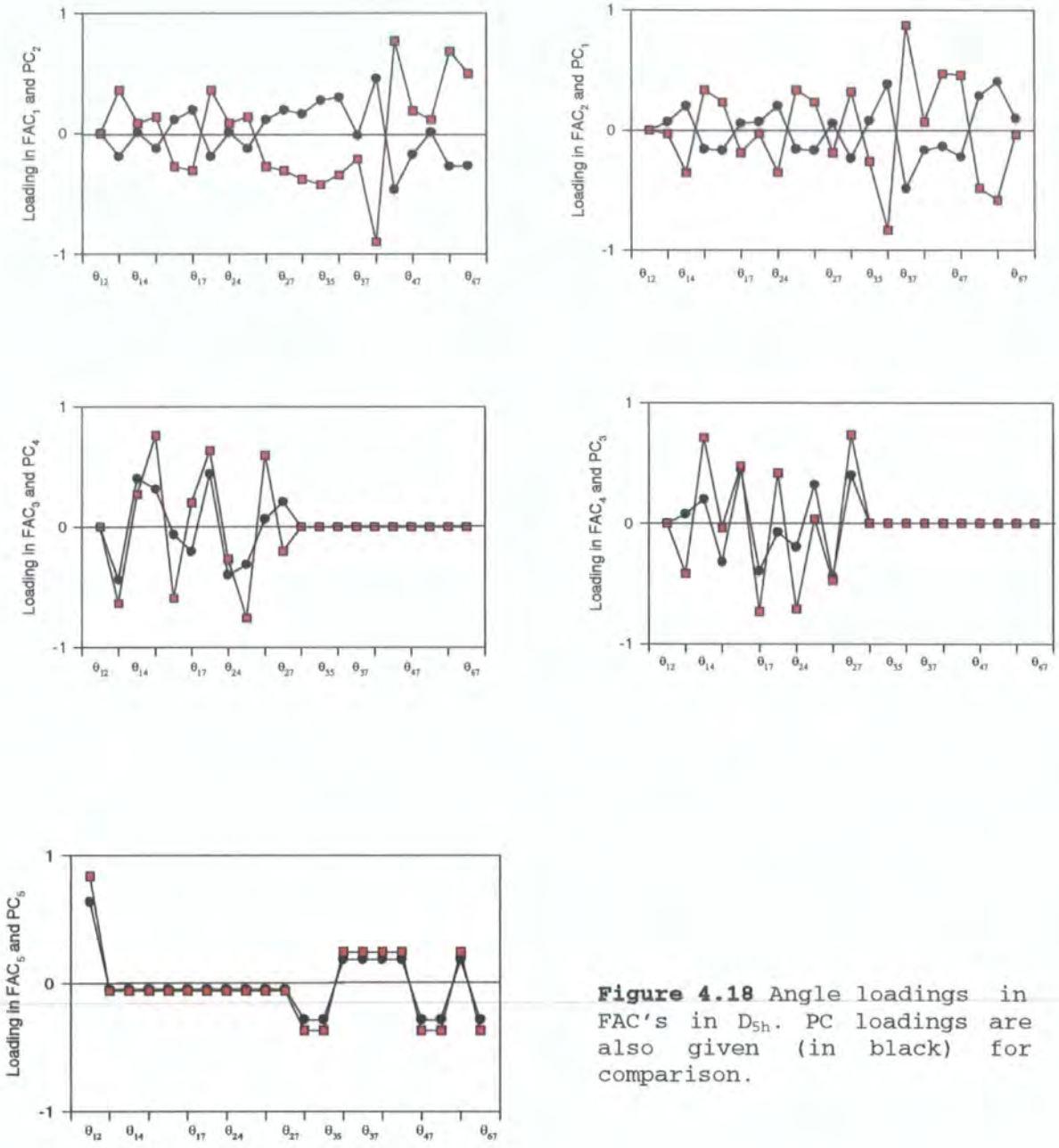


Figure 4.18 Angle loadings in FAC's in D_{5h}. PC loadings are also given (in black) for comparison.

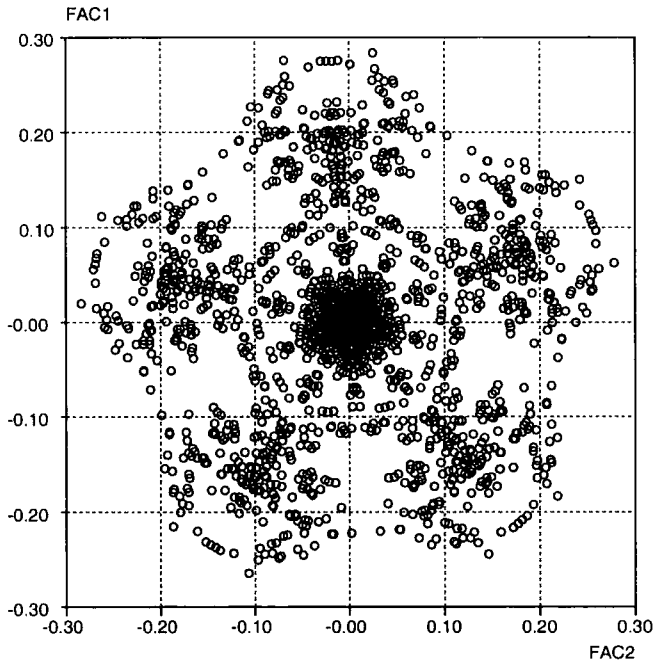
chelation effects. In this plot, a U shape curve shows that CTP geometry is located at the bottom with FAC_3 at ~0.0, and along the edge of the curve, the deformation goes through COC and finally up to PBP at both ends of the U-shape. Some extreme complexes with very tight bite angle, for CTP complexes, appear at the top of the open side along the FAC_7 axis (0.0, 0.48); for PBP complexes with this bidentate ligand in the equatorial plane, they are located at the top ends of both sides (~0.12, 0.44; -0.12, 0.44)

of the U shape as indicated in the Schematic plot at the right of the scatterplot (e). The distribution of FAC_3 in Figure 4.19(f) shows that the geometry for the CTP and PBP has been clustered by this factor.

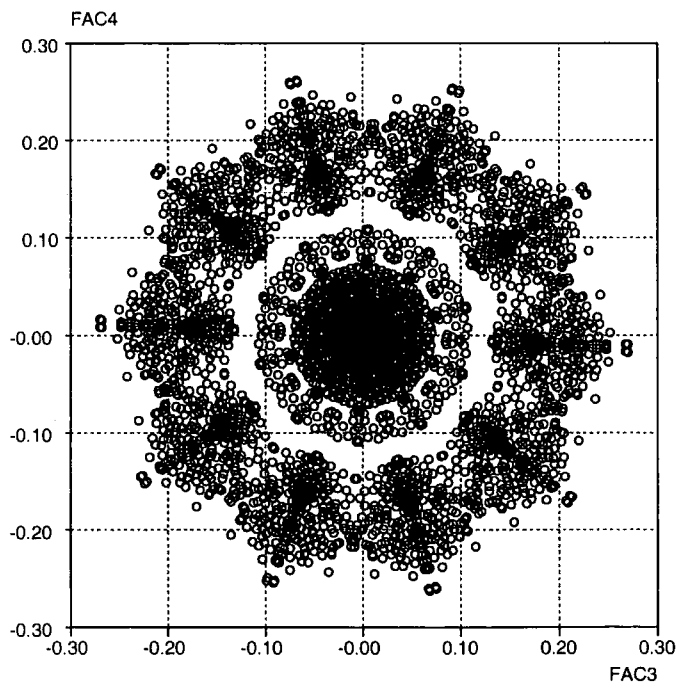
The similarity of PCA and FA for this dataset can be quantified by a calculating of the correlation coefficients between the scores derived by these two techniques. Table 4.8 lists correlation coefficients between PCA and FA scores for all 7-coordination complexes in D_{5h} symmetry. As already seen in PC and FA loadings, FAC_1 is highly negatively correlated to PC_2 and FAC_2 to PC_1. If an angle is taken as a measure of the correlation, it can be seen that is equivalent to -17.3° between the two axes. This means that there is a small axial rotation (-17.3° each) of the FAC_1-FAC_2 plot with respect to the PC_1-PC_2 plot. FAC_3 and FAC_4 are both correlated to PC_3 and PC_4, respectively. Therefore, these two pairs of axes may exchange with each other. It is obvious that the FAC plot for this pair has a larger axial rotation with respect to the PC-plot when compared with the first two factors. But since in the PC_3-PC_4 and FAC_3-FAC_4 plots, a higher symmetry is presented, this rotation of the axes is not as clearly observed as in the PC_1-PC_2 and FAC_1-FAC_2 plots.

Table 4.8 Correlation coefficients between FA and PCA scores for all 7-coordination complexes in D_{5h} symmetry

	PC_1	PC_2	PC_3	PC_4	PC_5
FAC_1	0.265	-0.955	0.000	0.000	0.000
FAC_2	-0.955	-0.266	0.000	0.000	0.000
FAC_3	0.000	0.000	-0.680	0.732	0.000
FAC_4	0.000	0.000	0.731	0.680	0.000
FAC_5	0.000	0.000	0.000	0.000	0.950

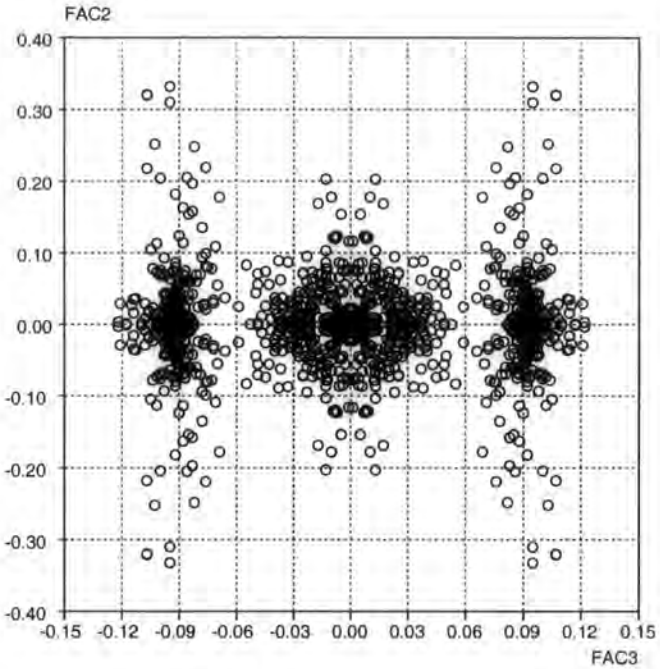


(a)

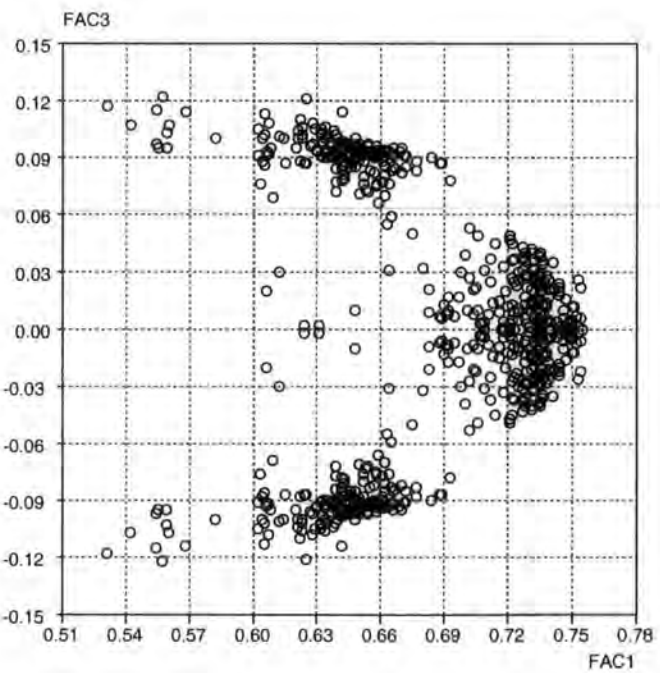


(b)

Figure 4.19 Scattergrams of factors in D_{5h}

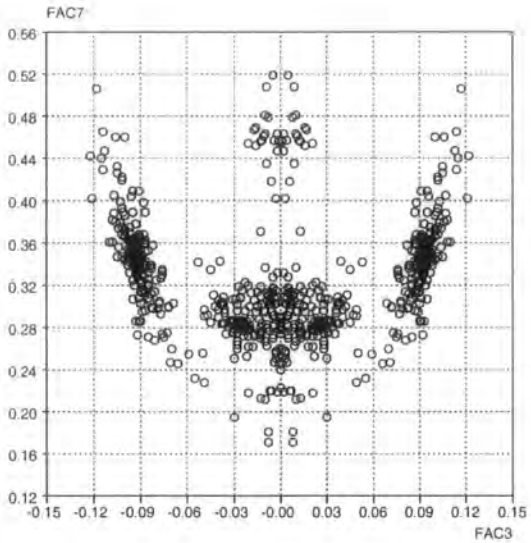


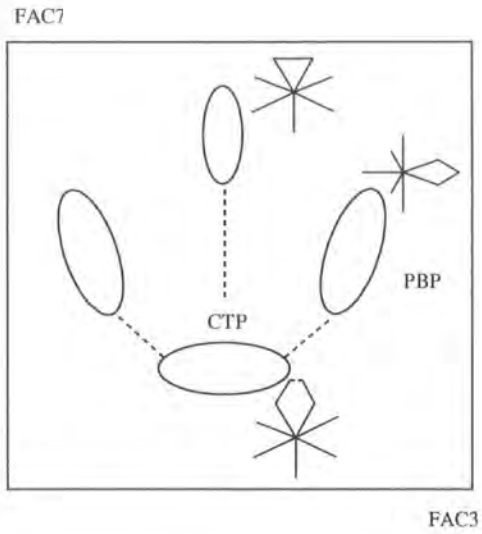
(c)



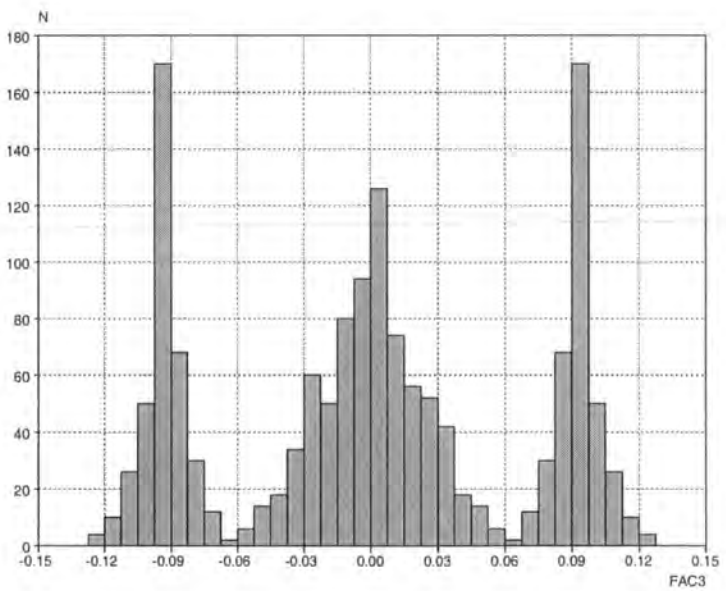
(d)

Figure 4.19 *Contd.* (c), (d) in C_{2v} symmetry





(e)



(f)

Figure 4.19 *Contd.* (e),(f) in C_{2v} symmetry

4.3 Eight-Coordination

4.3.1 Geometry of Coordination Sphere by $R_{ang}(x)$ Calculations

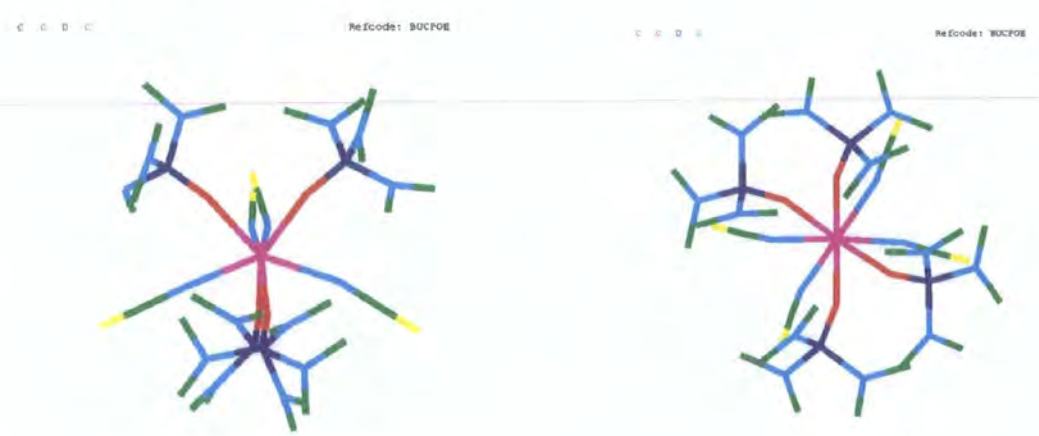
With an increase of coordination number, the available complexes are more commonly found for the lanthanide metals. This coordination number can also be found for some main group metals, e.g. Pb[*3*-(*2*-pyridyl)pyrazole]$_4$[$(MeO)_2PO_2]_2$·H2O (Psillakis *et al*, 1997) and In$(COCPhCO)_3$(4-Me-pyridine)$_2$·$4H_2O$ (Andras *et al*, 1993). In transition metals complexes, like $K_4[Zr(acac)_4]\cdot H_2O$ (Kojic-Prodic *et al*, 1978), $Ta[S_2CNMe_2]_4Cl\cdot CH_2Cl_2$ (Lewis & Fay, 1976) etc, they are mostly restricted to compounds with bidentate ligands, which are usually β-diketonates (such as, acetylacetonate), XCS_2^- and NO_3^- etc. While in the lanthanide and actinide complexes, more ligand types can be observed. Therefore, to investigate the coordination environments of compounds in wide variety of ligand types, this study is focused on the data set of lanthanide and actinide complexes.

The results from the CSD search revealed *254* eight-coordinated molecular fragments lanthanide and actinide with various ligand types, in which *51* hits and *55* fragments have all unidentate ligands. For the transition metals data set, *146* fragments could be found. The search queries were as for the seven-coordination case except for assigning the coordination number for the metal atom to be exactly eight as a 2D-constraint in the QUEST program.

The $R_{ang}(x)$ calculation were also performed on these data sets. The $R_{ang}(x)$ values for each of three idealised polyhedral forms (CUB, DOD and SQUP) (Figure 3.3) for eight-coordination metal sphere is obtained by the calculation over 40,320 possible ligand permutation for each complex.

Firstly, the results from the $R_{ang}(x)$ values on all unidentate ligand complexes revealed that the geometry of square antiprism is slightly preferred for this form of complex, that is, more complexes have smaller R_{ang} values for the square antiprism. When the metals have eight identical ligands, such as, $M(H_2O)_8$, the dodecahedron seems to be formed more easily. In MA_4B_4 complexes, two different types of ligands (A and B) can locate equally at the vertices of the two tetrahedra of the dodecahedron. It was considered (Kepert, 1978) that the dodecahedron has four sites from one tetrahedron associated with higher repulsive energy than those of the square antiprism, and four sites from the other tetrahedron associated with lower repulsive energy.

Therefore, the arrangement of two different types of ligand on these different sites may stablize the dodecahedron relative to the square antiprism. R_{ang} calculations for the data set give: $UCl_4(NCMe)_4$ [CIHLIO, $R_{ang}(SQUP)$: *9.23%*, $R_{ang}(DOD)$: *4.57%*]; $Th(NCS)_4[O{=}C(NMe_2)_2]_4$ [CNURTH, $R_{ang}(SQUP)$: *9.09%*, $R_{ang}(DOD)$: *3.18%*]. An interesting example in this type of complex is $U(NCS)_4[O\text{-}P(NMe_2)_3]_4$ (BUCPOE). There are *4* independent molecules in the unit cell, two of them were described as the square antiprism geometry and the remaining two as deformed towards the dodecahedron because the latter do not have very good coplanarity of the two square faces in the SQUP (Kepert, Patrick and White, 1983). However, R_{ang} calculations showed that two molecules have geometries close to the dodecahedron. Two pairs of $R_{ang}(SQUP)$ and $R_{ang}(DOD)$ values are: *7.26%, 4.22%* and *6.22%, 4.10%*, respectively. Four NCS ligands locate at positions *1, 3, 5, 7* and four $O\text{-}P(NMe_2)_3$ at *2, 4, 6, 8* as in those examples given above [Figure 4.20(a)]. The other two molecules are considered as intermediate between these two polyhedral conformations. The R_{ang} values are: *6.60%*(SQUP), *6.98%*(DOD) and *7.92%*(SQUP), *7.49%*(DOD), respectively. The structure shows better coplanarity of the two square faces in these two molecules. When the geometry of the coordination sphere is expressed in a square antiprism based on the $R_{ang}(SQUP)$ calculation, it has the form shown in Figure 4.20(b). Since the ligand



(a) (b)

Figure 4.20 Structure BUCPOE in DOD (a) and SQUP (b)

O-P(NMe_2)$_3$ is large compared with NCS, distortions of the dodecahedron towards a square antiprism increase the shortest O...O distances between the two ligands from 2.936Å [$R_{ang}(DOD)$: *4.10%*] to 3.249Å [$R_{ang}(DOD)$: *7.49%*]. This will decrease the repulsion between the ligand atoms.

Geometry obtained by the $R_{ang}(x)$ calculations and the corresponding descriptions are apparently more accurate and reasonable.

When two larger ligands with respect to others are included in a coordination sphere, i.e. the form of MA_6B_2 [such as, *para*-toluenesulfonato-O (O-p-Tos) in $M(H_2O)_6(O\text{-}p\text{-}Tos)_2$ (Ohki, Suzuki, Takeuchi & Ouchi, 1988; Faithfull, Harrowfield, Ogden, Skelton, Third & White, 1992), it has been shown that the square antiprism is more stable, in terms of R_{ang} values, than the dodecahedron because in the SQUP form, the two bulky ligands can sit at the vertices *2, 5* or *4, 7*, respectively. This gives a greater separation for these two ligands relative to the position *5, 7* in the dodecahedron (see Figure 3.3), when this geometry is adopted. Table 4.9 lists $R_{ang}(x)$ values for square antiprism and dodecahedron for this type of complex in the data set. When p-Tos is changed to a larger ligand, e.g. 2-naphthalenesulfonato (Napt), the oxygen atoms can more easily deviate from the planes of the square faces in the square antiprism (Ohki, Suzuki, Nakamura, Shimoi and Ouchi, 1985). Thus, the geometry has a tendency to transform to the dodecahedron. R_{ang} calculations give two similar values compared with those in the p-Tos complexes (see Table 4.9).

Table 4.9 $R_{ang}(X)$ values in MA_4B_2 complexes

$[M(H_2O)_6(O\text{-}p\text{-}Tos)_2](O\text{-}p\text{-}Tos)\cdot H_2O$				$[M(H_2O)_6(Napt)_2]Napt\cdot H_2O$			
Refcode	**M^{3+}**	**R_{ang}(SQUP)%**	**R_{ang}(DOD)%**	**Refcode**	**M^{3+}**	**R_{ang}(SQUP)%**	**R_{ang}(DOD)%**
GAKXUL	Sm	3.69	7.64	DIYMUT	Pr	4.53	6.34
GAKYAS	Gd	3.51	7.71	QQQBNP01	Eu	4.64	6.25
GAKYEW	Dy	3.51	7.66	DIYNEE	Gd	4.56	6.30
GAKYIA	Ho	3.54	7.60	GAKZEX	Dy	4.90	6.44
GAKYOG	Er	3.45	7.72	DIYNII	Er	4.51	6.48
GAKYUM	Yb	3.48	7.66	DIYNOO	Yb	4.64	6.33
SUDDOK	Lu	3.61	7.56				

Tetra-bidentate ligand complexes are the most common ML_8 coordination for the transition metals. General geometry can be found mainly in DOD and SQUP forms. Four bidentate ligands normally span the *1,6*; *2,7*; *3,8* and *4,5* edges in the dodecahedron. There are several types of bidentate ligands in the complexes, which can be classified according to the number of atoms between the two coordinated atoms. The most common ones are: $S_2C\text{-}X$, $^-OC{=}OC{=}OCO^-$, $O{=}CR(CH_2)_nCR'{=}O$ and NO_3^- etc. *4*, *5*, *6*-membered chelate rings are usually formed in these complexes. The contact distances between the two bite ligand atoms directly affect the geometrical conformation of the coordination sphere. It is obvious that the shorter the distance, the more steric the effect. In Kepert's theoretical predictions of geometry for this type of complexes (Kepert, 1978), examination of the minima on the potential energy surfaces corresponding to three different geometries, shows that the relative stabilities of specific structures, depends on the choice of normalised bite *b*, which is defined as the distance between the coordinated atoms of the chelate divided by the metal-ligand atom distance.

At low normalised bite values *b*, the dodecahedron is a dominant geometrical form for the complexes. As the normalised bite value is progressively increased, intermediate geometry that approximates the square antiprism appears, and for the larger normalised bite values it is possible to obtain a D_4 square antiprismatic structure. In the $R_{ang}(x)$ calculations, no distance parameters are used. But the results for the assignments of geometry for each coordination sphere is in general agreement with the theoretical prediction based on the normalised bite *b* in the repulsion energy expression. For the complexes $M(S_2C\text{-}X)_4^{n\pm}$, *4*-membered chelate rings are formed and normally have low normalised bite values and the corresponding $R_{ang}(x)$ calculations project all smaller values as being in the dodecahedral form. Table 4.10 (left part) lists these values for transition metal complexes. Since all these complexes have large R_{ang} for the cubic geometry, $R_{ang}(CUB)$ is not given here. Also, the solvent molecules and other packing fragments in the structures are not included in the Table.

With an increase in the number of atoms making up a chelate ring in the coordination sphere, the normalised bite values in six-coordination also increase greatly (Kepert, 1977). But in 8-coordination, this relationship is not as clear as in 6-coordination (Kepert, 1978). The two geometrical forms for this type of complexe have been given in Figure 4.20. It can be seen that when the square antiprism is formed, the most common conformation is that two bidentate ligands sit on the two opposite edges of the square face, respectively. Therefore, fewer atoms (≤ 4) in the chelate rings,

always give a rectangle instead of a square, with the chelate sides shorter than the open sides. In this case, the dodecahedron is the best description. While with larger chelate rings, the lengths of edges on the square planes become closer to one another, the square antiprism is a dominant geometry. For example, in the complex $V(S_2CPh)_4$ DTBENV10

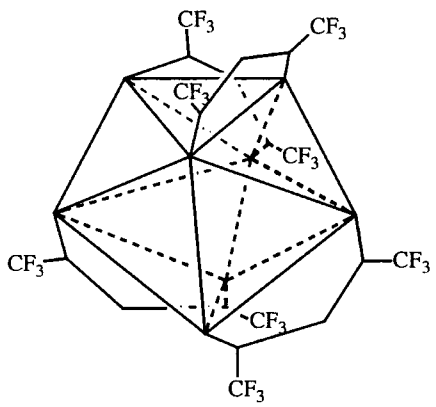
Table 4.10 R_{ang} values in $M[S_2C\text{-}X]_4$ and $M[OCRCHCRO]_4$ tetra-bidentate coordination spheres

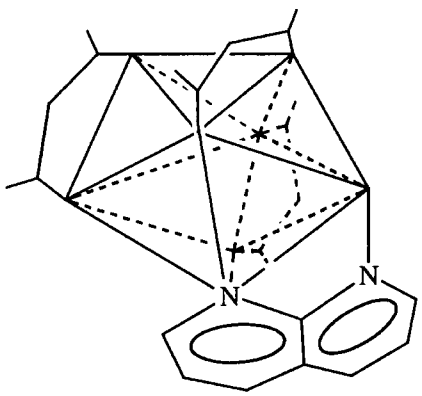
Refcode	Complexes*	R_{ang}(SQUP)%	R_{ang}(DOD)%	Refcode	Complexes*	R_{ang}(SQUP)%	R_{ang}(DOD)%
CIRJOC	$Mo[S_2CC(CN)_2]_4^{3-}$	9.59	7.56	BOHROF	$Zr(SCMeCHCMeO)_4$	4.43	7.06
COWXER	$Mo[S_2CNEt_2]_4^+$	8.86	7.02	BPHXHF01	$Hf(OCPhNPhO)_4$	8.90	4.84
DEBBIV	$Mo[S_2CNEt_2]_4^+$	9.88	6.11	CYSFAC10	$Y(OCCF_3CH_2CCF_3O)_4$	8.00	2.00
DEDRAP	$Nb[S_2CNEt_2]_4^+$	10.7	7.88	FIZLEF	$Nb(OCOCOO)_4^{4-}$	3.59	7.69
DETCMO	$Mo[S_2CNEt_2]_4^+$	7.42	4.45	HYZCSC	$Sc(OCONHNH_2)_4$	5.43	6.61
DETCNB	$W[S_2CNEt_2]_4^+$	10.3	7.41	KOXHFP10	$Hf(OCOCOO)_4^{4-}$	8.91	4.24
DIGBEA	$Mo[S_2CNEt_2]_4^+$	10.8	7.90	NAOXZR	$Zr(OCOCOO)_4^{4-}$	6.85	5.49
DTBENV10	$V[S_2CPh]_4^{4+}$	9.60	3.91	PIVMNB10	$Nb(OCBuCH_2CBuO)_4$	2.97	8.24
DTBZMO	$Mo[S_2C\text{-}Ph]_4$	9.88	5.10	PPRDZR	$Zr(OCPhCH_2CphO)_4$	2.53	7.95
ETCMOC	$Ti[S_2CNEt_2]_4$	10.5	7.34	SROXAM	$Zr(OCOCOO)_4^{4-}$	8.97	4.33
ETXANW	$W[S_2CSEt]_4$	10.4	6.82	ACACCE01	$Ce(OCMeCH_2CMeO)_4$	3.33	7.30
JAXDOB	$Mo[S_2C\text{-}C_{10}H_7]_4$	7.70	7.21	ACACNP	$Np(OCMeCH_2CMeO)_4$	2.56	8.06
JIKPIC	$Hf[OOCNPr^i_2]_4$	12.1	9.37	ACACTH01	$Th(OCMeCH_2CMeO)_4$	5.83	6.01
JUVTUP	$W[S_2CNMe_2]_4^+$	9.57	6.50	ACACUB	$U(OCMeCH_2CMeO)_4$	3.52	7.68
MTCTAC10	$Ta[S_2CNMe_2]_4^+$	9.57	6.97	GALYOH	$Tb(OCMeCH_2CMeO)_4$	4.29	7.95
TDTAMO	$Mo[S_2CMe]_4$	6.77	8.09	PIEUAC01	$Eu(OCPhCH_2CmeO)_4$	2.74	8.64
VUSBUG	$W[S_2CNEt_2]_4^+$	9.67	6.12	TACT6HB	$Th(OCMeCH_2CMeO)_4$	4.50	7.76
YAWKEM	$Mo[S_2C\text{-}C_5H_4N]_4$	9.31	6.74	TFPBOU	$U(OCPhCHCPhO)_4$	9.61	1.79
VUSDOC	$Re[S_2CNMe_2]_4$	9.87	6.88	VAYYID	$U(OCCF_3CH_2COOEt)_4$	6.15	4.80

* Only coordination spheres are given

(Bonamico *et al*, 1974), when a square antiprism is considered, the smallest difference between the long and short edges is 0.68Å, whereas in $Nb[COC(^tBu)CH_2C(^tBu)O]_4$, PPRDZR (Chun *et al*, 1979), the largest difference is 0.12Å. This is a common feature in this type of complex. With different substitutions on the bidentate ligands and coordination environments, one geometry could be distorted towards the other to form

an intermediate, such as, JAXDOB that has two similar R_{ang} values. Figure 4.21 gives two other forms of the dodecahedron, in which CAMHFA has closer geometry to the DOD. These can be seen as a creasing of the rectangular faces of the square antiprism tending towards the dodecahedron. Both R_{ang} values reflect the degree of the distortion of the square antiprism. However, sometimes they deviate from both geometries. There are some earlier criteria (Hoard *et al*, 1968; Leipoldt *et al*, 1980; Muetterties & Guggenberger, 1974; Corden *et al*, 1970) used to examine the distortion from each of those idealised geometries. Once the ligand labels are assigned by the R_{ang} calculations, it is convenient to apply these methods.





CAMHFA R_{ang}(SQUP) : 7.20%
R_{ang}(DOD) : 3.55%

KUPSUJ R_{ang}(SQUP) : 6.85%
R_{ang}(DOD) : 5.91%

Figure 4.21 Two other forms of the dodecahedron in M(bidentate)$_4$

The cube is a rare case in these complexes. However, there are two instances in this data set. One is NPFORM, $Np(CHOO^-)_4$, tetrakis(Formato-O,O')-neptunium(iv) (Hauck, 1976), [R_{ang}*(CUB): 7.05%*, R_{ang}*(SQUP): 17.6%* and R_{ang}*(DOD): 13.3%]*, the other is TBPYUB, tetrakis-(*2,2'*-Bipyridyl)-uranium (del Piero, *et al*, 1975), [R_{ang}*(CUB): 4.89%*, R_{ang}*(SQUP): 21.2%* and R_{ang}*(DOD): 17.5%*]. The first one is a greater distortion from the cube because of the tighter-bite ligand and the second is nearer to a regular cube. They both have the form shown in Figure 4.22, which is a view

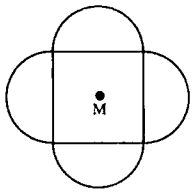


Figure 4.22 The stereochemistry of M(bidentate)$_4$ represented as a cube

from the top of the cube. Two pairs of bidentate ligands span the opposite edges of the upper and lower square planes and are located at different sides. These structures did not fit into any predictions and explanations from the theory based on the normalised bite. However, it is easy to identify them from the relevant R_{ang} values.

The hexagonal bipyramid is another possible geometry in 8-coordination. It mostly can be found in some d^{10} transition metals and main group metals, such as, $HgCl_2$(18-crown-6) (CESZOP) and $CdCl_2$(18-crown-6) (CESZUV, Paige & Richardson, 1984). This polyhedron has the most unique set of angle values between the metal and donor atoms, so it is the easiest form that is distinguishable from the other geometries.

A bicapped trigonal prism (BCTP), is sometimes considered independently, and is, in fact, an intermediate on the interconversion pathway from a dodecahedron to the square antiprism. In most calcium octacoordination complexes, the BCTP geometry can be approached (Klebe, 1994).

The geometry for octacoordination can be classified in more detail according to various types of ligands (Kepert, 1978). In this work, the systematic analysis has also emphasised the relationships and correlation of specific geometrical parameters in PC-plots as those in the systematic geometrical study of seven-coordination. The geometrical forms for all complexes in the whole data set are identified and classified and the interconversion pathways between different polyhedra are expected to be explored from the data located between these main clusters that correspond to the reference polyhedra.

4.3.2 Interconversion Between the DOD and SQUP

The dodecahedron and square antiprism are two major geometrical forms adopted by the transition metals and lanthanide 8-coordination complexes. In practice, these two geometrical forms can be interconverted and no real preference for either form can be

predicted because the difference in repulsive energy between the two forms is small. With diversification of coordination environments, some of the structures have geometries between these two polyhedral forms. These data are therefore useful to investigate the interconversion pathways between the DOD and SQUP.

As in 7-coordination, R_{ang} value calculations are also performed for the model transformation from the DOD to SQUP. Similarly, R_{ang} values are recorded at each point. Thus, the geometrical change can be estimated from the relationship between the R_{ang} values and each percentage point change for one model to the other. Figure 4.23

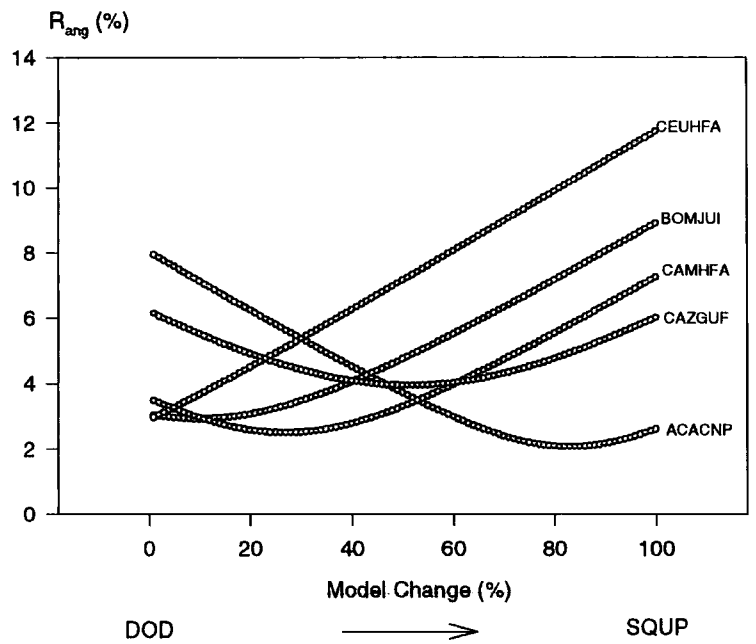


Figure 4.23 $R_{ang}(\%)$ in model change from DOD to SQUP

shows some curves of this type, in which each curve represents the change of R_{ang} values from the DOD to SQUP for each complex. The complex CEUHFA can be seen as a regular DOD the model changes to SQUP, R_{ang} values increase greatly and the smallest value is at the beginning of the change; BOMJUI has some small distortions from the DOD form; While ACACNP has about 20% distortion from the SQUP. The

smallest R_{ang} value for CAZGUF falls in the middle of the model interconversion route, which indicates an intermediate structure.

In a transformation from the DOD to SQUP it is supposed that two edges linking two vertices belonging to different tetrahedra in the DOD are stretched, such as edge *15* and *37* (see Figure 3.3). Figure 4.23 only reflects part of the interconversion pathway because the other pair, edge *17* and *35* could also be stretched. In this case, the model change needs considering in a different direction (Figure 4.24). Some complexes, e.g. FIZLEF, GEBPEI, GINROK and GAKXOF etc., which could not be fitted in Figure 4.23, can be addressed in the other form of the plot, which is the mirror plot of Figure 4.23. The mirror plane is perpendicular to the paper and through the R_{ang} axis.

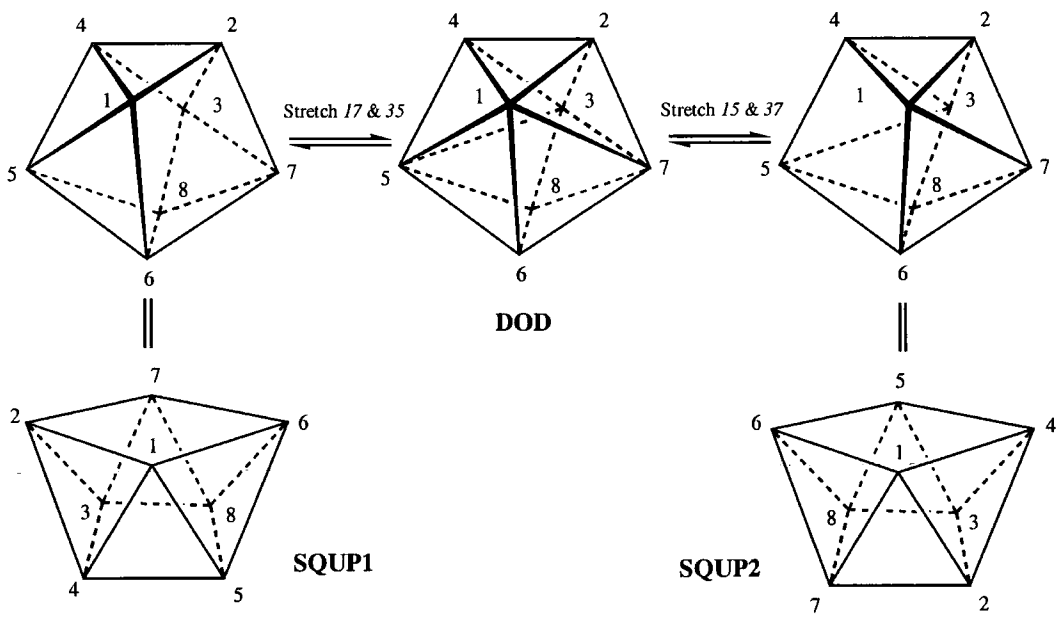


Figure 4.24 Interconversion between DOD and SQUP

In addition, the symmetry coordinates for 8-coordination in DOD (D_{2d}) and SQUP (D_{4d}) symmetry have been derived (see Appendix II*) from *28* internal bond angles, *L-M-L*, respectively.

* 28 symmetry coordinates from 28 bond angles are given, 13 of them should be independent.

In D_{2d} symmetry, a pair of coordinates from representation B_1, S_{10} and S_{11}, represents the distortions of these edges corresponding to the relative angles in two tetrahedra of the DOD form:

$$S_{10} = \frac{1}{\sqrt{2}}(\theta_{15} - \theta_{17} - \theta_{35} + \theta_{37});$$

$$S_{11} = \frac{1}{\sqrt{2}}(\theta_{26} - \theta_{28} - \theta_{46} + \theta_{48}).$$

The correlation of these two coordinates for the transition metal complexes data set is shown in Figure 4.25. All the DOD clusters are located at the centre (red) (0.0,0.0), which are close to two regular tetrahedra. As the angles are distorted towards SQUP, two kinds of conformations of the SQUP appear at both sides (blue) of the centre. This further confirms the interconversion pathway between the DOD and SQUP

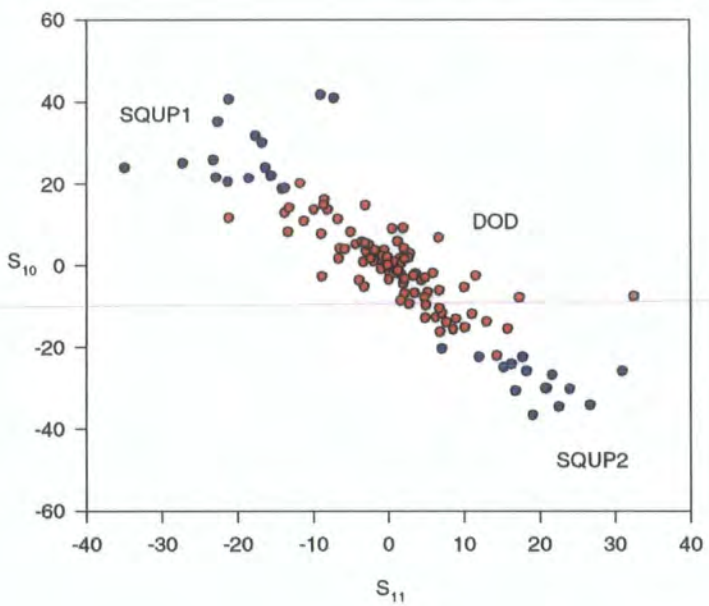


Figure 4.25 S_{10} *vs.* S_{11} in D_{2d} symmetry

given in Figure 4.24. SQUP1 or SQUP2 represents the conformation stretched from the edge *17* and *35* or *15* and *37*, respectively.

Again, R_{ang} value calculations on transition metal and lanthanide and actinide complex data sets of 8-coordination provided a basic identification of geometry of each coordination sphere and based on this geometrical conformation, the labels of ligands are assigned automatically according to the label order given in the reference polyhedra. From these results, further systematic analysis can be carried out.

4.3.3 PCA and FA in 8-Coordination Sphere

Multivariate analysis by PCA and FA was also applied in these data sets to explore the geometrical diversity of the 8-coordination sphere. PCA was performed on the *28* metal-ligand valence angles. Each fragment in the data set was expanded according to two symmetries D_{2d} and D_{4d} for two reference polyhedra DOD and SQUP, respectively. PCA reduces the dimensions: the first ten PCs take account ~90% of total variance in both symmetries and large variances are accounted for by the top five PCs. The PCA results from the lanthanide metal data set are listed in Table 4.11.

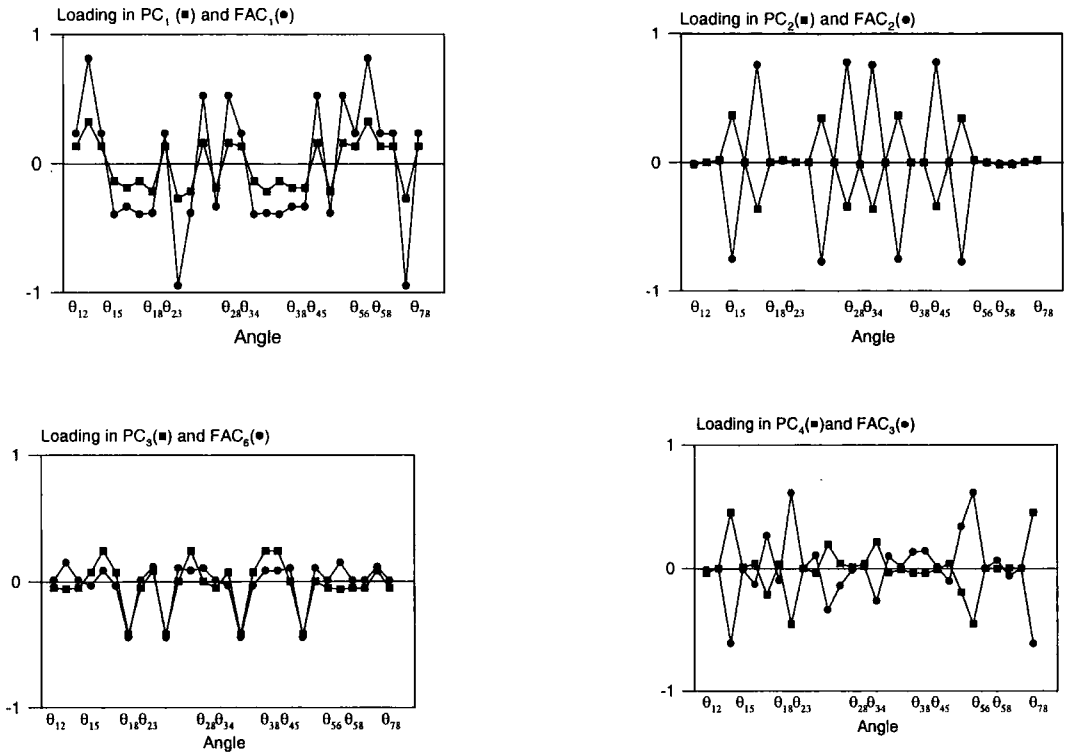
In D_{2d} symmetry, the first two PCs account for 50.9% of total variance in the data set and the third one 9.0%. The remainder account for 40% but each one for only a

Table 4.11 Variance percentage taken account for PC_1-PC_{10}

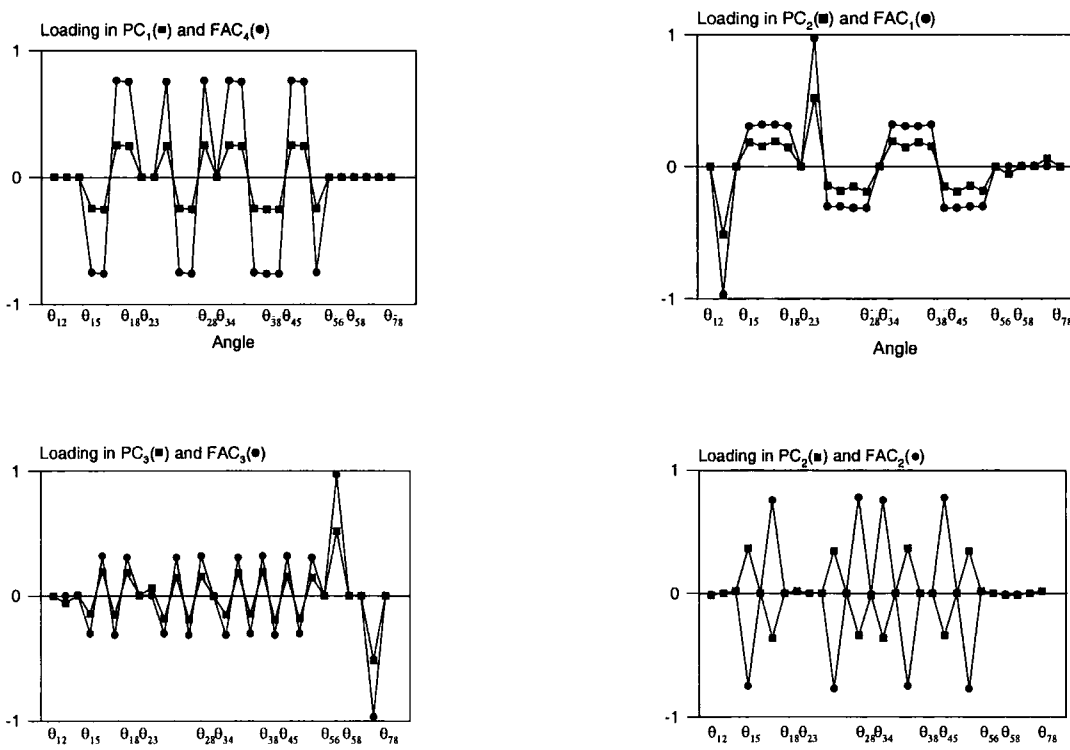
D_{2d}	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7	PC_8	PC_9	PC_{10}
Var.(%)	27.8	23.1	9.0	5.4	5.4	3.6	3.5	2.8	2.8	2.7
Cuml. Var.	27.8	50.9	59.9	65.3	70.7	74.3	77.8	80.6	83.4	86.2
D_{4d}										
Var.(%)	28.2	14.2	13.2	13.2	4.5	4.5	3.6	2.7	2.7	2.7
Cuml. Var.	28.2	42.4	55.6	68.8	73.3	77.8	81.4	84.1	86.8	89.5

small percentage. In D_{4d} symmetry, the top four have larger percentage variances and take account of 68.8% of variance in the total data set, PC_3 and PC_4 are a degenerate pair.

With the appropriate selection of the number of the factors, FA can provide very similar results to PCA. Figure 4.26 gives the angle loadings of the PCs and FACs in D_{2d} (a) and D_{4d} (b) symmetry, respectively. It can be seen that the contributions of the angles



(a)



(b)

Figure 4.26 Angle loadings in PCs and FACs (**a**) D_{2d}, (**b**) D_{4d}

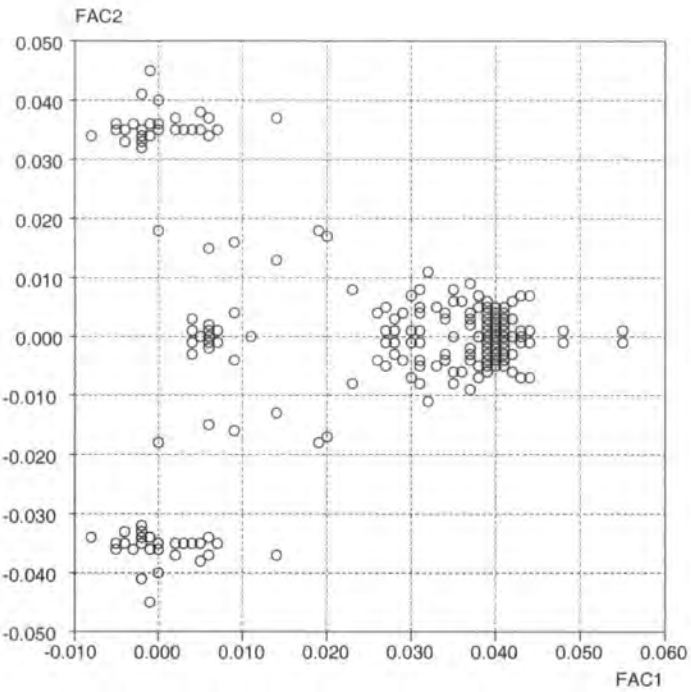
to the derived factors are enhanced in the FA results compared to the PCs.

In D_{2d}, PC_1 is equivalent to FAC_1, PC_2 is negatively correlated to FAC_2, PC_3 to FAC_6 and PC_4 is negatively correlated to FAC_3. Whereas in D_{4d}, equivalent pairs are: PC_1-FAC_4, PC_2-FAC_1, PC_3-FAC_3 and PC_4-FAC_5.

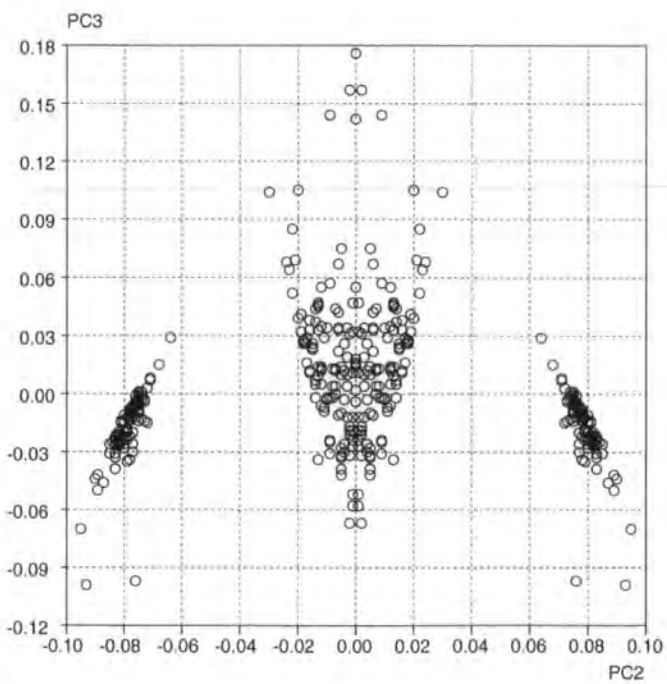
From the given polyhedral models (Table. 3.7), the larger differences of angles between the DOD and SQUP can be found in those angles making up the two tetrahedra in the DOD, e.g. θ_{13}, θ_{57}, and θ_{24}, θ_{68} etc. Whereas those angles between these two tetrahedra have almost no differences, such as, θ_{14}, θ_{67}, and θ_{23}, θ_{58} etc. For the geometry of a dodecahedron, when the ligand atoms at vertices *1, 3* are moved closer (i.e. θ_{13} reduced) and *2, 4* separated (θ_{24} increased), these four atoms tend to move to the same plane and the same feature also exists for vertices *5, 7* and *6, 8*, which describes a tendency to convert from a dodecahedron to a square antiprism.

Firstly, we will look at the results from D_{2d} symmetry. From the angle loadings to FAC_1/PC_1, it can be seen that this factor/PC represents the distortions from the DOD to SQUP given above. FAC_1 is highly correlated to θ_{13} and θ_{57} and negatively to θ_{24} and θ_{68}. FAC_2/ PC_2 also reflects the same movements but the difference is that it represents all other angles of the tetrahedra, i.e. θ_{15}, θ_{17} and θ_{35}, θ_{37}; θ_{26}, θ_{28} and θ_{46}, θ_{48}, which are related to those edges stretched in going from the DOD to SQUP, as discussed in the last section. A scatterplot of these two PCs is illustrated in Figure 4.27(a). Geometries are clearly classified by the clusters. SQUP is located at (0.0, ±0.35), DOD is around (0.04, 0.0). The cube and hexagonal bipyramid structures are also along the FAC_1 axis with FAC_2 at 0.0.

From the angle loadings to the FACs, FAC_1 and FAC_2 have an identical range of values for the linear combinations of relevant angles from each different geometries. A regular SQUP is 0.0 for FAC_1 and DOD is 0.0 for FAC_2. But for PC_3/FAC_6 and PC_4/FAC_3, the corresponding angles are those which are not distinguished in the DOD and SQUP forms. So they do not show sensible correlations with each other. For a geometry close to either of the reference geometries, a value of zero should be expected for these factors. A large value only indicates a more distortion from one of the reference polyhedra. Larger distributions of these two factors for this data set are given in Figure 4.28. No distinguishable values for different geometries can be seen from the distributions. But the scatterplot with PC_2 [Figure 4.27(b)] shows how the molecular structures distort from the idealized geometry classed by FAC_2 along the FAC_3 axis.

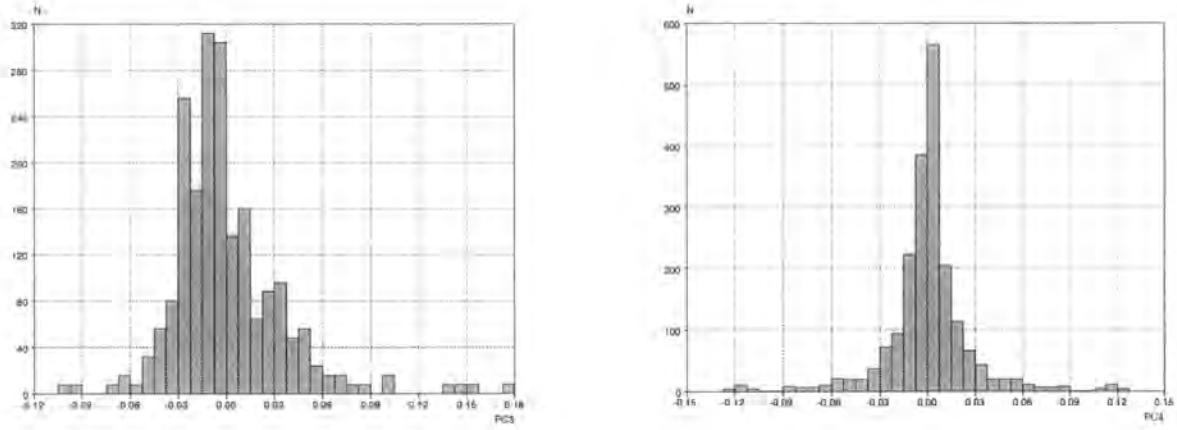


(a) FAC_2 *vs.* FAC_1



(b) PC_3 vs. PC_2

Figure 4.27 Scatterplots of FACs/PCs in D_{2d} symmetry



(a) (b)

Figure 4.28 Distributions of FAC_3 (a) and FAC_6 (b)

In the FA results, another factor FAC_5 is a linear combination of symmetry coordinates S_{13} and S_{16} in representation B_2 of point group D_{2d}. S_{13} reflects the distortion of angles relative to two interpenetrating planar trapezia *5427* and *1386* (see Figure 3.3), which is an alternative and useful way of viewing the dodecahedron. S_{16} corresponds to the movement of atoms at one tetrahedron (*2468*) towards the square planes. The scatterplot of FAC_5 against FAC_2 [Figure 4.27(b)] also illustrates clear clusters for DOD and SQUP structures. The DOD cluster is at around $FAC_5 = 0.0$, while SQUP are placed on two sides with $FAC_5 \approx \pm 0.035$. Distortions from the regular forms are observed along the FAC_1 axis.

More symmetry-expanded points are included in D_{4d} symmetry. Similarly, the clustering patterns observed in D_{2d} symmetry will emerge in more highly symmetric forms, because each fragment is forced to be considered in this higher symmetric form. The real structures that have large deviations from the defined symmetry may also be well separated from those with the symmetry as defined. Again, PCA results gave clustering patterns from the scatterplots of derived components, which identified different geometrical structures. A highly symmetric clustering pattern emerged in the plot of a pair of degenerate PCs, PC_3 and PC_4 [Figure 4.29(a)], which can be seen as a view along the *4*-fold symmetry of the D_{4d} point group. Four well-separated clusters related by $\bar{4}$ symmetry are identified as DOD geometry and the large clusters placed in

the centre of the square cross forward by the SQUP. Figure 4.26(b) shows the *28* internal angle loadings in these components. It is obvious that for an idealised SQUP geometry, the linear combinations of the angles give values of 0.0 for both components, ±0.04 appear alternately for an idealised DOD geometry. All of the observed molecular geometries are aggregated around the centres of these clusters. Four clusters representing the DOD geometry, correspond to different conformations in this geometry. Examination of individual clusters in detail will indicate which cluster represents a relevant isomer (listed in Appendix II). Similar identification was carried out in the geometrical descriptions of the 5-coordination sphere (Auf der Heyde & Bürgi, 1989). In this way, along each of the four lines that connect DOD and SQUP geometry, the way in which a geometrical form transforms to the other will be observed easily. For instance, starting from the SQUP in the centre, either of four DODs may potentially be formed, depending on which of atoms on the diagonal of the square planes moves during the distortion, to lead subsequently to the vertex atoms of the tetrahedra in the DOD.

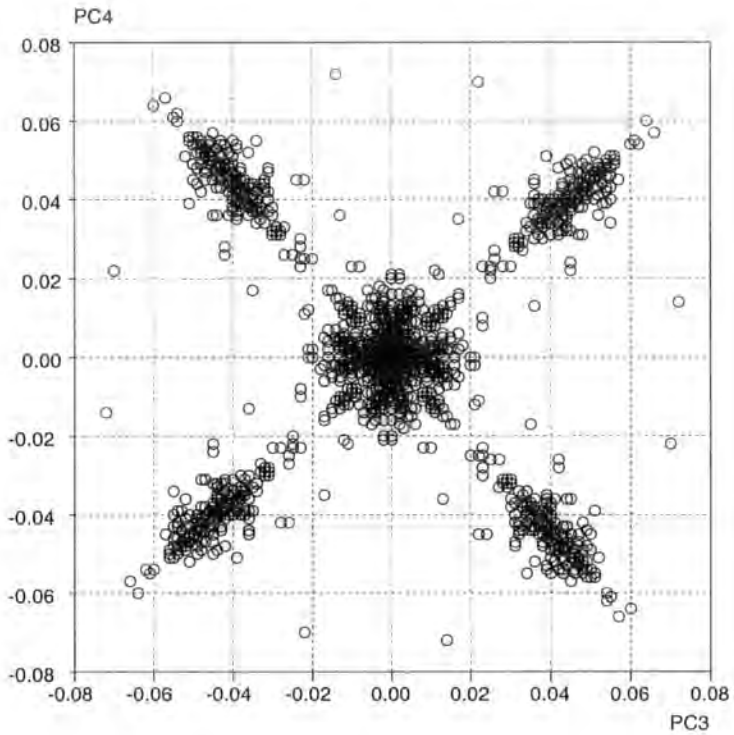
The plot of PC_3 against PC_2 [Figure 4.29(b)] gives a view from another direction. PC_2 is considered as a perpendicular axis to the PC_3-PC_4 projection. The scatterplot shows 2-fold symmetry of the D_{4d} point group. In this projection, four clusters for DOD structures are merged to two and situated on two sides of a U-shape and SQUP is located at the bottom centre of the U. The interconversion between these two geometries falls along the line of the U shape.

Figure 4.29(c) shows the correlation of PC_2 and PC_1. It has similar features to the PC_3-PC_2 plot.

FA was also applied on the data set. With the selection of *12* factors, very similar results to the PCA in this symmetry were obtained. The corresponding factors to the top four PCs have been observed in the relationships of the angle loadings. It can be seen that FAC_1 is equivalent to PC_2, FAC_4 to PC_1. The other two pairs are: FAC_3-PC_3 and FAC_5-PC_4, in which PC_4 is negatively correlated to FAC_5.

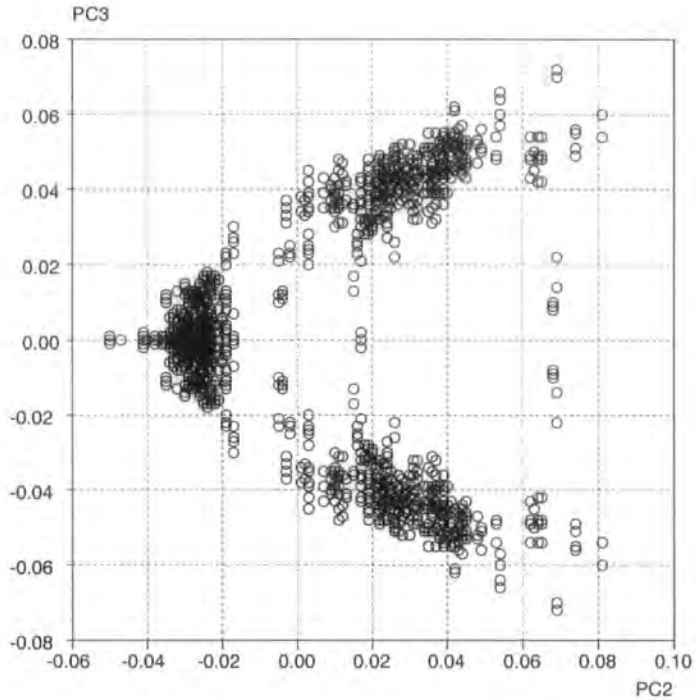
In D_{2d} and D_{4d} symmetries, PC/FA results have shown that the data set is clustered in terms of the geometry, i.e. a cluster for the DOD conformation and a cluster for the SQUP. The appearances of the clustering in the PC/FAC plots looks quite different in the different symmetries of the data spaces. In D_{4d} symmetry, which represents the symmetry of a standard SQUP form, geometries that tends toward the SQUP will naturally assemble around the archetype D_{4d} SQUP at the origin. On the other hand,

those structures that have geometries close to the DOD will cluster around an archetype whose geometry is as close as possible to that of a perfect DOD (D_{2d}). These differences are caused by the symmetry expansion but the results are, in fact, similar. They give different expressions in the different symmetry data spaces. All CSD hits in the cluster representing a DOD in D_{2d}-space, can be found again in the corresponding clusters representing the DOD in D_{4d}-space. Similarly, the CSD reference codes appearing in the clusters representing SQUP in D_{2d}-space, are found again in the central cluster

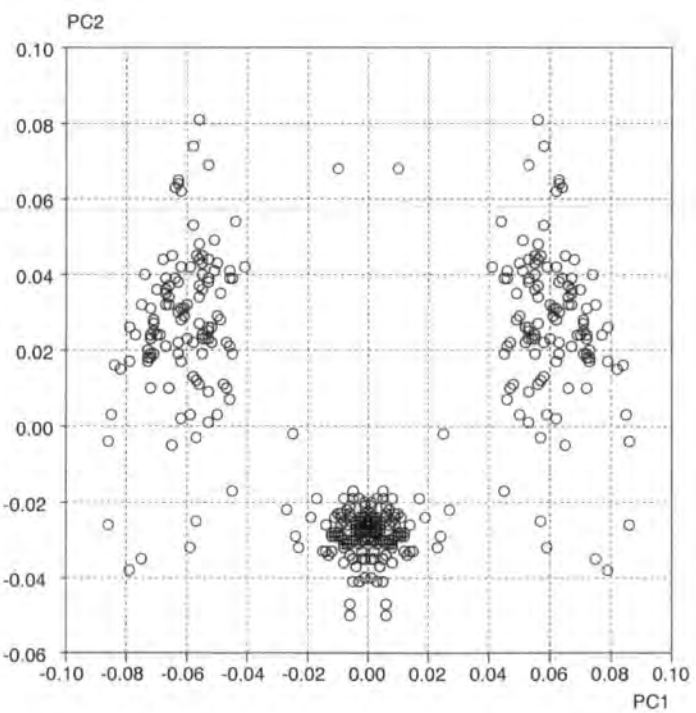


(a) PC_4 vs. PC_3

Figure 4.29 Scatterplots of PCs in D_{4d} symmetry



(b) PC_3 *vs.* PC_2



(c) PC_2 *vs.* PC_1 **Figure 4.29** *contd.*

representing the SQUP in D_{4d}-space. When an existing geometry for a compound (e.g. D_{4d}) is considered in the symmetry appropriate to another geometry (e.g. D_{2d}), the symmetry of the data set simply does not accord with its existence. Thus, geometries approaching a SQUP conformation are forced to cluster around a C_{2v} SQUP in the FAC_2-FAC_1 plot of D_{2d}–space. It also shows that FAC_2 represents the symmetry of D_{2d} and all geometries having the DOD conformation present around zero in this axis. Again, those compounds approaching the SQUP geometry aggregate around the single central cluster in D_{4d}-space, which represents their own symmetry.

The points located outside of any clusters represent geometry distorted greatly from either of the given reference geometries. If the detailed conformation of these complexes is examined, it is found that strained chelate rings are usually to be the cause.

4.4 Nine-Coordination Sphere

Complexes with a coordination number of nine are also found mainly with lanthanide and actinide elements. Some transition metals like rhenium can form nine-coordinate hydride compounds, e.g. $ReH_7(PPh_3)_2$, (Howard, Mead & Spencer, 1983) and $WH_6[PPh(CMe_3)_2]_3$ (Gregson, Mason, Howard and Spencer, 1984), but most transition metal examples have multiple metal-core forms e.g. $Co_2Au_2Ru_2(\mu^2\text{-}CO)_2(CO)_{10}(PPh_3)_2$ (Roland, Fischer & Vahrenkamp, 1983) or are coordinated with aromatic molecules, e.g. $Cr(CO)_3(\eta^6\text{-fulvene})$ (Koch, Edelmam and Behrens, 1982). These complexes have special geometrical shapes and cannot be described by the polyhedral forms given in Chapter 3 for the mono-metal central coordination sphere.

The CSD results gave a data set containing 265 lanthanide and actinide 9-coordinate complexes with a wide variety of ligand types. Again, the searches were performed using the same bit-screen restrictions as those for 7- and 8-coordination compounds except TOTAL-COORD-NO was assigned as nine for the central atoms in the 2D-CONSTRAIN menu of QUEST.

4.4.1 Geometry Identified by $R_{ang}(x)$ Values

Lanthanides have complexes in their most stable trivalent state, and tetravalent for uranium and thorium, among those in the data set. Table 4.12 lists the numbers of metal complexes presented in the data set. It can be seen that the complexes which appear most frequently are those of Nd^{3+}, La^{3+}, Eu^{3+} and Gd^{3+}.

In the systematic analysis of 7- and 8-coordination geometry, $R_{ang}(x)$ calculations for each complex were used to provide a quantitative indication of the degree to which an observed structure deviates from an idealized polyhedral form. In this section, the same procedure is applied in this higher coordination system.

As has been shown in Chapter 3, the nine coordinate geometrical sphere can be described principally by two polyhedral forms, the tricapped trigonal prism (TTP) and the capped square antiprism (CSA) (Figure 3.4). Examination and some discussions of these stereochemistries from different aspects can be also found in other work (Favas & Kepert, 1981; Guggenberger & Muetterties, 1976; Robertson, 1977). Although each

Table 4.12 Number of the complexes for various metals in the data set

Metal	No.	Metal	No.
La	29	Dy	17
Ce	17	Ho	10
Pr	12	Er	19
Nd	31	Tm	5
Sm	17	Yb	11
Eu	29	Lu	11
Gd	29	Th	13
Tb	10	U	5

geometry accounts for the observed structures in different ways, it was shown that difficulty usually arose in identifying exactly which archetypal form was the best description for a given structure, especially when strained ligands are included. Thus, the geometry was usually considered based on the ligand type. For instance, with unidentate and multidentate mixed complexes, for a CSA form, the unidentate ligand was often located at the capping position in the CSA form. This, in fact, is not always true from the current analysis.

Then, what is the performance from the $R_{ang}(x)$ calculation? It has been indicated, from the angles given for these two idealized polyhedral forms, that the difference between the angles defining each archetype becomes smaller compared with those for the lower coordination numbers and so the $R_{ang}(x)$ values for different geometries may become indistinguishable. However, an advantage in using $R_{ang}(x)$ values to identify the geometry of a coordination sphere is that it always provides the closest geometrical conformation to a reference polyhedron and assigns ligand labels accordingly. This also provides the possibility to use further criteria to see how an observed structure deviates from either of the idealized polyhdral forms. This at least also provides a convincing starting point for each possible geometrical form, since drawing conclusions directly from the determined structure concerning the correct geometrical conformation for this coordination number, is indeed very difficult and wrong assignments easily result.

Normally, identification of the geometry by $R_{ang}(x)$ calculation does not need to consider the ligand types of complexes. Furthermore, in 9-coordination, an $R_{ang}(x)$

calculation could indicate which individual compounds need to be examined. Firstly, the complexes with nine unidentate ligands are investigated. Again, this arrangement involves the least influence of the ligand conformations on the coordination sphere. However, unlike 7- and 8-coordination, there are varieties of ligands which can be selected for the unidentate ligands. Only water (H_2O) lanthanide complex were structurally characterized for this kind of complexes. Nine equivalent unidentate ligands from the oxygen of H_2O are restricted to the lanthanide nonahydrates $[Ln(H_2O)_9](EtSO_4)_3$ and $[Ln(H_2O)_9](CF_3SO_3)_3$ (Gerkin & Reppart, 1984; Harrowfield *et al*, 1983) and the hydride complex $K_2[ReH_9]$ (Knox & Ginsberg, 1964). The hydride complexes are not included in this data set.

$R_{ang}(x)$ values for TTP and CSA are listed in Table 4.13. The results show that the tricapped trigonal prismatic structure is preferred. The Ln-O distances for the capping positions (*7, 8* and *9*) are longer than those at the trigonal prism (*1-6*), which are in reasonable agreement with this geometrical form. The common structure is shown in Figure 4.30 (also an ethyl sulfate fragment is included).

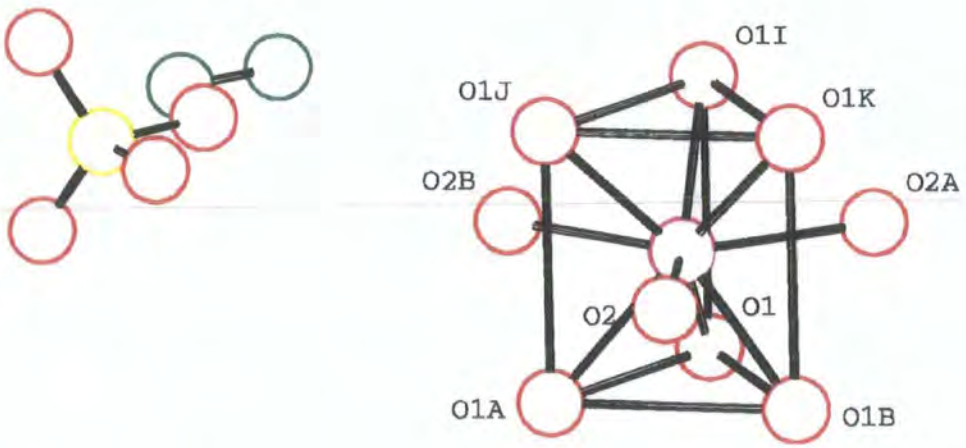


Figure 4.30 Structure for $Ln(H_2O)_9^{3+}$ sphere

In addition, there are several complexes with nine non-equivalent unidentate ligands. Two *1*-hydroxy-*2*-naphthoato (NAP2) substitute two H_2O molecules to form $[La(H_2O)_7(NAP2)_2]NAP2{\cdot}H_2O$ (FIYHOK) and $[Nd(H_2O)_7(NAP2)_2]NAP2{\cdot}H_2O$

Table 4.13. R_{ang} values(%) in geometry TTP and CSA

$[Ln(H_2O)_9](SO_3CF_3)_3$				$[Ln(H_2O)_9](SO_4Et)_3$		
Refcode	**Ln^{3+}**	**R_{ang}(TTP)**	**R_{ang}(CSA)**	**Refcode**	**R_{ang}(TTP)**	**R_{ang}(CSA)**
BUVVIX01	La	3.48	5.63	ZZZAQP01	3.95	5.80
BUVWEU01	Ce	3.33	5.59	CIBSAH	3.78	5.76
BUVWIY01	Pr	3.26	5.57	APRETS02	3.71	5.74
BUVWOE01	Nd	3.22	5.56	ZZZAQS01	3.57	5.70
BUVWUK01	Sm	3.16	5.53	DEYYIP	3.46	5.67
BUVXAR11	Eu	3.09	5.52	ZZZAQY01	3.39	5.65
BUVVOD01	Gd	3.08	5.52	ZZZARA01	3.44	5.65
BUVXEV01	Tb	3.11	5.51	ZZZARD01	3.38	5.63
BUVXIZ01	Dy	3.16	5.51	ZZZARG01	3.32	5.60
BUVXOF01	Ho	3.17	5.50	HOESUL02	3.33	5.61
	Er			AERETS	3.30	5.60
	Tm			ZZZARJ01	3.22	5.57
BUVYEW01	Yb	3.53	5.54	ESULYB01	3.24	5.58
BUVVUJ	Lu	3.60	5.58	ZZZARM01	3.32	5.57

(FIYHUQ) (Ohki, Suzuki & Ouchi, 1987). The formation of hydrogen bonds between the coordination sphere and the NAP2 anion distorts the structure from the regular TTP. The best description for the TTP is that one coordinated NAP2 locates at a capping position and the other at one of the trigonal prism vertices. Free NAP2 forms the H-bonds with two coordinated H_2O molecules. Similarity can also be found in complex *hepta-aqua-bis(maleato)-gadolinium(III) maleate monohydrate* (PEBDOP, Xue, Zhu & Yang, 1992).

HEDMUY has a coordination sphere of uranium with five H_2O, two $N{\equiv}CCH_3$ and two Br, i.e. $NH_4[U(H_2O)_5(NCMe)_2Br_2]Br_2$, (Zych *et al*, 1993). It was considered as an intermediate of the TTP and CSA, $R_{ang}(TTP) = 4.81\%$ and $R_{ang}(CSA) = 4.28\%$. The best description for the TTP is that Br_1, O_1, O_{2B} and Br_{1B}, O_{1B}, O_2 locate at the vertices of two parallel triangle faces, respectively, N_2, N_1 and O_3 at the capping positions.

Other examples can be seen in Figure 4.31. All are considered to have the TTP form. $R_{ang}(TTP) = 3.26\%$ and $R_{ang}(CSA) = 4.91\%$ for JOSSEP (Starynowicz, 1992) and

$R_{ang}(TTP) = 4.37\%$ and $R_{ang}(CSA) = 5.28\%$ for TAZYOI (Starynowicz, 1991). Bulky ligands are all located at the vertices of the trigonal prism.

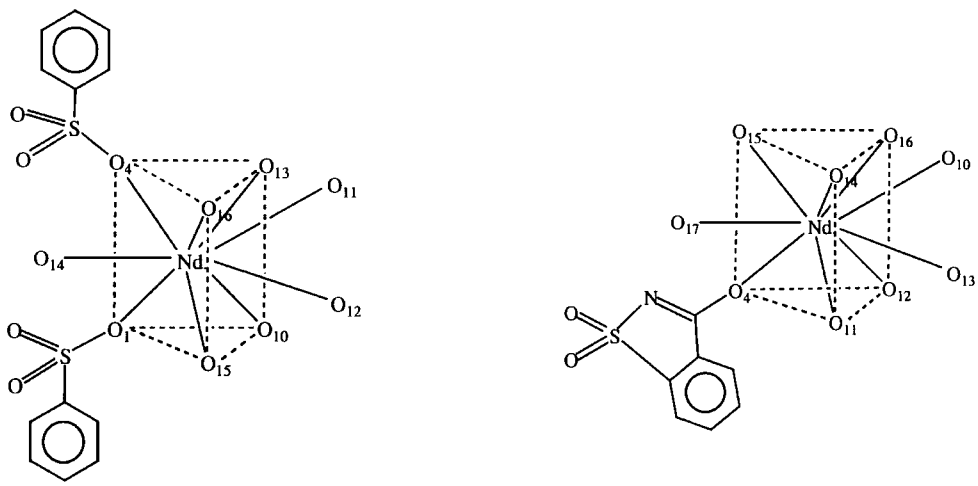
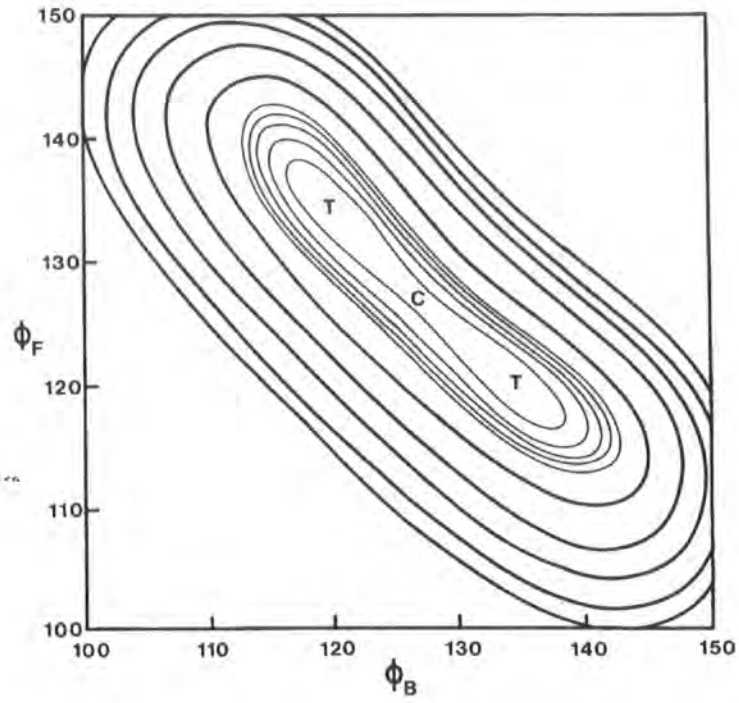


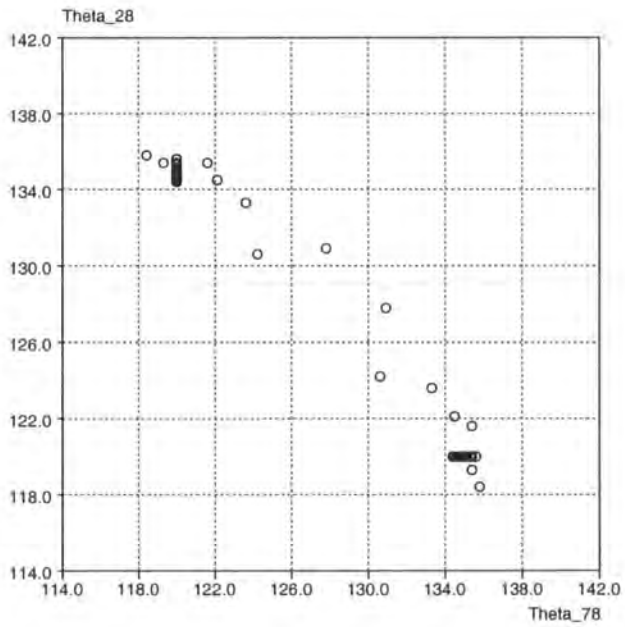
Figure 4.31 Coordination spheres geometries for JOSSEP and TAZYOI

The potential energy surface for the ML_9 coordination sphere has been calculated (Kepert, 1981). The projection onto the ϕ_B-ϕ_F plane is given in Figure 4.32(a), in which C is the CSA and T the TTP in this representation, and where the angles ϕ_B and ϕ_F in Kepert's paper correspond to θ_{28} and θ_{78} in this work, respectively (see Figure 3.4). There is no apparent potential energy barrier between these two geometries so it is easy to convert one to the other and also to be ambiguous in classifying the geometry for a determined structure.

The data set is divided into three subsets according to the ligand types of the complexes: (i) with all nine identical unidentate ligands (U); (ii) with three bidentate ligands (B) and three unidentate ligands, MB_3U_3; (iii) with four bidentate ligands and a unidentate ligand, MB_4U. In (i), there is no chelate effect to the coordination sphere geometry and in (ii) and (iii) the influences of ligands to the coordination sphere geometry could come from the chelate rings formed by the bidentate ligands.



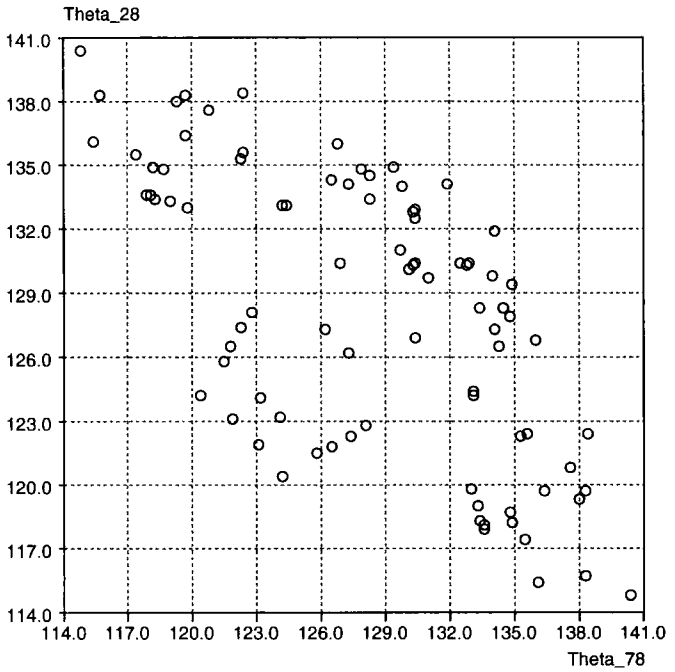
(a)



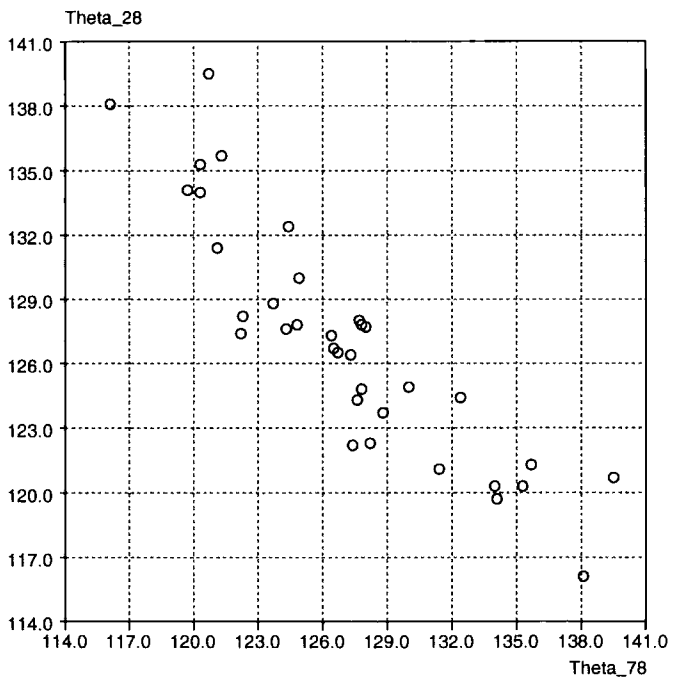
(b)

Figure 4.32 Projection of the potential energy surface (a) (reproduced from Kepert, 1981) and comparison from the observed structures, MU_9* (b);

* U — unidentate ligand; B — bidentate ligand



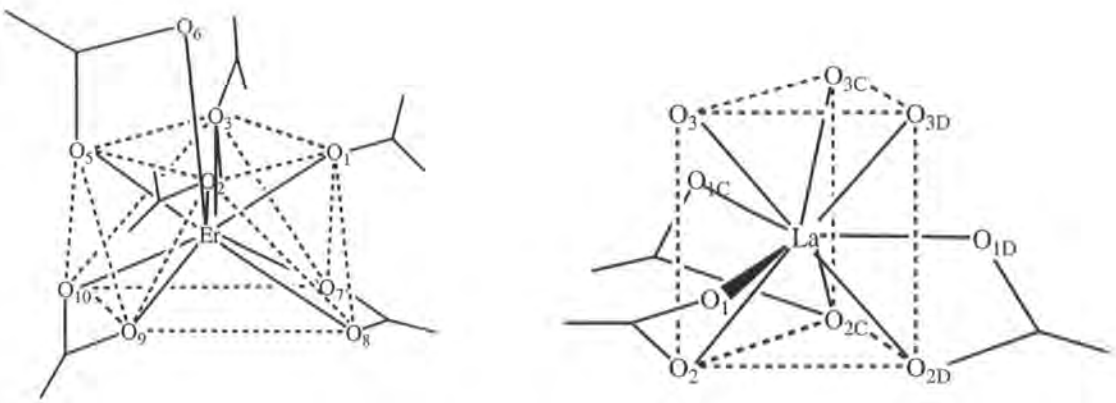
(c)



(d)

Figure 4.32 *contd.* MB_3U_3 (c) and MB_4U (d)

Geometries for the 42 observed structural fragments of [M(unidentate)$_9$] retrieved from the CSD are identified by the $R_{ang}(x)$ values. The angles θ_{28} and θ_{78} corresponding to the ϕ_B and ϕ_F in Kepert's potential energy surface map, are plotted for MU_9, MB_3U_3 and MB_4U complexes in Figure 4.32(b), (c) and (d), respectively and plots are in good agreement with the potential energy surface. There is actually no perfect CSA form in this kind of complex (i). Only a few have close geometry to this polyhedron, which can be seen as an intermediate close to the CSA. Most of them are clustered around the perfect TTP sites. Some structures between these two geometries represent the intermediates of these geometries. In the form of MB_3U_3, both TTP and CSA can be found. The most common conformation forms for the CSA and TTP in this data set,



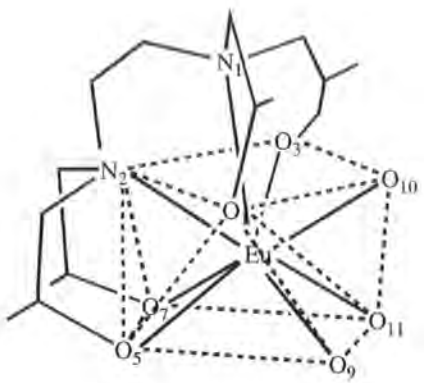
DISFUG (CSA) PEWJIK (TTP)

Figure 4.33 Representatives of geometrical conformations in MB_3U_3

identified by $R_{ang}(x)$ values, are presented in Figure 4.33. The three bidentate ligands and the prismatic ligands are related by 3-fold symmetry.

In the MB_3U_3 data set, some multi-dentate ligand complexes also included, but the complexes always have three unidentate ligands except for the multi-dentate ligands. A cluster in the CSA area [Figure 4.32(c)], but separate from the cluster described above is, in fact, another form of coordination, which consist of three unidentate ligands and a pentadentate ligand. In this data set, they are ethylenediaminetetra-acetato complexes.

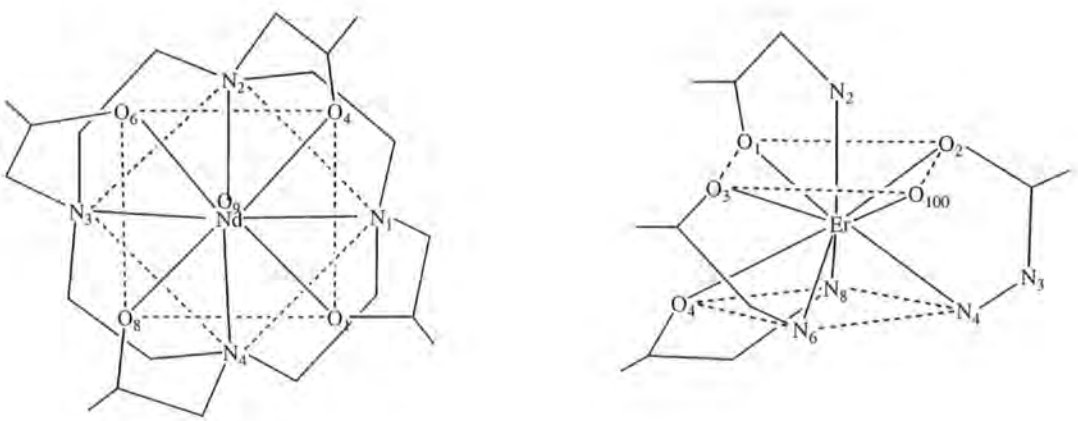
Only the CSA form can be formed here and the structural conformation is shown in Figure 4.34.



ZAMHOK

Figure 4.34 Geometrical form for M(pentadentate)U_3 in ZAMHOK

It is interesting that the capping position is not occupied by a unidentate ligand but by one atom of the multi-dentate ligand. Three unidentate ligands (H_2O) are sitting on the same side. It will allow the atoms on the square faces to be co-planar.



(a) (b)

Figure 4.35 Structural representations of observed complexes in CSA geometry from the CSD (a) JOPJIH; (b) CECLEB.

From the composition of the coordination form M(octa-dentate)U, it is easy to see that the CSA is a suitable polyhedron for the coordination sphere. Because the octadentate ligands can span edges linking the two square faces of antiprism and the unidentate one locates at the capping position. A typical structure of this type complex, JOPJIH, is shown in Figure 4.35(a). $R_{ang}(TTP)$ and $R_{ang}(CSA)$ are: 6.27% and 3.68%, respectively. For MB_4U complexes, the CSA geometry is more preferred [Figure 4.32(d)]. It was considered that the same conformation as that for M(octa-dentate)U described above should be formed (Kepert, 1981). However, most observed structures can not have this arrangement for the coordination sphere. Normal conformation for this type complexes identified by the $R_{ang}(x)$ values is shown in Figure 4.35(b)

It can be seen that two bidentate ligands span the edges of the two parallel square antiprismatic faces, one on the edge linking two atoms at the same square plane. The other contributes an atom as capping (N_2) and one on the capping base square plane (O_1). In this structural conformation, if the unidentate ligand O_{100} from H_2O molecule were considered as the capping atom, four bidentate ligands could not have the same conformation as JOPJIH, i.e. spanning the edges linking two parallel square faces of antiprism, which was previously predicted calculation (Kepert, 1981) for this kind of complex.

4.4.2 Further Criteria in the Identification of Geometry for Chelate Ligand Complexes

The geometry for most complexes having the defined coordination types given above have been identified by $R_{ang}(x)$ values. However, some complexes with chelate effects are usually in an intermediate state between the given reference polyhedra, TTP and CSA. Thus, $R_{ang}(x)$ calculation could not provide significant differentiation between TTP and CSA. This definitely gives difficulty in identifying the geometry for these complexes. Reference to the original publications for these complexes shows that they were mostly assigned by the subjective view of the structural projection illustrations. This actually included some random views and authors' interpretations, since with this coordination number and from different viewing directions, both geometries can be represented quite well. An example is shown in Figure 4.36. $R_{ang}(x)$ calculation for

complex KITGOJ $[Ho(H_2O)_4(EO_4)^*]Cl_3$ (Rogers *et al*, 1991) gave two values: *5.73%* in TTP and *5.98%* in CSA. From the diagram and these values, it is difficult to say which geometry it is really closer to.

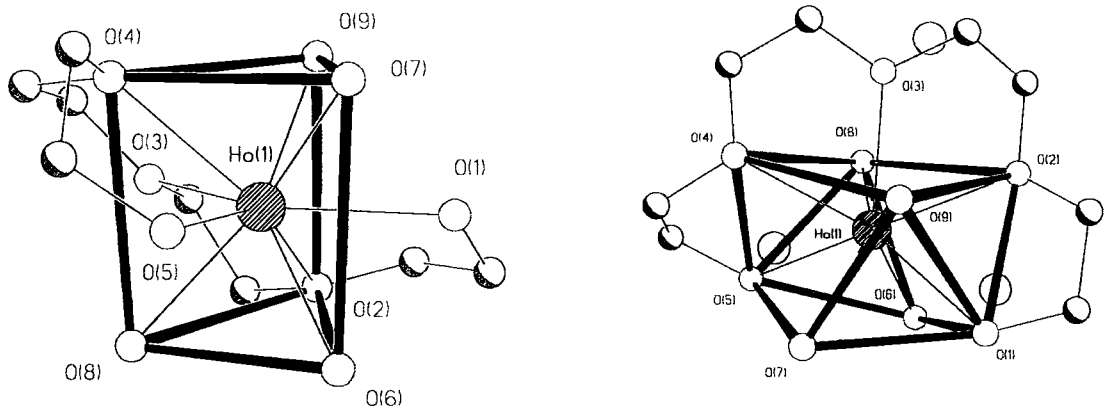


Figure 4.36 A structure KITGOJ from the CSD in both TTP and CSA

However, $R_{ang}(x)$ values provided us with TTP as the closest conformation to either of the geometrical forms and from this conformation, the ligand labels could assigned. Thus, in the TTP form of KITGOJ, O_4, O_7, O_9 and O_8, O_6, O_2 were assigned as the prismatic atoms and O_1, O_3, O_5 as the capping atoms. This definitely gives more accurate assignment than that obtained by eye and forms a basis for further geometrical analysis.

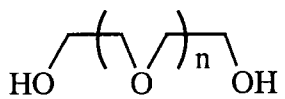
With this aim, further criteria for identifying geometry of these complexes are designed and applied in the procedure.

As is well known, for either of regular polyhedra, TTP or CSA, some idealized planes should be included. In TTP, there are three side planes of the prism, which are made up by vertices *1254*, *1346* and *2356,* and in CSA, two parallel square planes can be defined. Deviations for each atom included in these plane definitions for each complex are calculated and how much a given geometry distorts from the reference archetype might be found from these deviation values. It is obvious that the larger

* EO_4 — Tetraethylene Glycol

deviations represent the greater distortion from the defined polyhedron. Details of the application of this method to the lanthanide acyclic polyethylene glycol complexes are given as follows.

A series of complexes with a different number of coordinated oxygen atoms of polyethylene glycol molecules (Rogers *et al*, 1991) has been included in the data set. The multidentate ligands are expressed according to the number of the oxygen atoms on the molecular chain.



n =2 (EO_3), n =3 (EO_4), n = 4 (EO_5)

All these complexes have the same features as those in KITGOJ namely, small differences in the two $R_{ang}(x)$ values. The deviations from the defined planes given above for each vertex atom are calculated and the average deviation values for three planes in the TTP are defined as d_1, d_2, d_3, and d_4, d_5 for two planes in the CSA, respectively. All other coordination positions are occupied by unidentate ligands, H_2O or halide atoms. Table 4.14 lists the results of the calculation for lanthanide EO_4 complexes. This series of complexes involves *12* lanthanide element ions and before complex KITGAV (Tb^{3+}), the coordination spheres are composed by $Ln(H_2O)_3(EO_4)Cl$, i.e. a chlorine is coordinated directly to the metals. The remaining complexes all have the coordination form of $Ln(H_2O)_4(EO_4)$. When the CSA geometry was considered, it can be seen from Figure 4.36 that the capping position is occupied by an oxygen atom O_3 from EO_4 rather than by an expected unidentate ligand atom (H_2O). But in chloride complexes (Ln = Ce – Gd), the capping position is occupied by oxygen atom of H_2O. Since there is a tight ion pair in these complexes, the conformation of the alcoholic glycol is different from those complexes with four water molecules. The chlorine atom with two waters sits at one side of the acyclic chain of the EO_4 and one water at the other side. While in Ln = Tb – Yb complexes, two water molecules locate each side of the acyclic chain of the EO_4. Thus, the chloride ion, two alcoholic oxygen atoms and one water oxygen atom, form a capping base square face and one water oxygen sits at the capping position. However, all of these complexes have a geometry closest to the tricapped trigonal prism from the calculation of deviations from the defined planes, i.e.

Table 4.14. Deviations from the specific planes in TTP and CSA for $Ln(OH_2)_3(EO_4)Cl$ or $Ln(OH_2)_4(EO_4)$ spheres

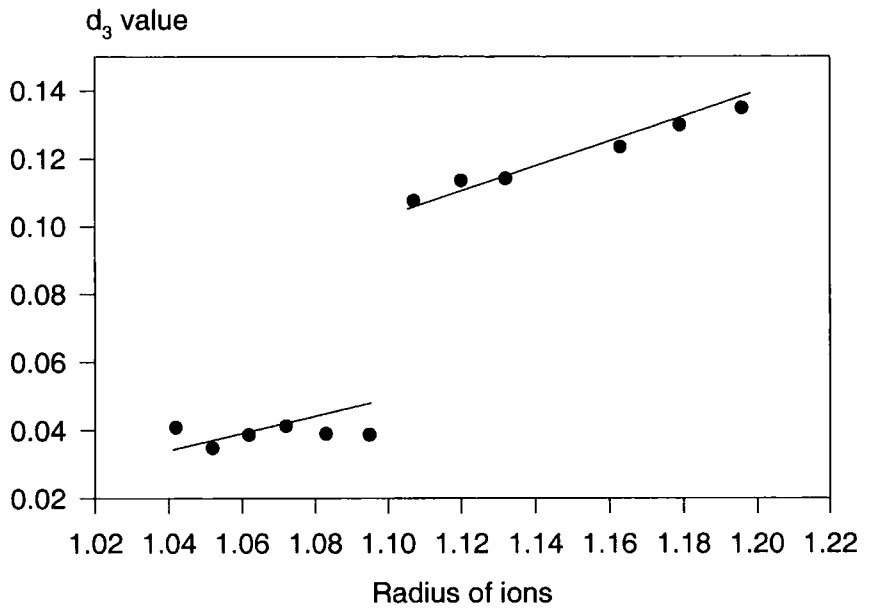
REFCODE	Ln^{3+}	TTP			CSA	
		d_1	d_2	d_3	d_4	d_5
KITDUM	Ce	0.118	0.041	0.135	0.041	0.236
KITFAU	Pr	0.112	0.041	0.130	0.041	0.238
KITEEY	Nd	0.104	0.041	0.123	0.041	0.235
KITFIC	Sm	0.097	0.044	0.114	0.044	0.233
KITFOI	Eu	0.095	0.040	0.114	0.040	0.230
KITFUO	Gd	0.087	0.041	0.108	0.041	0.232
KITGAV	Tb	0.122	0.037	0.039	0.122	0.221
KITGEZ	Dy	0.117	0.040	0.039	0.117	0.212
KITGOJ	Ho	0.115	0.038	0.041	0.115	0.206
KITGUP	Er	0.108	0.036	0.037	0.108	0.205
KITHAW	Tm	0.113	0.032	0.035	0.113	0.198
KITHEA	Yb	0.103	0.037	0.041	0.103	0.197

smaller d_i values in the TTP. In the first set of these complexes (with chloride coordinated), the glycol oxygen atoms alternate along the acyclic chain from prismatic to capping sites and in the second set (with four waters coordinated), the glycol oxygen atoms alternate along the acyclic chain from capping to prismatic sites. All water oxygen atoms are at prismatic sites, whereas the former has one water oxygen atom capping one of the quadrilateral faces.

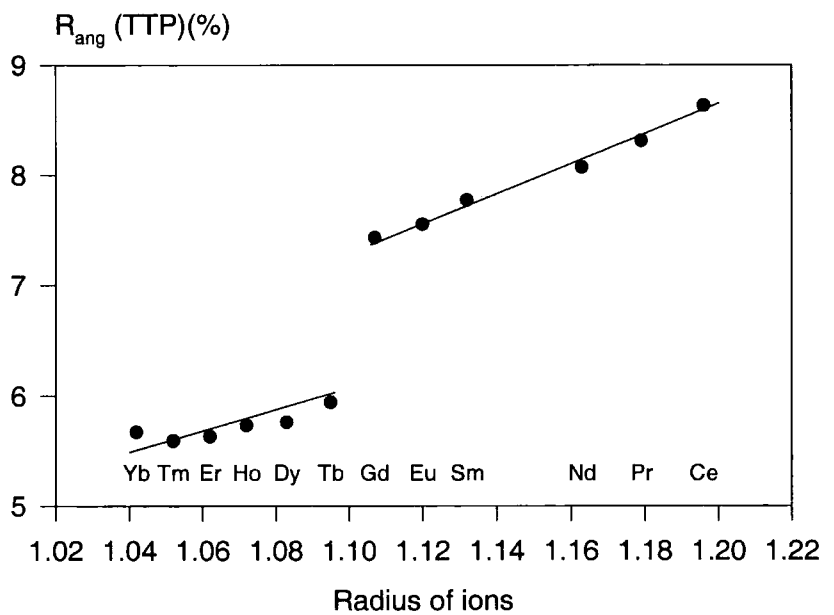
Similarly, the deviation calculation was carried out on the EO_3 and EO_5 complexes. The results are given in Table 4.15. All these complexes are, again, closest to TTP geometry.

The distortions in the geometry of the coordinated ion increase from lutetium to lanthanum. The deviation value d_3, which represents distortion of a complex from an

idealized TTP geometry correlates well with the ionic radius of the lanthanide ion. As shown in Figure 4.37, the distances of each coordinated atom, in complexes



(a)



(b)

Figure 4.37 d_3 value *vs.* radius of ions (a) and Rang(TTP) *vs.* Radius of ions

Table 4.15 Deviations from the specific planes in TTP and CSA for $Ln(OH_2)_3(EO_5)^{3+}$ or $Ln(OH_2)_5(EO_3)^{3+}$

Refcode	Ln^{3+}	$Ln(H_2O)_3(EO_5)^{3+}$				Refcode	$Ln(H_2O)_5(EO_3)^{3+}$			
		TTP			CSA*		TTP			CSA*
		d_1	d_2	d_3	d_5		d_1	d_2	d_3	d_5
	Nd					FUXPIX	0.068	0.068	0.139	0.244
HAMYOJ	Sm	0.122	0.079	0.139	0.254	KUYBAH	0.058	0.003	0.122	0.268
HANBIH	Eu	0.116	0.068	0.153	0.265	FUXPOD	0.059	0.014	0.116	0.249
HANBON	Gd	0.081	0.068	0.149	0.258	FUXPUJ	0.057	0.004	0.114	0.243
HANBUT	Tb	0.092	0.062	0.153	0.253	KUYBEL	0.049	0.005	0.117	0.240
HANCAA	Dy	0.097	0.064	0.147	0.253	FUXRAR	0.045	0.008	0.102	0.235
HANCEE	Y	0.099	0.052	0.138	0.240	FUXREV	0.058	0.011	0.115	0.228
HANCII	Ho	0.092	0.059	0.141	0.246					
HANCOO	Er	0.088	0.061	0.139	0.236					
HANCUU	Tm	0.093	0.057	0.133	0.241					
HANDAB	Yb	0.086	0.054	0.135	0.234					
HANDEF	Lu	0.092	0.054	0.121	0.230					

* one of the planes in CSA is also a plane in TTP, so the deviation value d_4 is not given repeatedly,

$Ln[(H_2O)_3ClEO_4]Cl_2 \cdot H_2O$ or $Ln[(H_2O)_4EO_4]Cl_3$, from one of the idealized planes in the TTP polyhedral form decrease regularly with the ionic radius (Cotton & Wilkinson, 1988) and show a linear dependence. It is interesting to note, however, that in a plot of the d_3 value versus the ionic radius of the trivalent metal, a break occurs in the curve at the gadolinium and terbium ions, which changes(a) and D_{4d} (b) symmetry, respectively. It can be seen that the contributions of the angles the coordination form from a chloride ion in the primary sphere to being replaced by a water molecule. The same correlation can be also seen in the plot of $R_{ang}(TTP)$ *versus* ionic radius, that is, the larger, earlier lanthanides are more distorted from the TTP geometry.

The same trend in the change of geometry of the coordinated ion is also observed in other series of complexes, such as, for 9 water molecules coordinated to lanthanide ions, Figure 4.38 shows such correlation again with $R_{ang}(TTP)$ against radius of ions.

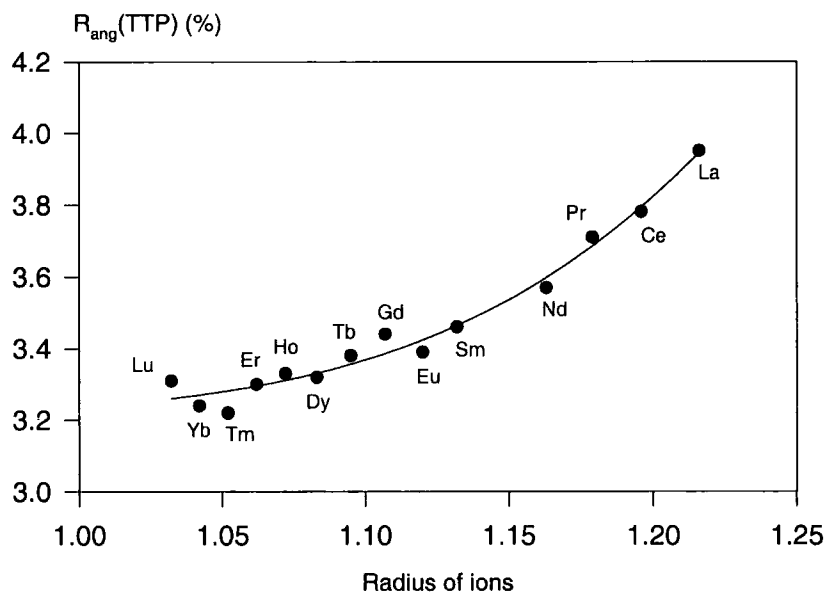


Figure 4.38 R_{ang}(TTP) *vs.* radius of ions in $Ln(H_2O)_9 \cdot 3EtSO_3$

It can be concluded that structures containing smaller ions are more likely to have TTP geometry than those of larger ions.

4.4.3 Symmetry Coordinates

As in 7- and 8-coordination, geometrical distortions and interconversions between the two idealized polyhedral forms can also be mapped by the correlation of some significant symmetry coordinates.

The symmetry coordinates are considered and can be derived in two symmetrical forms, D_{3h} and C_{4v}, which represent the two ideal geometrical symmetries, TTP and CSA, respectively. The total angle representations are:

$$\Gamma_\theta = 2A_1' + 2E' + A_2'' + E'' \qquad \text{in } D_{3h} \text{ point group;}$$

$$\Gamma_\theta = 3A_1 + B_1 + B_2 + E \qquad \text{in } C_{4v} \text{ point group.}$$

The symmetry isomers for 9-coordination sphere in point group D_{3h}(TTP) and C_{4v}(CSA) are given in Appendix III. As the coordination number increases, the symmetry coordinates become more complicated. The experiences in the study of 7- and 8-coordination sphere geometry showed me that it is not necessary to consider all coordintes. Meaningful results can be revealed by just small subset of the coordinates. Therefore, two pairs of these coordinates from two point groups (in representation E' of D_{3h} and E of C_{4v}) are presented as follows:

$$\begin{cases} S_{2a} = \dfrac{1}{\sqrt{8}}(2\theta_{12} - \theta_{13} - \theta_{23} - \theta_{46} - \theta_{56} + 2\theta_{45}) \\ S_{2b} = \dfrac{1}{2}(\theta_{13} - \theta_{23} + \theta_{46} - \theta_{56}) \end{cases} \qquad \text{in } D_{3h}$$

and

$$\begin{cases} S_{2b} = \dfrac{1}{2}(\theta_{28} - \theta_{78} - \theta_{89} + \theta_{58}) \\ S_{2c} = \dfrac{1}{2}(\theta_{13} - \theta_{14} - \theta_{36} + \theta_{46}) \end{cases} \qquad \text{in } C_{4v}.$$

The choice of these particular coordinates is because they represent the significant angular distortions in two geometries and these distortions could also result in the transformation of geometries. Figure 4.39 illustrates the angle distortions represented by these two pairs of symmetry coordinates. S_{2a} and S_{2b} represent the angle vibrations between prismatic atoms. These movements are equal to moving the vertex atoms of the prism away from an ideal trigonal prism. Figure 4.39(a) gives a view along the axis perpendicular to the triangle faces. While S_{2b} and S_{2c} show the distortions of atoms on the two parallel square planes in the CSA form. Figure 4.39(b) is a view along capped atom (8) to the metal. It shows that one square plane is distorted by the shrunken angle between atoms 1 and 4, 3 and 6, or stretched between atoms 1 and 3, 4 and 6; the other plane is distorted by the equivalent movements of atom 7, 9 and 2, 5, vertically above ⊙ and below ⊗ the plane of the paper, respectively.

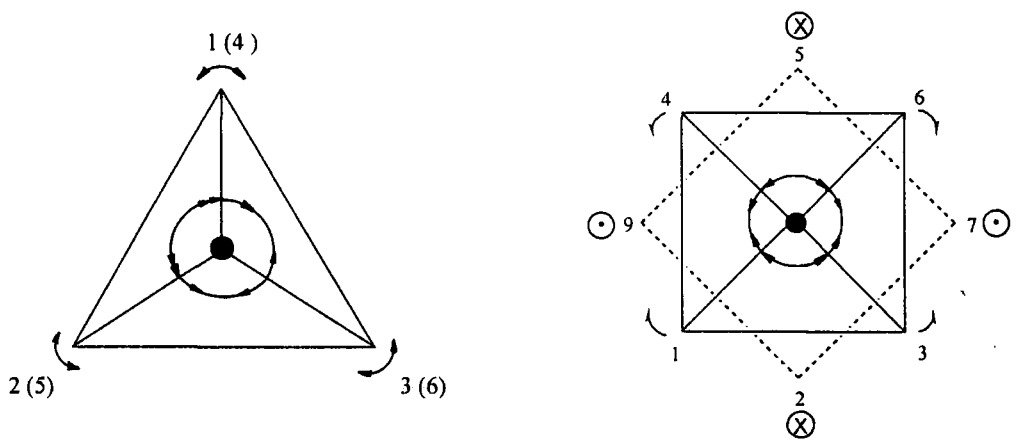
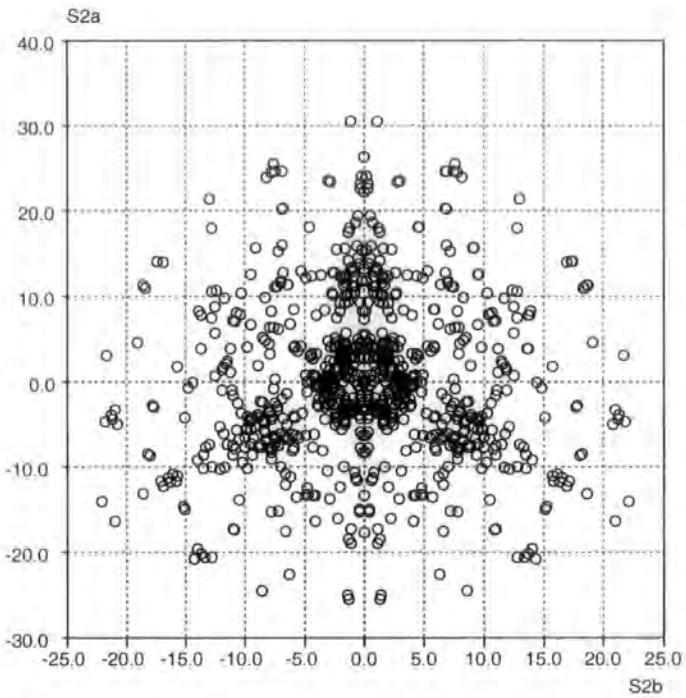


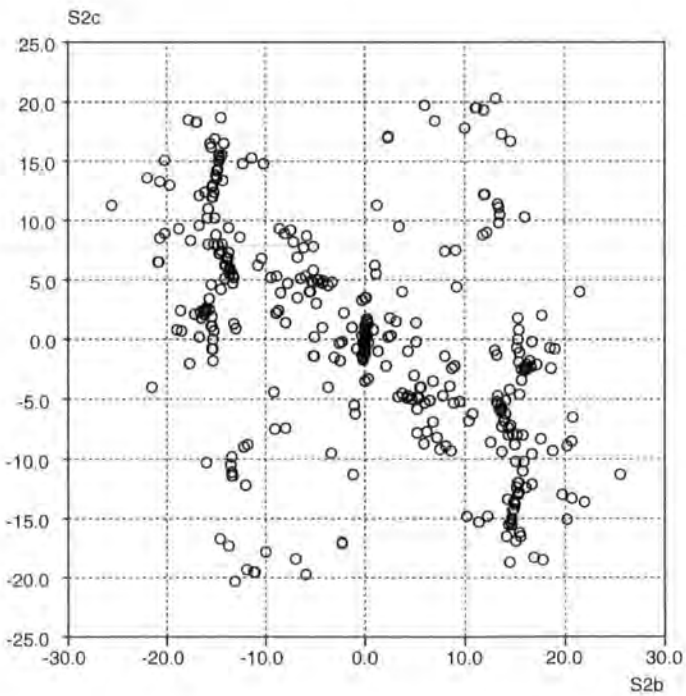
Figure 4.39 Visualized representations of angle distortions from Symmetry coordinates S_{2a} & S_{2b} in D_{3h} (a) and S_{2b} & S_{2c} in C_{4v} (b).

In the same manner applied in 7- and 8-coordination systems, all the CSD retrieved data for 9-coordination complexes are expanded in symmetry D_{3h} and C_{4v}. Thus, each point is related by 12-fold symmetry in D_{3h} and 8-fold symmetry in C_{4v}. The scattergrams of S_{2a} vs. S_{2b} [Figure 4.40(a)] and S_{2b} vs. S_{2c} [Figure 4.40(b)] on the data set for all examples in 9-coordination complexes both map the geometrical

deformations and interconversion from one to the other polyhedral form. In Figure 4.40(a), the central density is readily attributed to the TTP complexes. Other clusters related by 3-fold axes represent the CSA complexes. Whereas in C_{4v} symmetry [Figure 4.40(b)], CSA complexes are located at the centre and TTP complexes are related by a 2-fold axis. Both plots have a number of the experimental observations that fall on lines connecting these two polyhedral forms, that is, a mapping of the geometrical changes that take place along TTP $\Leftrightarrow$ CSA interconversion pathways in two different symmetry forms. In addition, there are some points on both plots located beyond the clusters, reference back to the details of the structures, it showed that all these complexes contain special ligands with large steric effects for the coordination sphere geometry. Some of the angles considerably deviate from those idealized angles due to the influence of the ligand conformation. For example, some crown ether complexes, Gd(EtOH)(18-crown-6)Cl_2]Cl, BITZUZ10 (Forsellini *et al*, 1985), [Ce(H_2O)Cl_2(18-crown-6)]Cl·H_2O HAMZEA (Rogers *et al*, 1993) and other macrocyclic complexes (JUVZIJ, Bligh *et al*, 1992).



(a) S2a vs. S2b in D_{3h}



(b) S2c vs. S2b in C_{4v}

Figure 4.40 Scatterplots of symmetry coordinates in D_{3h} (a) and in C_{4v}(b)

4.5 Geometrical Preferences and Coordination Environments

In the previous sections, the overall geometrical preferences, deformation and interconversion pathways for data sets of 7-, 8- and 9-coordination spheres have been mapped based on the geometry identified from the discrepancy index $R_{ang}(x)$. The all-unidentate ligand complexes always showed a clear geometry identification. However, most coordination spheres have more complex ligand types. With variety of ligand types, geometry may distort from the idealized polyhedral forms, To investigate the influence of the ligands, an index method has also been given. That is a measure of how an observed atomic position deviates from the idealized vertex position for a given polyhedral form. It provides geometrical information for each individual ligand.

For a systematic study of the influences of ligand type the geometrical preferences of a coordination sphere, firstly, a plot of two $R_{ang}(x)$ values is produced. Figure 4.41 is a scatterplot of $R_{ang}(PBP)\%$ against $R_{ang}(CTP)\%$ for the complete 7-coordination data set. Two major clusters with $R_{ang}(PBP)\%$ value close to zero and $R_{ang}(CTP)\%$ to zero, respectively, indicate a well populated PBP and CTP. COC appears between these two clusters and closer to the CTP cluster. A series of data points with intermediate $R_{ang}(x)$ values connect these clusters, which provide interconversion pathway of PBP $\Leftrightarrow$ CTP. It is interesting that not only a linear connection (A) is observed but also reveals interconvertions (B, C) between the PBP and CTP structures linked to a loose cluster of fragments for which both $R_{ang}(x)$ values are ~15%.

To examine the features provided on this plot, colour coding based on ligand denticity is applied and the coordinated ligand types are classed according to the number of unidentate ligands containing in the complexes. Colours red, green, yellow, blue, pink and black represent that there are 1, 2, 3, 4, 5 and 7 unidentate ligands in the complexes. The symbols □ and Δ indicate that at least a four-member or a three-member chelate ring is included in a complex, respectively. Thus, data points can be located according to the ligand types in this classification. For examples, for the complexes with two unidentate ligands (green), they mainly populate the PBP area.

Examination of the contributors shows that they arise mostly from the iron complexes, which have been discussed in previous sections (§4.2.3 and §4.2.4). The nature of the ligands makes it easy to form PBP geometry by arranging a penta-dentate ligand on the equatorial plane and two unidentate ligands at two axial positions. There

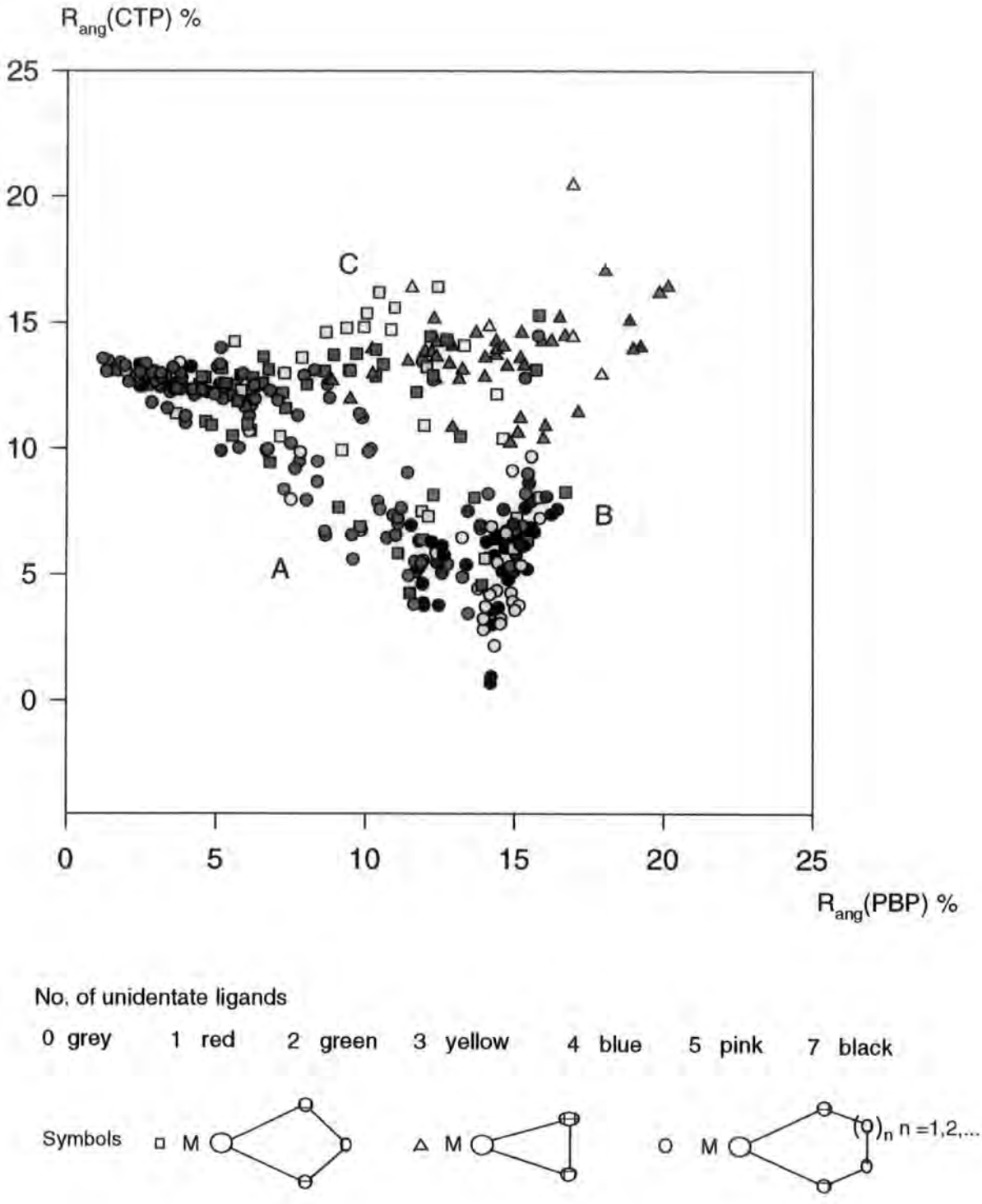


Figure 4.41 $R_{ang}(CTP)$ % vs. $R_{ang}(PBP)$ % in 7-coordination

are two kinds of distortions from PBP geometry. One is along line C, the other is along line A. The former has a bidentate ligand lying in the pentagonal plane and the latter has one equatorial atom out of the plane. Complexes with a three-membered chelate ring contribute to the cluster that has two large $R_{ang}(x)$ values. For the complexes with 4 unidentate ligands (blue), since three of them can be located at the apical positions of CTP/COC, most points can be found in this area. All unidentate ligand complexes are expressed in black. The details for this type of complexes have been given in the previous sections. The remaining classified complex with one (red), three (yellow) and five (pink) unidentate ligands have more varied geometry. Three possible geometries are adopted according to the chelate rings located on the equatorial plane or along the axis direction. They all have the same features as those represented in green, that is, the fragments which involve tight ligand bite angles (Δ) appear in the area with two large $R_{ang}(x)$ values and the fragments which have a bidentate ligand on the equatorial plane are distorted along line C. For the chelate rings with greater than 4-members symbolized as (○), the chelate effect for the coordination sphere geometry is not very apparent.

If individual structures along the interconversion line from PBP to CTP (A) from the plot are studied, a clear conversion of geometry from one to the other can be observed. Figure 4.42 shows this change of geometry from the observed coordination sphere structures [from structure DETKOC (PBP) to DOWKAB (CTP)]. Structure HARVEB represents a COC geometry, which is at between PBP and CTP structures. This is in agreement with the interconversion pathway proposed in Figure 3.9.

It can now be seen that the 2D $R_{ang}(CTP)$ vs. $R_{ang}(PBP)$ plot for the 7-coordination dataset not only presents all of the information detailed in the previous mappings by other techniques, e.g. PCA etc., but also allow us to classify the geometry according to different ligand types by colour coding. More importantly, the 2D $R_{ang}(x)$-plot is in a simpler form and is much easier to understand.

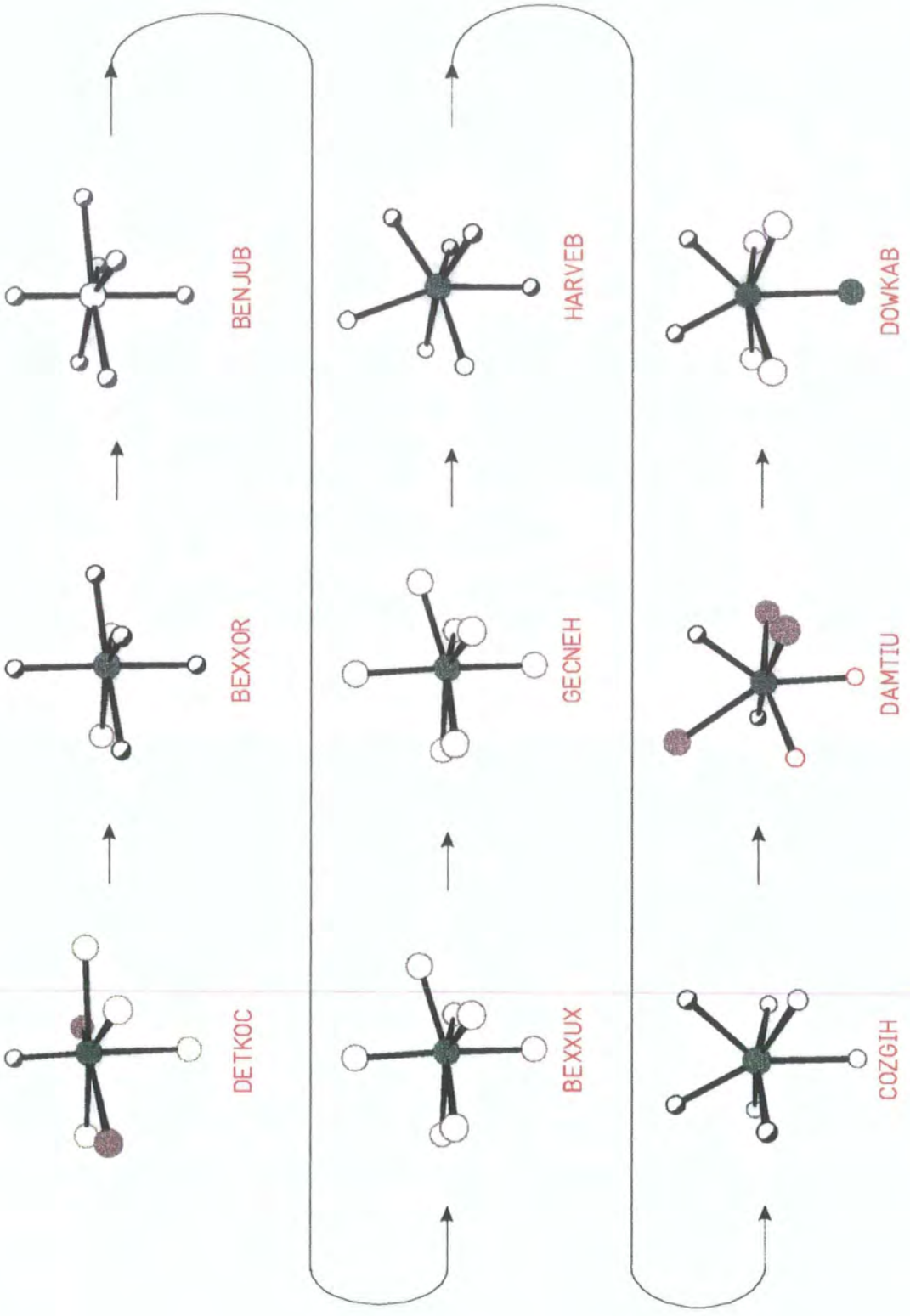


Figure 4.42 Geometry conversion from PBP to CTP by the observed Structures

4.6 $R_{ang}(x)$ Index Applied to Lower Coordination ML_n Systems (n =3, 4, 5 and 6)

The $R_{ang}(x)$ index method has been successfully applied in the higher coordination number system. The results in systematic study of ML_n sphere (n = *7, 8* and *9*) geometry have shown that it not only provides geometry identification for a specific crystallographically observed coordination sphere, by a clear numerical measure of deviation of the observed structure from that of given idealized polyhedron, but also generates a unique atomic enumeration for each fragment. Thus, a multivariate data matrix can be formed and further multivariate analyses can be carried out, so that the geometrical properties in these systems can be mapped.

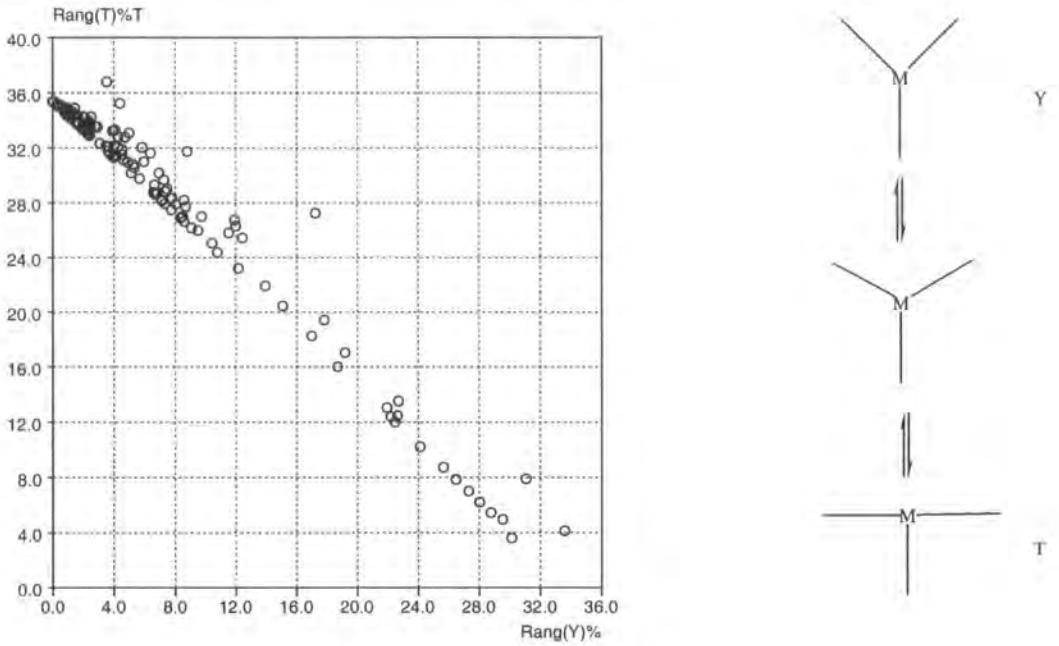
For a general application of this method, the $R_{ang}(x)$ index method is also applied to the lower coordination sphere ML_n (n =*3, 4, 5* and *6*). As stated previously, the well-known archetypal polyhedra used in the geometrical description of these coordination spheres all have specific defined angles and are easier to identify than those in the higher coordination number systems.

The lower coordination number complexes are much more common than higher coordination number complexes. The CSD contains a lot of structure data for these complexes. To restrict the numbers of retrieved structures, the search may usually be run separately for specific metal atoms at the centre, or to locate all unidentate examples. The total coordination number for the central atom (transition metals only) are restricted to 3, 4, 5 and 6, accordingly. The results from the CSD search reveals that 140 fragments in 3-coordination, 1306 in 4-coordination, 218 in 5-coordination and 1157 in 6-coordination have all-unidentate ligand complexations.

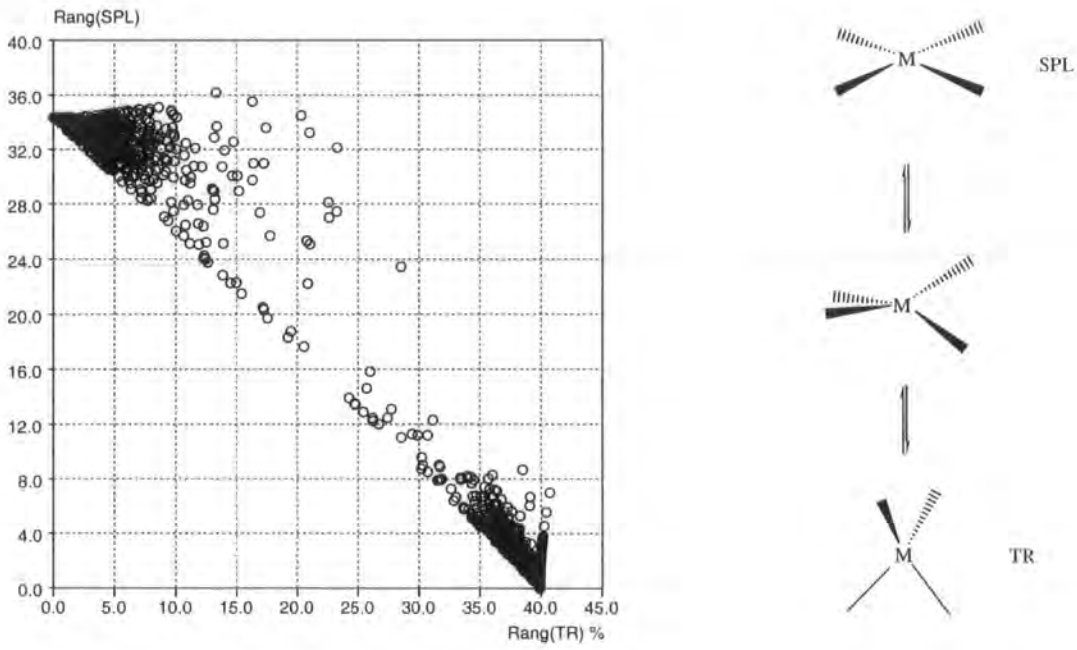
All $R_{ang}(x)$ values are calculated on these data sets. As in 7-coordination system, each calculation is over $n!$ ligand permutations and minimized the $R_{ang}(x)$ value to the different idealized polyhedral forms as described in Chapter 3.

4.6.1 2D $R_{ang}(x)$ Plots

It is obvious that geometry for each coordination sphere can also be identified by the relevant $R_{ang}(x)$% values for these lower coordination spheres. 2D $R_{ang}(x)$ plots are used again to elucidate geometrical preferences and distortions of these coordination spheres. The resulting plots for all-unidentate ligands 3, 4, 5 and 6-coordination data are

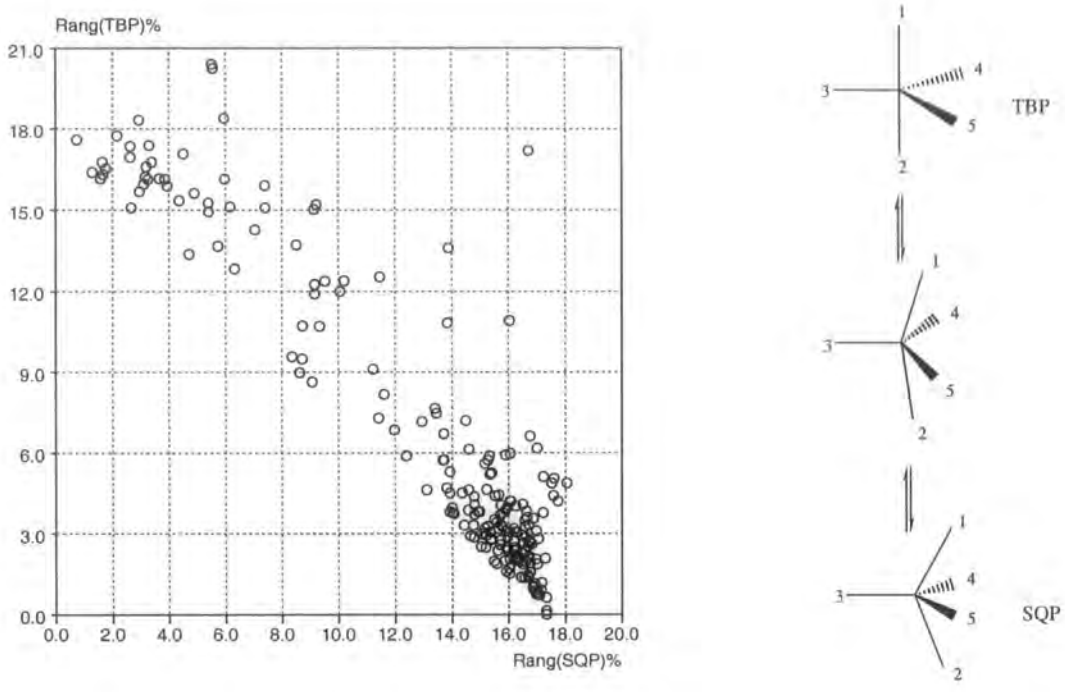


(a)

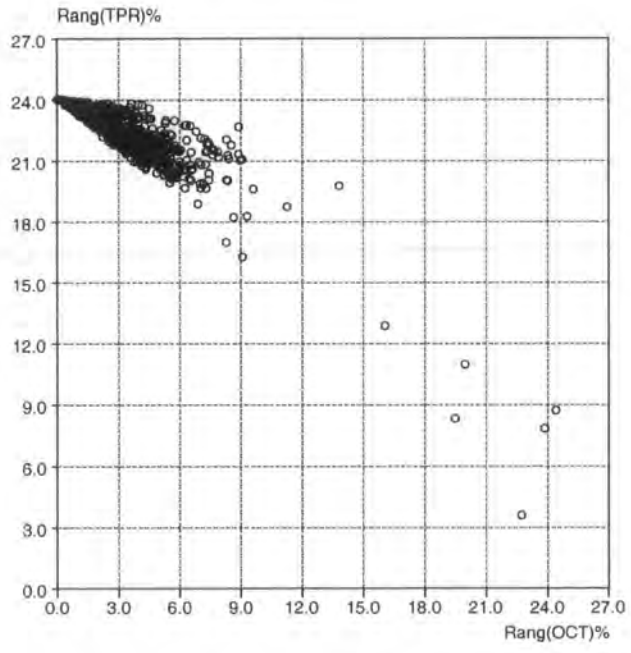


(b)

Figure 4.43 2D $R_{ang}(x)$% plots in (a) 3-coordination, (b) 4-coordination



(c)



(d)

Figure 4.43 *contd.* (c) 5-coordination (d) 6-coordination

given in Figure 4.43 (a), (b), (c) and (d), respectively.

As in the previous examples in 7-coordination, these plots reveal the two populated clusters for the given archetypal polyhedra and also provide an interconversion pathway between these two geometries along a line for each coordination sphere.

In the case of 3-coordination, trigonal plane geometrical forms of in Y and in T shape are used in $R_{ang}(x)$ calculations. Data points with intermediate geometry connect these two extremes. Examination of the contributors along the line shows that the geometry transforms smoothly from one shape (Y) to the other (T) or vice versa. Although the observed structures differ with respect to the central atom and the nature of the ligands, they may be arranged in a sequence that vividly illustrates the smooth transition along the line provided on the plot from T (structure MEPHGC10) to Y (structure SOJVES) shape (Figure 4.44).

Similar results can be also seen in 4-coordination data and 5-coordination data, which map geometrical forms of squar plane (SPL) and tetrahedron (TR) in 4-coordination, and trigonal bipyramid (TBP) and square pyramid (SQP) in 5-coordination, respectively.

In 6-coordination, the octahedron (OCT) is obviously the favored geometrical form. Only one exception (JAMWOJ), 2Li(TMED)* ·$ZrMe_6$ (Morse & Girolami, 1989), has the geometry of a trigonal prism (PTR) [$R_{ang}(PTR)$ = *3.6 %*]. another five (BERDEH, KOFWEH, PDACCO, SONROC, YULGUH) have distorted PTR forms, which mostly are caused by ligand conformations. All others were assigned as octahedral.

4.6.2 $R_{ang}(x)$ spectra

$R_{ang}(x)$ values, in fact, can be seen as a kind of dissimilarity measure which are usually used in the cluster analysis as given in Chapter 2. When torsional dissimilarity calculations were applied in the systematic study of conformation of a flexible substructure, a rapid visualization of the conformational complexity of a given chemical substructure could be obtained through the representation of the dissimilarity values as a simple histogram. Similarly, dissimilarity calculations on coordination sphere from

* TMED---- tetramethylethylenediamine

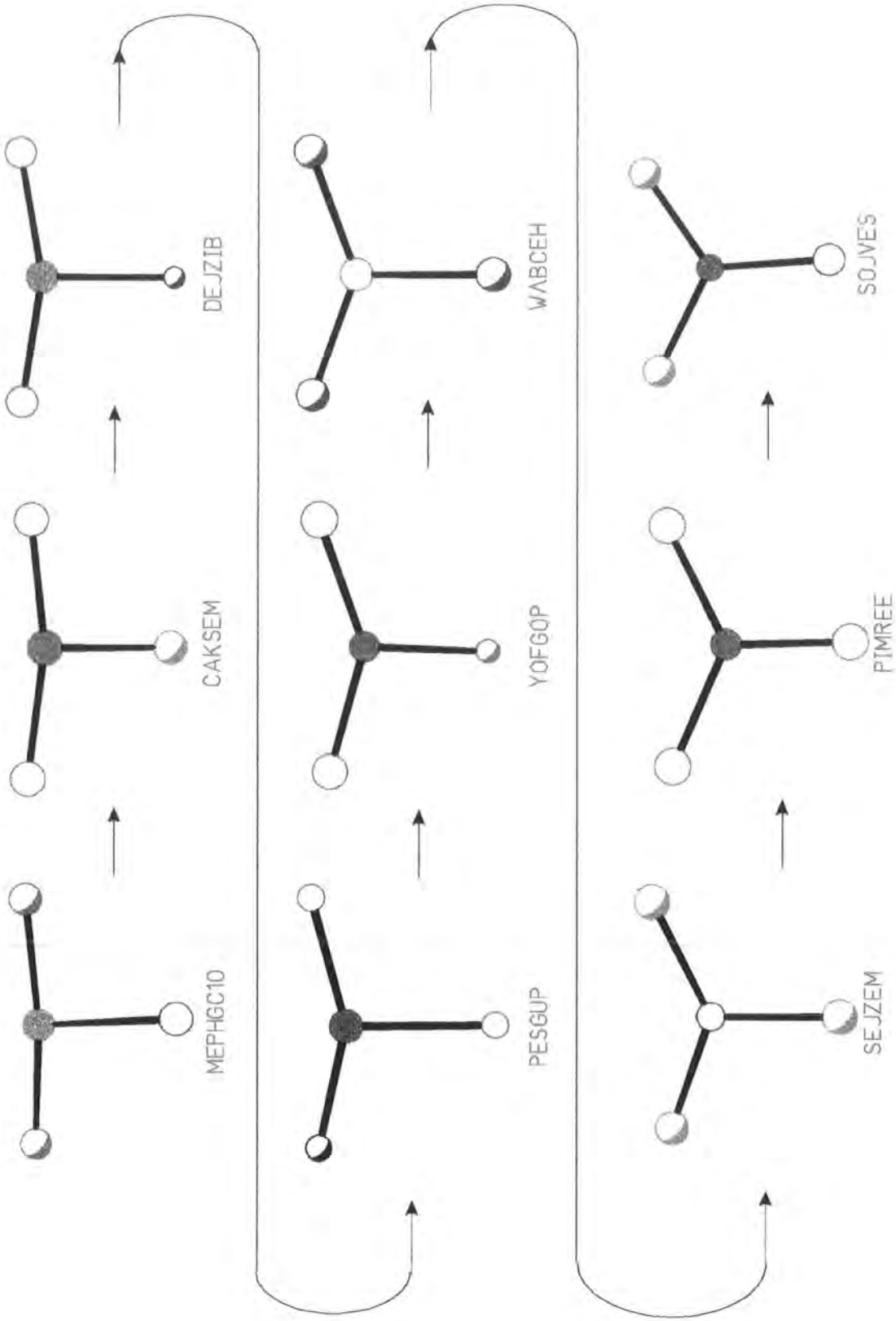


Figure 4.44 Transition from T to Y shape of trigonal plane by observed structures of ML_3 sphere from the CSD data.

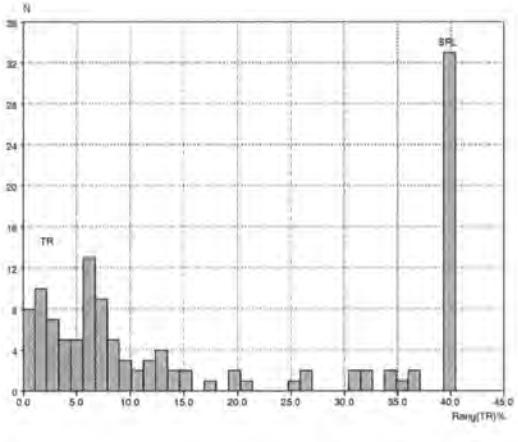
L-M-L valence angles can also reveal the different geometries represented by different well-known archetypal polyhedra. Such representations were called "angular spectroscopy" (Allen, Bath and Willett, 1994). For this study, the $R_{ang}(x)$ values, except for the 2D plots given above, can also provide a one-dimensional representation as a simple histogram as shown in Figure 4.45. It will be called "$R_{ang}(x)$ spectra".

The $R_{ang}(x)$ spectrum is a simple expression of the multivariate parameter space. The frequency of various values appearing in the histogram will indicate geometry preferences for different polyhedral forms in coordination sphere. However, for an n-coordination spheres, since the dimensionality to define each geometry is reduced from n(n-1)/2 to one, information for higher coordination numbers (n > 7) may lead to information loss and overlap of peaks for different geometries. Nevertheless, this can yet be regarded as a simple and effective representation of conformational or configurational diversity in ML_n ($n \leq 7$) sphere geometry.

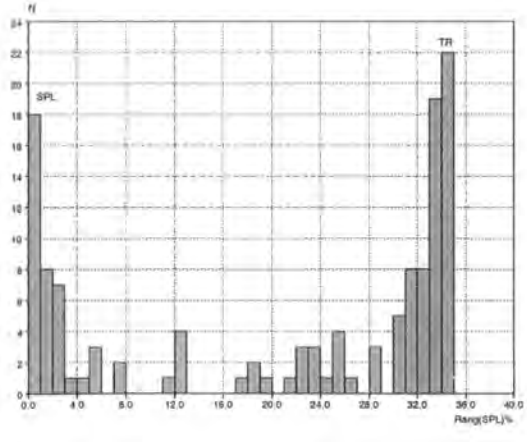
As an example of $R_{ang}(x)$ spectrum, the details of tetracoordination spheres for some common metals (Cu, Co, Ni, Pt and Zn) are given as below.

4-coordination Cu sphere

The data for this species is the same as that used in PCA in Chapter 2. There are a large number of complexes in the CSD. To restrict the number of examples retrieved from the CSD, all-unidentate ligand complexes are only involved in the data set (*127* fragments). Figure 4.45 illustrates histogram of $R_{ang}(TR)\%$ (a) and $R_{ang}(SPL)\%$ (b).



(a)



(b)

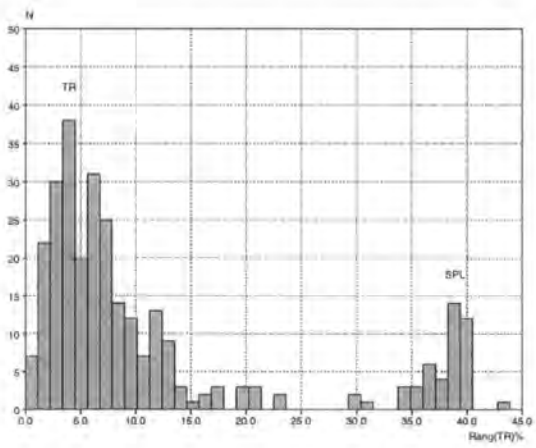
Figure 4.45 Histograms of $R_{ang}(TR)\%$ (a) and $R_{ang}(SPL)\%$ (b) in CuL_4

Both histograms show the peaks corresponding to two major different geometries, one close to the origin and the other at a $R_{ang}(x)$ of 30-40%. The peaks indicate the frequencies of the major components represented by the relative $R_{ang}(x)$ values and the information presented in mappings by more complex multivariate analysis, e.g. PCA, concerning the transition from one to the other geometry, can also be visualized simply by the series points with intermediate $R_{ang}(x)$ values between the two major peaks. Furthermore, in these 4-coordination Cu species, the peaks for the tetrahedral geometry are broadly spread in both spectra, particularly that close to zero in the $R_{ang}(TR)\%$ histogram [Figure 4.45(a)], which indicates a large number of distorted tetrahedral structures, while the peak for square plane is sharp in both histograms. This implies that a perfect square plane shape is easier to maintain once the coordination sphere adopts this geometry.

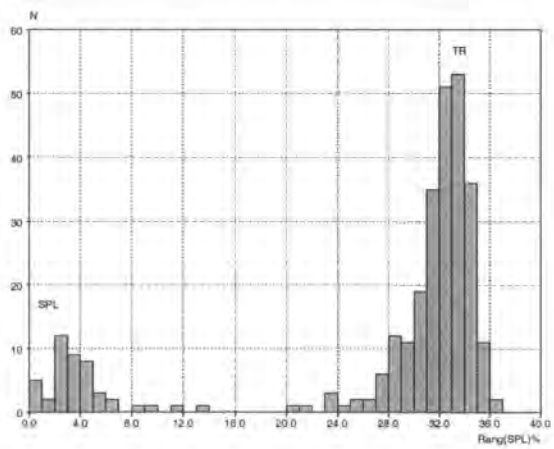
4-coordination Co sphere

The CSD retrieval provides *291* examples of tetracoordinate Co atoms. Figure 4.45 (c) and (d) give the distributions of $R_{ang}(TR)\%$ and $R_{ang}(SPL)\%$, respectively.

Tetracoordinate Co(II) always exists in the d^7 configuration. From the considerations of *d* orbital energies, high-spin electron arrangements of d^7 have stable tetrahedral form and square planar complexes often occur for the low-spin state of Co(II). Thus, those complexes with halide and carbonyl etc. as ligands, which have



(c)



(d)

Figure 4.45 *contd.* Histograms of $R_{ang}(TR)\%$ (c) and $R_{ang}(SPL)\%$ (d) in CoL_4

relative weaker ligand field, are easy to be found in the tetrahedral peaks (TR). All square planar complexes in SPL-peak are those ligands with stronger ligand field like -NH- etc.

4-coordination Ni and Pt spheres

1078 Ni complexes and Pt complexes were retrieved from the CSD. They all have d^8 electron configurations and exist in Ni(II) and Pt(II) forms in 4-coordination complexes. Both tetrahedron and square plane are found in 4-coordination Ni spheres [Figure 4.45 (e), (f)], most of them form square planar complexes. Tetrahedral complexes mainly contain ligands like –P, CO etc. Bulky ligands always make the complexes have the intermediate $R_{ang}(x)$ values.

It was considered in the crystal field stabilization energy (CFSE) (Mackay & Mackay, 1981) that the larger ΔE values gained by heavier elements in CFSE extend the scope of formation of square complexes. Thus, almost all complexes of Pt(II) are square planar [Figure 4.45 (g), (h)]. The exceptions are found in those complexes with chelating ligands in which a square plane is difficult to be formed, such as, structure CUTHEE, which has the closest tetrahedral geometry in the data set [$R_{ang}(TR) = 4.69\%$] and GAVFAK01 [$R_{ang}(TR) = 8.44\%$] and so on.

CUTHEE (Ling *et al*, 1985)

GAVFAK01 (Asker *et al*, 1990)

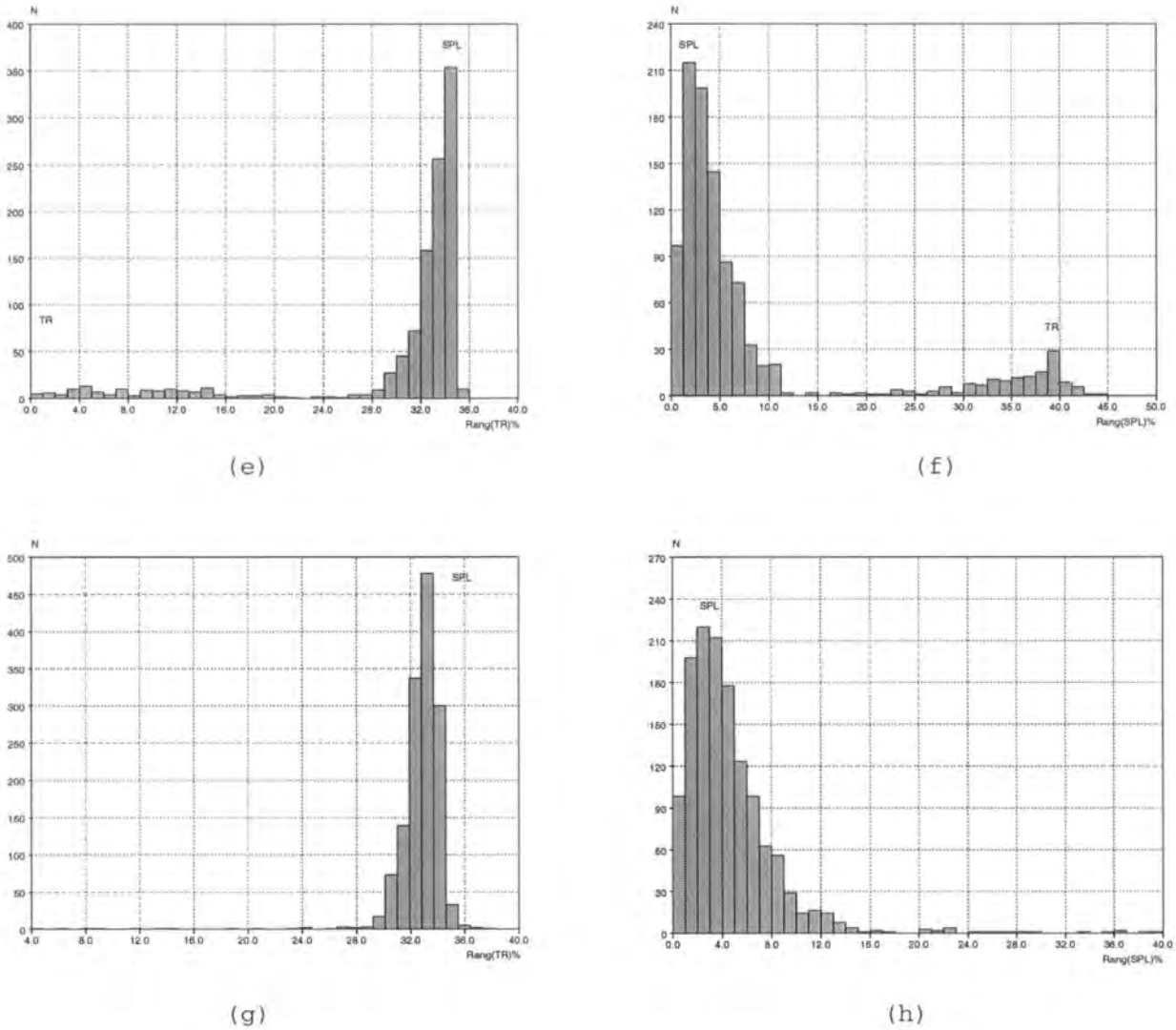


Figure 4.45 *contd.* NiL_4 (e), (f) and PtL_4 sphere (g) (h)

The *2D-$R_{ang}(x)$%* plot of PtL_4 data set (Figure 4.46) shows the effects of coordinated ligand environments to the geometries. Simple inspections of the CSD structures show that *trans*-coordinated forms normally have regular square planar complexes and *cis*-coordinated complexes, especially those with bulky ligands, are distorted from an idealized square plane. The distortions of chelate complexes from a SPL are located along *$R_{ang}(SPL)$%* axis, which shows that the smaller chelate ring results in larger distortion.

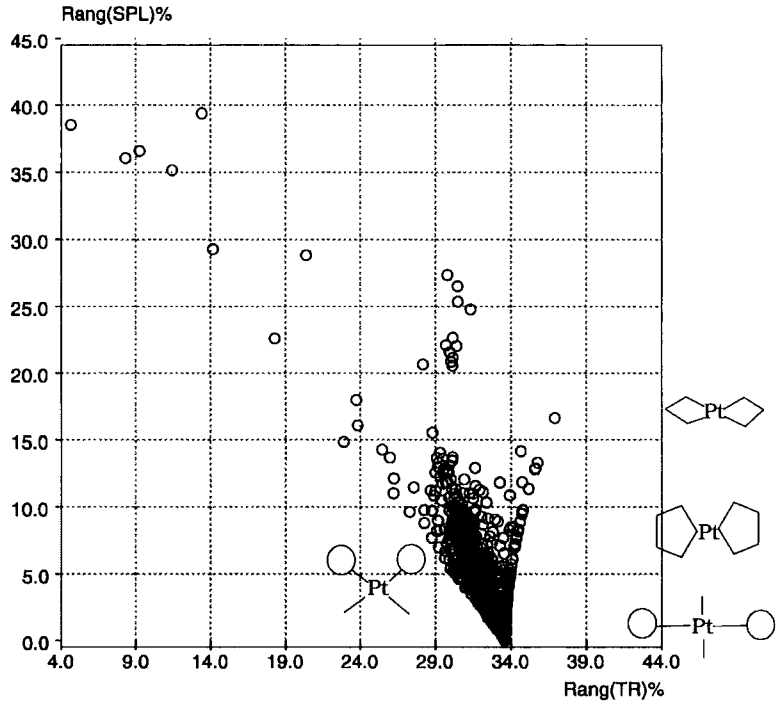
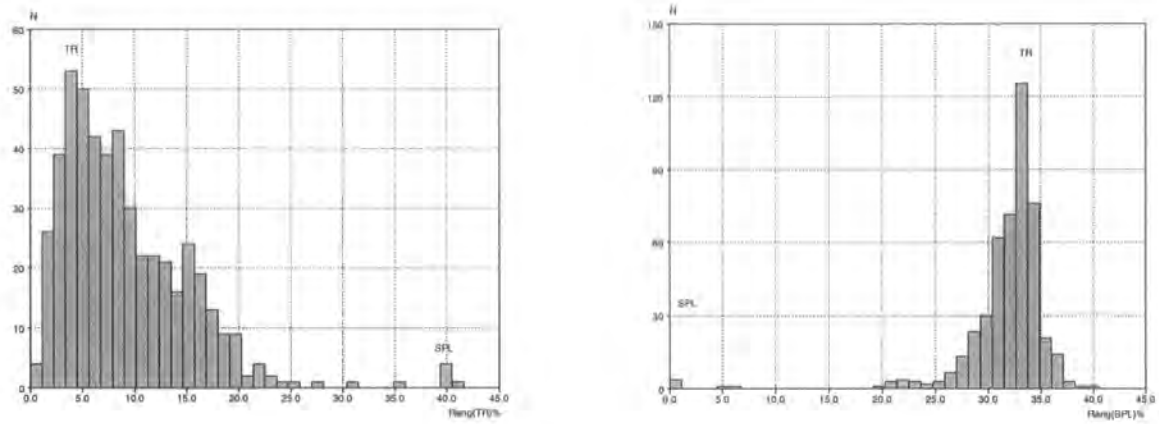


Figure 4.46 2D $R_{ang}(x)\%$ plot for PtL_4 sphere

4-coordination Zn sphere

The element zinc has the outer electronic configuration $d^{10}s^2$ and has the common oxidation state of II, corresponding to the loss of the two *s* electrons, i.e. d^{10} electrons for Zn(II). This d^{10} species, like Cu(I), readily adopts tetrahedral structures in 4-coordinated complexes. The CSD data for Zn complexes shows that out of *499* examples, *493* have perfect or close to perfect tetrahedral geometry. Six exceptions with a square planar coordinated sphere in this species, HAMCON, HAMCUT, HAMGEH (Byrn *et al*, 1993), ZALJIG (Birnbaum *et al*, 1995), PILMOM (Crociani *et al*, 1994) and LAZGAU (Vogel *et al*, 1993) all coordinate with the porphyrin derivatives. The metal ions can fit well the cavity provided by the porhyrin rings and are co-planar with four nitrogen atoms. The two peaks for these tetrahedral structures can be seen in both $R_{ang}(x)\%$ distributions [Figure 4.45(i), (j)] but Figure 4.45(i) illustrates a broader TR-peak which will provide more details in the tetrahedral structures with the distortions.



(i) (j)

Figure 4.45 *contd.* ZnL_4 sphere (i),(j)

All these examples for selected metals in the 4-coordination spheres show a further application of $R_{ang}(x)$ calculations. The 1D representations of $R_{ang}(x)$ provide a quick visualization of coordination sphere geometrical diversity prior to some in-depth analysis. Each representative geometry and distortion forms in these species can be obtained very simply by inspection of the relative frequencies corresponding to the different geometries. The same expression in the $R_{ang}(x)\%$ spectra can obviously be applied to other lower coordination number systems.

4.7 Concluding Remarks

A recent development and progress in crystallography is concerned with crystal engineering (Braga, Grepioni, and Desiraju, 1998), which is to design a new functional solid and predict an unknown molecular structure on the basis of systematic analyses on existing crystallographic structural data. It is indeed an example of knowledge acquisition and accumulation from the available information provided by the vast amounts of accurate crystallographic data. Although the term crystal engineering was originally drawn from the organic solid state photochemical reactions (Rabinovich & Schmidt, 1964), it is now frequently applied to a wide range of attempts to design molecular crystals having various functions. In coordination chemistry, metal

coordination sphere geometry is fundamental to an understanding of physical and chemical properties of complexes.

In the absence of suitable theoretical models to answer such questions as: "what factors determine a given metal coordination geometry?" and "how do ligand conformations and other environmental factors affect the coordination sphere geometry?" Then identifying geometrical preferences of the known coordination structures for the transition metals in a series of complexes will be an important crystal engineering strategy in the design of predictable coordination geometry and specific ligands or the polymeric networks formed by coordination bonds (Keller, 1997; Losier & Zaworotko, 1996). In this study, the acquisition of such knowledge is carried out by a simple but general computational procedure which gives a clear numerical measure of the deviation of a specific crystallographically observed coordination sphere geometry from that of a well-known archetypal polyhedron.

The computer program to apply this method has provided an automatic procedure to find the closest geometrical form for each fragment to a given reference polyhedron. This has resulted from taking full account of 2D atomic permutational symmetry so that all possible coordination isomers are included and so that it is applicable in the case of a generalized ML_n search fragment. The user only needs input the geometry parameters retrieved from the database.

In Chapter 4.2, it was shown that the discrepancy $R_{ang}(x)$ index method could give a clear identification in 7-coordination sphere geometry. The high dimensional parameter space in geometry representation can be captured in a single $R_{ang}(x)$ value. A quick visualization of $1D\text{-}R_{ang}(x)\%$ histogram distribution or $2D\text{-}R_{ang}(x)\%$ plots can easily and rapidly locate the geometrical clusters and thereafter find geometrical preferences and interconversion between the different forms. The further classification of the ligand types on the $2D\text{-}R_{ang}(x)\%$ plot, showed how these different types affect the coordination sphere geometry. This may provide useful information in the design of specific ligands.

In the $R_{ang}(x)$ calculation for seven coordination, each fragment is minimized over 5040 possible ligand permutation and thus a unique atomic enumeration is generated. This forms the basis for a multivariate data matrix and makes it possibile to apply more complex multivariate analyses to this system. Thus, in addition to the discrepancy index $R_{ang}(x)$ method, symmetry coordinates and PCA have been initially applied in 7-coordination sphere geometry for a systematic study. Both methods mapped geometry

for this coordination sphere and from the maps, a proposed interconversion pathway of PBP⇔COC⇔CTP can be observed. Combining the results from the symmetry coordinates and PCA, this geometry change can be readily interpreted.

For the results of 8-coordination sphere geometry in Chapter 4.3, the similar features to those in the 7-coordination case were again shown. The classification of ligand types was considered in more detail for this coordination sphere. It showed that observed structures illustrate the same trends as those based on calculations of ligand bite distances in affecting the formation of a given coordination sphere geometry, although the distance between ligand-ligand or metal-ligand were not taken into account in the $R_{ang}(x)$ calculations to identify the geometry. PCA also appear as a useful tool for detecting the geometrical preferences, deformations and interconversion pathways of metal coordination spheres.

Chapter 4.4 revealed some useful results for 9-coordination sphere geometry. $R_{ang}(x)$ values can be still effective to identify geometry of the coordination sphere which have no great chelating ligand effects. Since much greater number of geometry parameters are accumulated in one-dimensional parameter by using $R_{ang}(x)$ value representation compared with the lower coordination numbers, this reduction in dimensionality, sometimes, can lead to information loss and to overlap of peaks due to different geometries. Extra criteria were used to support further identification of geometry for these fragments. A series of complexes have been examined to see the influences of metals and coordination environments on the coordination sphere geometry. Some representative symmetry coordinates for this coordination sphere in symmetry D_{3h} and C_{4v} have been also successfully applied to map the geometry in this system.

The implementations of the index method for lower coordination number data sets showed its generality as an automatic identification of geometry for an ML_n sphere. The 1D- or 2D plots provide an effective view of the diversity of geometries in these systems. The multi-dimensional descriptors of coordination sphere geometry can be represented in a simple form and from these representations, a wide variety of coordination environments can still be observed. This raises the possibility to use these representations routinely to identify and classify geometrical diversity before going to the more complex multivariate procedures.

The computational procedure and some analysis methods and the results from the systematic study of geometry on various ML_n spheres are expected to contribute as a basis for knowledge engineering in coordination chemistry.

References

Allen, F. H., Bath, P. and Willett, P., *J. Chem. Inf. Comput. Sci.*, 35, 261, (1995).

Andras, M. T., Hepp, A. F., Duraj, S. A., Clark, E. B., Scheiman, D. A., Hehemann, D. G. and Fanwick, P. E., *Inorg. Chem.*, **32**, 4150, (1993).

Asker, K. A., Hitchcock, P. B., Moulding, R. P. and Seddon, K. R., *Inorg. Chem.*, **29**, 4146, (1990).

Auf der Heyde, T. P. E. and Bürgi, H. B., *Inorg. Chem.*, **28**, 3960, (1989a)

Auf der Heyde, T. P. E. and Bürgi, H. B., *Inorg. Chem.*, **28**, 3970, (1989b).

Baker, P. K., Coles, S. J., Durrant, M. C., Harris, S. D., Hughes, D. L., Hursthouse, M. B. and Richards, R. L., *J. Chem. Soc., Dalton Trans.*, 4003, (1996).

Basilevsiky, A., *Statistical Factor Analysis and Related Methods*, John-Wiley & Sons Inc, New York, USA, (1994).

Birnbaum, E. R., Hodge, J. A., Grimstaff, M. W., Schaefer, W. P., Henling, L., Labinger, J. A., Bercaw, J. E. and Gray, H. B., *Inorg. Chem.*, **34**, 3625, (1995).

Bonamico, M., Dessy, G., Fares, V. and Scaramuzza, L., *J. Chem. Soc., Dalton Trans.*, 1258, (1974).

Bligh, S. W. A., Choi, N., Evagorou, E., McPartlin, G., Cummins, W. J. and Kelly, J. D., *Polyhedron*, **11**, 2571, (1992).

Braga, D., Grepioni, F. and Desiraju, G. R., *Chem. Rev.*, **98**, 1375, (1998).

Brower, D. C., Winston, P. B. Tonker, T. L. and Templeton, J. L., *Inorg. Chem.*, **25**, 2883, (1986).

Brown, G. M. and Walker, L. A., *Acta Cryst.*, **20**, 220, (1966).

Byrn,M. P., Curtis, C. J., Hsiou, Y., Khan, S. I., Sawin, P. A., Tendick, S. K., Terzis, A. and Strouse, C. E., *J. Am. Chem. Soc.*, **115**, 9480, (1993).

Cambridge Structural Database User Manuals, Crystallographic Data Centre, Cambridge, (1994).

Chun, H. K., Steffen, W. L. and Fay, R. C., *Inorg. Chem.*, **18**, 2458, (1979).

Corden, B. J., Cunningham, J. A. and Eisenberg, R., *Inorg. Chem.*, **9**, 356, (1970).

Cotton, F. A., Diebold, M. P. and Roth, W. J., *Polyhedron*, **4**, 1103, (1985).

Cotton, F. A., Duraj, S. A. and Roth, W. J., *Inorg. Chem.*, **23**, 4046, (1984).

Cotton, F. A. and Wilkinson, G., *Advanced Inorganic Chemistry, 5th Edition*, John Wiley & Sons, New York, USA, (1988).

Crociani, B., Di Bianca, F., Fontana, A., Forsellini, E. and Bombieri, G., *J. Chem. Soc., Dalton, Trans.*, 407, (1994).

Del Piero, G. Perego, G., Zazzetta, A. and Brandi, G., *Cryst. Struct. Commun.*, **4**, 521, (1975)

Drew, M. G. B. and Wilkins, J. D., *J. Chem. Soc., Dalton Trans.*, 198, (1974).

Drew, M. G. B. and Wilkins, J. D., *J. Chem. Soc., Dalton Trans.*, 1654, (1974)

Ellis, J. E. and Yuen, P., *Inorg. Chem.*, **32**, 4998, (1993).

Faithfull, D. L., Harrowfield, J. M., Ogden, M. I., Skelton, B. W., Third, K. and White, A. H., *Aust. J. Chem.*, **45**, 583, (1992).

Favas, M. C. and Kepert, D. L., *Prog. Inorg. Chem.*, **28**, 309, (1981).

Finnen, D. C., Pinkerton, A. A., Dunham, W. R., Sands, R. H. and Jr. Funk, M. O., *Inorg. Chem.*, **30**, 3960, (1991).

Forsellini, E., Benetollo, F., Bombieri, G., Cassol, A. and De Paoli, G., *Inorg. Chim. Acta*, **109**, 167, (1985).

Gao, Y., Guery, J. and Jacoboni, C., *Acta Cryst.*, C**49**, 963, (1993).

Gerkin, R. E. and Reppart, W. J., *Acta Cryst.*, C**40**, 781, (1984).

Greenwood, N. N. and Earnshaw, A., *Chemistry of the Elements*, Pergamon Press, Oxford, UK, (1994)

Gregson, D., Mason, S. A., Howard, J. A. K., Spencer, J. L. and Turner, D. G., *Inorg. Chem.*, **23**, 4103, (1984).

Gillespie, R., *J. Chem. Soc. Rev.*, 59, (1992).

Gillespie, R. and Hargitai, I., *The VSEPR Model of Molecular Geometry*, Allyn & Bacon, Massachusetts, USA, (1991).

Guggenberger, L. J. and Muetterties, E. L., *J. Am. Chem. Soc.*, **98**, 7221, (1976).

Harrowfield, J. McB., Kepert, D. L., Patrick, J. M. and White, A. H., *Aust. J. Chem.*, **36**, 483, (1983).

Hauck, J., *Inorg. Nucl. Chem. Lett.*, **12**, 617, (1976).

Hoard, J. L., Hamor, T. A. and Glick, M. D., *J. Am. Chem. Soc.*, **90**, 3177, (1968).

Hoffmann, R., Beier, B. F., Muetterties, E. L. and Rossi, A. R., *Inorg. Chem.*, **16**, 511, (1977).

Howard, J. A. K., Copley, R. C. B., Yao, J.W. and Allen, F. H., *Chem Comm.*, 2175, (1998)

Howard, J. A. K., Mead, K. A. and Spencer, J. L., *Acta Cryst.* C**39**, 555, (1983).

Keller, S. W., *Angew. Chem., Int. Ed. Engl.*, **36**, 247, (1997).

Kepert, D. L., *Prog. Inorg. Chem.*, **23**, 1, (1977).

Kepert, D. L., *Prog. Inorg. Chem.*, **24**, 179, (1978).

Kepert, D. L., *Prog. Inorg. Chem.*, **25**, 41, (1979).

Kepert, D. L., in *Comprehensive Coordination Chemistry*, Eds. G. Wilkinson, R. D. Gillard and J. McCleverty, **1**, 31, Pergamon Press, Oxford, UK, (1987).

Kepert, D. L., Patrick, J. M. and White, A. H., *J. Chem. Soc., Dalton Trans.*, 385, (1983).

Klebe, G., *J. Mol. Struc. (Theochem.)*, **308**, 53, (1994).

Klebe, G. and Weber, F., *Acta Cryst.* B**50**, 50, (1994).

Knox, K. and Ginsberg, A. P., *Inorg. Chem.*, **3**, 555, (1964).

Koch, O., Edelmam, F. and Behrens, U., *Chem. Ber.*, **115**, 1313, (1982).

Kojic-Prodic, K., Ruzic-Toros, Z. and Sljukic, M., *Acta Cryst.*, B**34**, 2001, (1978)

Leipoldt, J. G., Basson, S. S. and Bok, L. D. C., *Inorg. Chim. Acta* , **44**, L99, (1980)

Lewis, D. F. and Fay, R. C., *Inorg. Chem.*, **15**, 2219, (1976).

Lind, M. D., Hannor, M. J. and Hoard, J. L., *Inorg. Chem.*, **3**, 34, (1964).

Ling, S. S. M., Jobe, i. R., McLennan, A. J., Manojlovic-Muir, L., Muir, K. W. and Puddephatt, R. J., *J. Chem. Soc., Chem. Comm.*, 566, (1985).

Losier, P. and Zaworotko, M. J., *Angew. Chem., Int. Ed. Engl.*, **35**, 2779, (1996).

Mackay, K. M. and Mackay, R. A., *Introduction to Modern Inorganic Chemistry*, Int. Textbook Com., London, (1981).

Miguel, D., Perez-Martinez, J. A., Riera, V. and Garcia-Granda, S., *Polyhedron*, **10**, 1717, (1991).

Morse, P. M. and Girolami, G. S., *J. Am. Chem. Soc.*, **111**, 4114, (1989).

Muetterties, E. L and Guggenberger, L. J., *J. Am. Chem. Soc.*, **96**, 1748, (1974).

Murray-Rust, P., *Acta Cryst.*, B**38**, 2765, (1982).

Murray-Rust, P., Bürgi, H. B. and Dunitz, J. D., *Acta Cryst.*, B**34**, 1787, (1978).

Murray-Rust, P., Bürgi, H. B. and Dunitz, J. D., *Acta Cryst.*, A**35**, 703, (1979).

NAG Fortran Library Introductory Guide, Mark 18, The Numerical Algorithms Ltd., Oxford, UK, (1997).

Ohki, Y., Suzuki, Y. and Ouchi, A., *Bull. Chem. Soc. Jpn*, **60**, 1543, (1987).

Ohki, Y., Suzuki, M., Nakamura, M., Shimoi, M. and Ouchi, A., *Bull. Chem. Soc. Jpn.*, **58**, 2968, (1985)

Ohki, Y., Suzuki, M., Takeuchi, T., and Ouchi, A., *Bull. Chem. Soc. Jpn.*, **61**, 393, (1988).

Paige, C. R. and Richardson, M. F., *Can. J. Chem.* **62**, 332, (1984).

Palenik, G. J., Wester, D. W., Rychlewska, U. and Palenik, R. C., *Inorg. Chem.*, **15**, 1814, (1976).

Psillakis, E., Jeffery, J. C., Mcdeverty, J. A. and Ward, M. D., *Chem. Commun.*, 1965, (1997).

Rabinovich, D. and Schmidt, G. M. J., *J. Chem. Soc.*, (B), 144, (1964).

Robertson, B. E., *Inorg. Chem.*, **16**, 2735, (1977).

Rogers, R. D., Etzenhouser, R. D., Murdoch, J. S. and Reyes, E., *Inorg. Chem.*, **30**, 1445, (1991).

Rogers, R. D., Rollins, A. N., Etzenhouser, R. D., Voss, E. J. and Bauer, C. B., *Inorg. Chem.*, **32**, 3451, (1993).

Roland, E., Fischer, K. and Vahrenkump, H., *Angew. Chem., Int. Ed. Engl.*, **22,** 326, (1983).

Shellmann, D., Keller, J., Moll, M., Campana, C. F. and Haase, M., *Inorg. Chim. Acta*, **141**, 243, (1988).

Shiu, K-B., Lee, J.Y., Wang, Y. and Cheng, M-C., *Inorg. Chem.*, **32**, 3565, (1993).

Shiu, K-B., Yih, K-H., Wang, S-L. and Liao, F-L., *J. Organomet., Chem.*, **414**, 165, (1991).

Speer, S. T., Perumaredi, J. R. and Adamson, A. W., *J. Am. Chem. Soc.*, **90**, 6626, (1968).

Starynowicz, P., *Acta Cryst.*, C**47**, 2063, (1991).

Starynowicz, P., *Acta Cryst.*, C**48**, 1414, (1992).

Tomi, F., Wah, H. K. W. and Postel, M., *New, J. Chem.*, **12**, 289, (1988).

Taylor, R. and Allen, F., in *Structure Correlation,* Eds., H.-B. Bürgi and J. D. Dunitz, VCH, Weinheim, **1,** 111, (1994).

Vogel, E., Koch, P., Hou, X-L., Lex, J., Lausmann, M., Kisters, M., Aukauloo, M. A., Richard, P. and Guilard, R., *Angew. Chem., Int. Ed. Engl.*, **32**, 1600, (1993).

Xue, W., Zhu, Y. and Yang, R., *J. Coord. Chem.*, **26**,199, (1992).

Zych, E., Starynowicz, P., List, T. and Drozdzynski, J., *Polyhedron*, **12**, 1661, (1993).

Chapter 5

A Database Study of Transition Metal Alkyne and Alkene Complexes

5.1 Introduction

Metal-alkene and -alkyne bonding in complexes was firstly characterized in terms of molecular orbital theory by Chatt *et al* (1953). Since then the wide applications of these kinds of complexes in organic synthesis, studies of various chemical properties and reactions of the complex have kept interest high. In structural chemistry, a lot of structures of complexes were characterized during the 1970s and 1980s(Mingos, 1982; Bruce, 1995). For example, crystal structure determinations of some platinum alkene complexes by X-ray and neutron diffractions (Love, Koetzle, Williams, Andrews and Bau, 1975; Howard, Spencer and Mason, 1983; Howard, Mitrprachon and Roy, 1982). Some theoretical calculations were applied to the specific bonding system (Frenking and Pidum, 1997; Albright, Hoffmann, Thibeault and Thorn, 1979) so that the bonding theory on this system was further developed. Based on structural knowledge, systematic comparisons on the geometrical characteristics of selected complexes (Ittel and Ibers, 1982; Templeton, 1989) provided clear outlines of the interaction of unsaturated molecules with transition metals. However, because these studies were based on a manual literature search, the range of compounds was limited. The CSD also records large amounts of structural data on transition metal-alkene and -alkyne complexes. It will provide a wider basis for systematically investigating the geometrical properties of such complexes with various transition metals and double and triple bond derivatives. The work described includes a systematic analysis of the geometry of these kinds of complexes by firstly using the Cambridge Structural Database.

The metal-alkene and -alkyne bond is usually described as involving a σ-bond, donated from π-electron of the double or triple bond to an empty hybrid orbital on the metal and a π-bond formed from back donation from a filled hybrid orbital on the metal into the empty C=C π^*-orbital.

For this special bonding form, the observed structural features on the geometry of the complexed alkene or alkyne is that the substituents on C=C or C≡C are bent back from the C-sp^2 plane in alkene and the C-sp^1 line in alkyne. So the factors of interest are to:

(1) address the structural features by finding correlations between specific structural parameters.

(2) find the implication of the nature of the substituents on the bent-back angle.

(3) investigate the interactions between the metals and C=C bond.

(4) from all the correlations, if possible, derive a predictable variation for the complexes in various levels of bonding in C≡C, C=C and C-C, i.e.

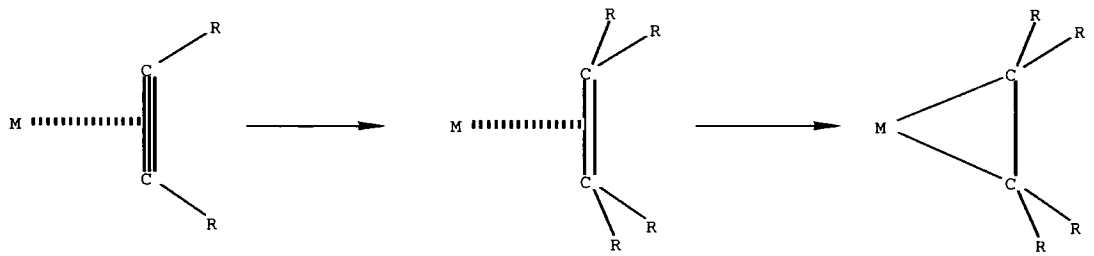


Figure 5.1 The changes in bond order from triple to single bond in metal π-bond complexes

Before further results from this study are given, a few individual structures of metal alkene complexes will be reviewed to show the basic structural characteristics of this kind of complex and therefore to understand the aims of this study.

Zeise's salt, $K[PtCl_3C_2H_4]$ is one of the earliest published alkene complexes and the structure of anion has been examined by X-ray and neutron diffraction (Eller, Ryan and Schaeffer, 1977). According to the σ-π bond scheme, the bonding orbitals are shown in Figure 5.2. An accurate quantum chemical calculation (Rösch, Messmer and Johnson, 1974) showed that major part the total bonding energy in Zeise's salt is contributed from the forward donation component (75%) and 25% from the back donation component. The X-ray structure gave that both M-C distances are equal but the C-C bond length [*1.375(3)*Å] is longer than the *1.337(2)*Å of ethylene in the free state in gas phase. The hydrogen atoms in the ethylene are bent back from the original planar

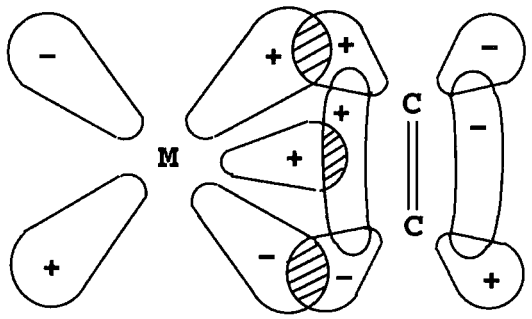


Figure 5.2 The orbitals involved in σ-π bond in alkene complex

alkene. The C=C bond is perpendicular to the $PtCl_3$ plane, that is, the ethylene is bound perpendicularly to the square-plane, so that the σ-π bond has the maximum orbital overlaps and relevant definite conformational preference is displayed as in Figure 5.3. However, the geometrical constraints imposed by the ligand can sometimes force the alkene to adopt the alternative conformation (Schilling, Hoffmann and Lichtenberger, 1979), i.e. the alkene is parallel to the plane formed by other three ligands and the metal atom.

In some alkene complexes, strong metal-ligand interactions dominate the coordination bond while ligand to metal donation is weak. The best representation of the structure is as a metallocyclopropane rather than an alkene complex, i.e. shown as Figure 5.3.

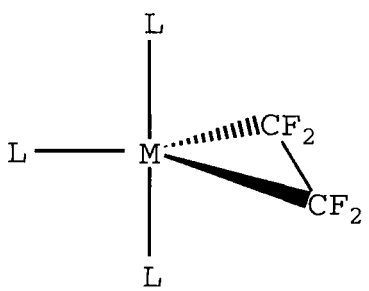


Figure 5.3 Structural representation as a metallocyclopropane

In this case, the C-C bond length is lengthened to almost the same as in ethane. This is equal to a re-hybridization of the alkene C atoms from sp^2 to sp^3. In fact, there are a lot of structures whose C-C bond lengths occur between these two extreme cases.

Similar structural characteristics to metal-alkene complexes can be also observed in the metal-alkyne complexes. They are often classified as two-electron or four-electron donation according to the orientation of two sets of perpendicular $p\pi$-orbitals of the alkyne (Templeton, 1989). The difference is that if the second π^*-orbital, which is perpendicular to the first pair of π-orbital, interacts with additional metal *d* orbitals, additional pair of electrons can be back to this orbital so that the metal-alkyne bond could be strengthened significantly compared with that of a metal-alkene bond.

In consequence, the formation of metal σ-π bond depletes the order of double and triple bonds so that the geometries of alkene and alkyne have been changed from the original planar and linear states. The levels of such change obviously depend upon a number of factors, firstly, the donation from metal *d*-electron to the empty π^* orbital of ligand. Secondly the presence of electron withdrawing substituents on the C atoms of alkene or alkyne will increase the contribution, while the presence of electron donating substituents will increase the σ-bond from π-electrons of C=C or C≡C bonds to empty metal orbitals. Thirdly, the nature of the metal, such as its oxidation state and the presence of other ligands will also affect the bonding. The formation of a metal σ-π bond also increases the δ^+ charges on the C atoms so that metal-alkene complexes are often used in nucleophilic addition not in electrophilic addition as in the free alkene state. This reactivity is clearly affected by the bonding in alkene complexes. Nucleophilic attack is enhanced by an increase in σ donation, while an increase in π back donation will enhance electrophilic attack. But this only gives a trend from nucleophilic to electrophilic attack. In practice, they are still in a region of electron density which nucleophilic attack may occur even if there is more π-bonding than σ-bonding (Hartley, 1969).

The model given in Figure 5.2 has been applied for decades to explain a number of experimental observations as stated above. However, this model is not proper to quantify the distortion of the ligand in the complex or to predict details of the difference bonding properties of different ligands (Pidum and Frenking, 1995). So the modern *ab initio* quantum chemical methods have been used in some specific complex species to

validate the model (Nielson, Bayd, Clark, Hunt, Melson, Richard and Schwerdtfeger, 1992; Widmark, Roos and Sirgbahu, 1985).

From these studies, it has been shown that metal-ligand bond lengths, C-C (simplified C=C or C≡C hereafter) bond lengths in ligands, ligand orientations and selected angles can provide useful information in revealing important geometric characteristics of these complexes. In the following sections, how these geometrical parameters are defined in the CSD search to meet the requirements of this study will be given. Correlation relationships derived from these definitions based on a large number of structural data show some meaningful results.

5.2 Database Retrieval

In attempting to examine the molecular geometry of metal π-bond complexes formed by alkene and alkyne, first, an appropriate model fragment must be defined so that suitable geometrical parameters for further analysis can be retrieved from the database. The fragments to be searched in this study are given according to their basic bond connections and shown in Figure 5.4 (a) metal-alkyne; (b) metal-alkene.
Fragments consist of two parts of a general transition metal (TR) and alkyne or alkene π-donation ligand. The interaction between metal atom and ligand is represented by the distances B_2 and B_3. The bond type options `1,9` in the parentheses indicate that it is a single or π bond in CSD definitions. The C-C bond length is defined as B_1 which is a double or triple bond (bond type `2,3`) in metal-alkyne complexes and a single or double bond (bond type `1,2`) in metal-alkene complexes, respectively. T_3 and T_4 are used here to specify that carbon atom in alkyne and alkene have a total coordination number *3* and *4*, respectively, based on sp^1-hybridization in alkyne and sp^2-hybridization in alkene. R's are substituent groups on ligands, which include terminal H atoms and all *p*-block elements.

The definition of the bent back angle, θ_4 and θ_5, is easier in metal-alkyne complexes than in metal-alkene complexes. They can be found directly from the definition of angle, R-C≡C because there is only one substituent on the C-atom. But in alkene complexes, there are two substituents on the C-atom, the angle, in fact, reflects the average deviation of these two substituents from the C-C sp^2 plane. There are several different ways to describe the displacements. The best way here is to choose the

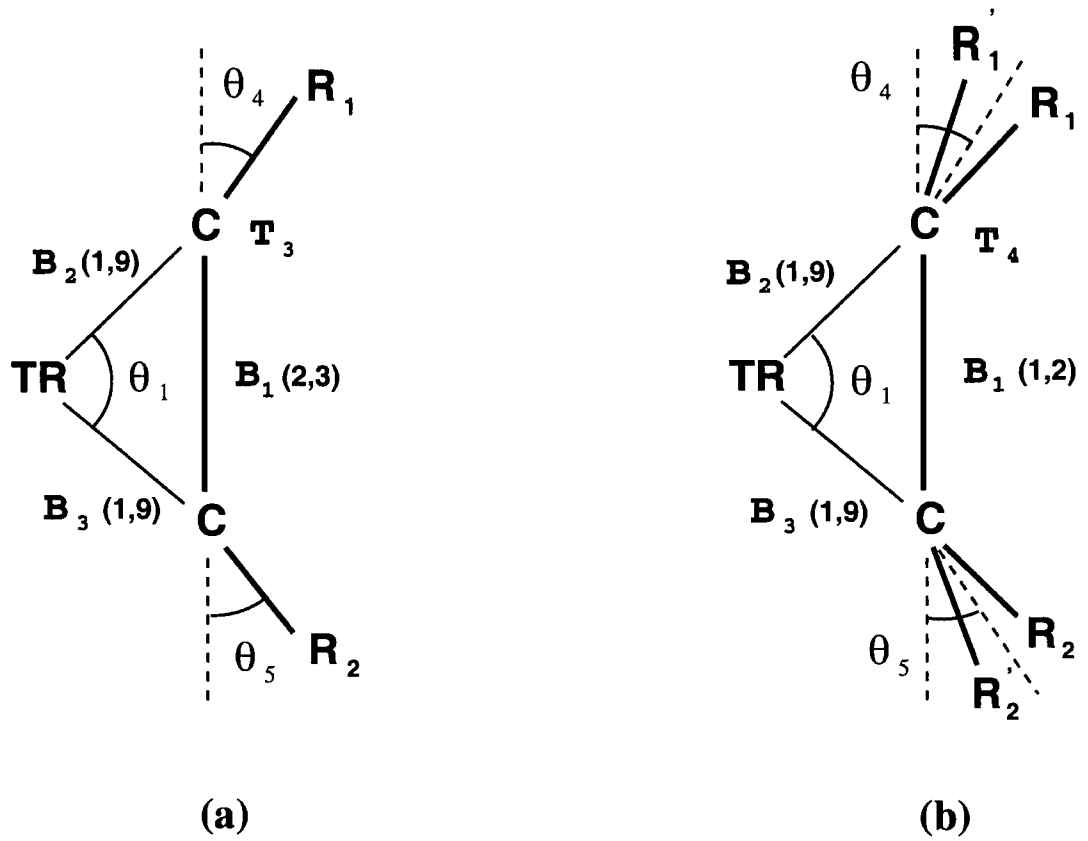


Figure 5.4 Fragments defined in the CSD search (a) metal-alkyne (b) metal-alkene

coordinates which are derived from the four torsion angles relating the central double bond and the substituents at each C-atom, which was used to describe the secondary amide group (Dunitz, 1979). In this definition, the four torsion angles about the C--C bond are shown in Figure 5.5 and θ_4 and θ_5 are defined as:

$$\theta_4 = \omega_1 - \omega_3 + \pi = -\omega_2 + \omega_4 + \pi(\mathrm{mod}\,2\pi)$$

$$\theta_5 = \omega_2 - \omega_3 + \pi = -\omega_1 + \omega_4 + \pi(\mathrm{mod}\,2\pi)$$

where ω_1, ω_2, ω_3 and ω_4 are four torsion angles, *R2--C--C--R1*, *R2'-C--C--R1'*, *R2'--C--C--R1* and *R2--C--C--R1'*, respectively.

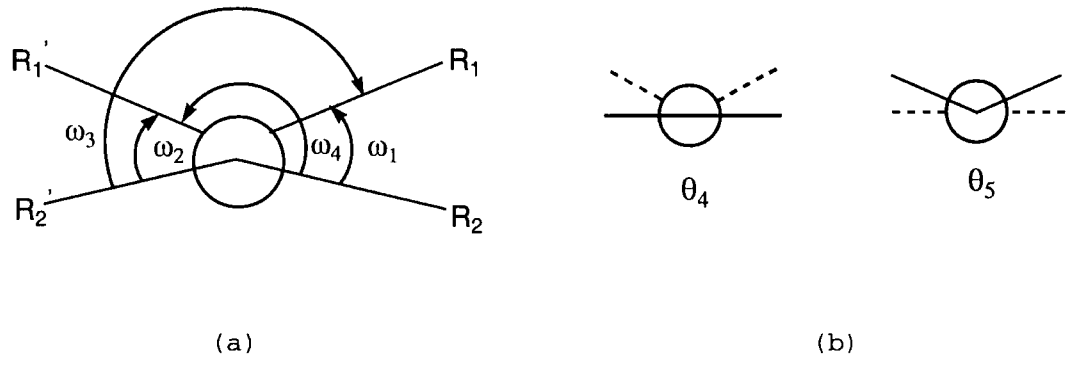


Figure 5.5 Diagram representations of four torsion angles $\omega_1, \omega_2, \omega_3$ and ω_4 (a) and (b) the bend angles θ_4 and θ_5 at two the C-atoms (Dunitz, 1979).

Version 5.10 of the CSD (1995) was used to perform the fragment search, which recorded all required molecules and confirmed the geometrical parameters matched exactly with the definitions of substructures. General search restrictions ensured that for all retrieved entries, there were:

(a) no residual numerical errors;
(b) no disorder in the chemical structure;
(c) no R values greater than 10%;
(d) atomic coordinates available;
(e) no polymer contained and
(f) "organometallic" compounds according to the CSD chemical class assignments.

In addition, there were a lot of conjugated dialkenes such as butadiene complexes resulting from the search of metal-alkene complexes, attention was restricted to complexes of monoalkenes. Conjugated dialkenes undergo substantial changes in bond lengths on coordination to a metal atom and therefore their complexes were not included in this study. Complexes with large steric effects that directly dominate the bent back angle, such as some cyclo-alkene complexes, were also excluded from the analysis. Finally, the CSD revealed 424 metal-alkyne and 373 metal-alkene complexes in an exactly defined environment. The numbers for different metals in both types of complexes are listed in Table 5.1.

Table 5.1 The Number of Metal-Alkyne and -Alkene Complexes for Different Metals Retrieved from the CSD

M	Ti	V	Cr	Mn	Fe	Co	Ni	Cu
Alkyne	7	8	7	2	8	6	33	28
Alkene	1	3	4	6	28	11	57	8
M	**Zr**	**Nb**	**Tc**	**Mo**	**Ru**	**Rh**	**Pd**	**Ag**
Alkyne	8	27	0	50	6	4	2	0
Alkene	0	2	0	28	23	68	15	0
M	**Hf**	**Ta**	**W**	**Re**	**Os**	**Ir**	**Pt**	**Au**
Alkyne	0	17	128	39	6	7	31	0
Alkene	2	5	17	14	7	14	56	0

In general, it is considered that metals with low oxidation states will enhance the formation of π-bond complexes because more electrons can be back donated. Thus metals on the right of the transition block tend to form more stable π-bond complexes than those on the left. But metals also need to provide an empty orbital with comparable energies to accept π-electrons from ligands, therefore, this kind complex is difficult to find for the metals in group 11 and 12.

In the alkene complexes dataset, most compounds are pure hydrocarbon alkene which have no substituent groups in the neighborhood of the C=C double bond. There are also many complexes (115) which have one substituent on the carbon atom. But the number of complexes in which both hydrogens on the carbon atoms are substituted falls to *21*. Whereas in alkyne complexes dataset, only eight complexes have pure hydrocarbon alkynes, most others contain functional group substituents.

5.3 Data Analysis

From both of the data sets for metal-alkene and alkyne, statistical results and correlation relationships can be obtained by using the VISTA 2.0 program provided by Cambridge Crystallographic Structural Database. Some new definitions related to the parameters retrieved from QUEST3D could be generated by the command CREATE in VISTA. The statistics for these parameters are summarized in the following sections.

5.3.1 Bend Back Angles and C-C Bond Lengths

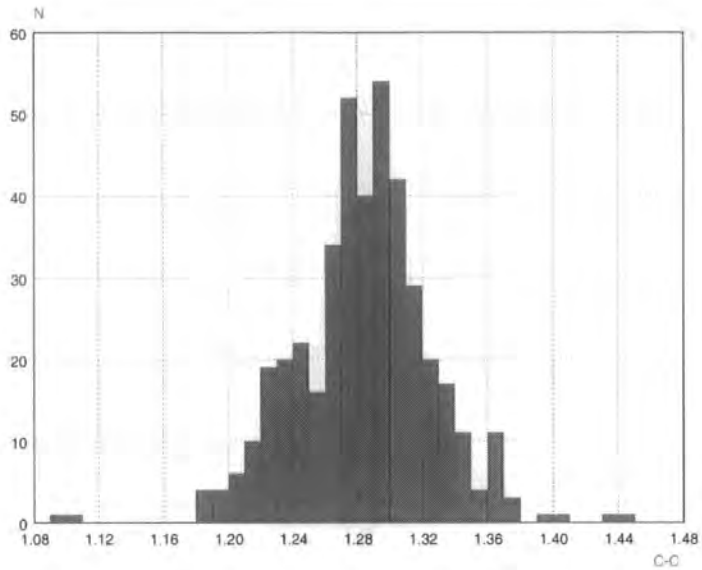
An important geometrical feature in these complexes is the reordering of alkyne and alkene, which is reflected in the bend back angle and the change in C-C bond lengths. An examination of the data sets reveals that most bond lengths of coordinated alkynes or alkenes are significantly longer than that of a free acetylene [1.181(2)Å] or ethylene [1.337(2)Å]. Moreover, with different substituents on unsaturated carbon atoms and different other ligands coordinated with the metal, the C-C bond lengths are also different. There are some significant variations in C-C distances. The distributions of these variations in alkyne and alkene complexes are shown in Figure 5.6 (a) and (b), respectively. The average C-C distances in the two datasets are 1.284(2)Å (alkyne) and 1.403(3)Å (alkene).

The distribution of values of the bend back angles for both complexes are listed in Table 5.2, which shows that the range of this angle for metal-alkyne complexes falls into *13.56°* to *52.71°*, mean value *36.81°* and most complexes have bend back angle between *32.0°* and *45.0°*. The metal-alkene complexes showed a wider distribution range, from *0.0°* to *67.09°*, for this angle and the mean value *24.78°*. The value of *0.0°* for the bent-back angle for this data set is contributed from those without substituted alkene molecules. X-ray structural determination may place the hydrogen atoms for these complexes at theoretical positions. Such positions only represent the normal positions in the free alkene state and so the bend back angle must be ignored. In consequence, this resulted in a large peak near zero degrees of bend back angle. However, there are still a lot of the hetero-atom substituented alkene metal complexes in the data set. The bend back angles can be derived from these determined structures. Overall, inspection of the distributions reveals that bending back has minimum effect for hydrogen atoms both in alkyne and alkene, also for phenyl groups alkyne (e.g.

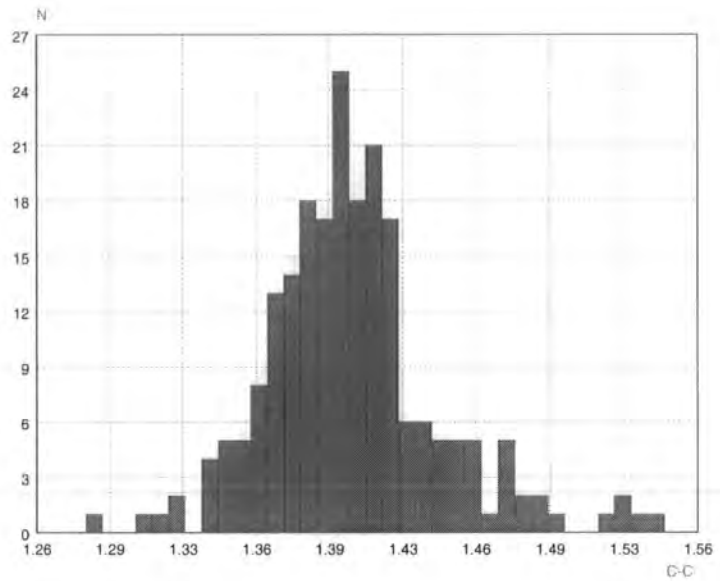
Table 5.2 Distribution values of bend back angle(°)

Metal complex	Mean Value	Min. Value	Max. Value	SD	N_f
Alkyne	36.85	13.56	52.71	0.36	424
Alkene	24.78	0.00	67.01	1.03	373

SD --- Standard Deviation
N_f --- Number of fragments.



(a)



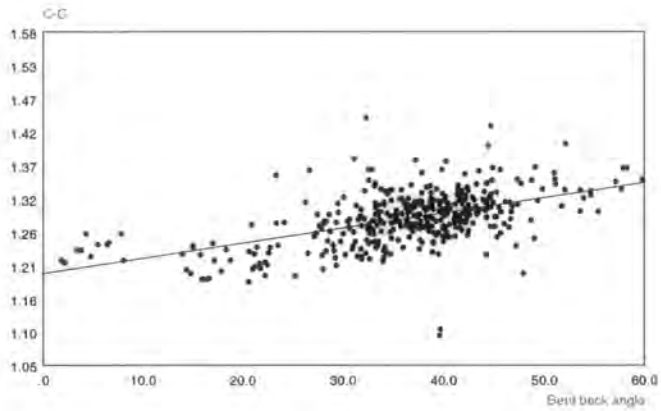
(b)

Figure 5.6 Distribution of C≡C bond length in alkyne (a) and C=C bond length in alkene complex (b)

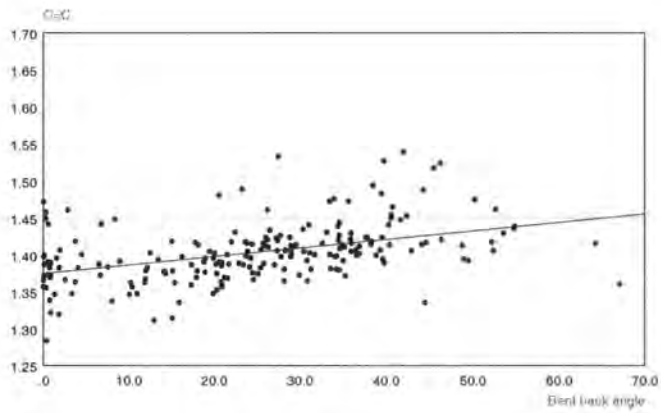
FAZFAN) which tries to keep planar conjugated stability. Values are greatest for halogen substituents, especially fluorine (e.g. SIDYUZ) in alkene and tertiary-butyl-ammonium groups (e.g. BEFKUS) in alkyne, the latter is a bulkier substituent.

These two distributions show that larger average bend back angle occurs in metal-alkyne complexes, but metal-alkene complexes have wider distribution range.

The correlation of C≡C bond length in alkyne complex and C=C bond length in alkene complex versus the related bend back angle are given in Figure 5.7 (a) and (b), respectively.



(a)



(b)

Figure 5.7 Correlations of C≡C bond length (a) and C=C bond length (b) *vs.* bent back angle θ in metal-alkyne and metal-alkene complexes

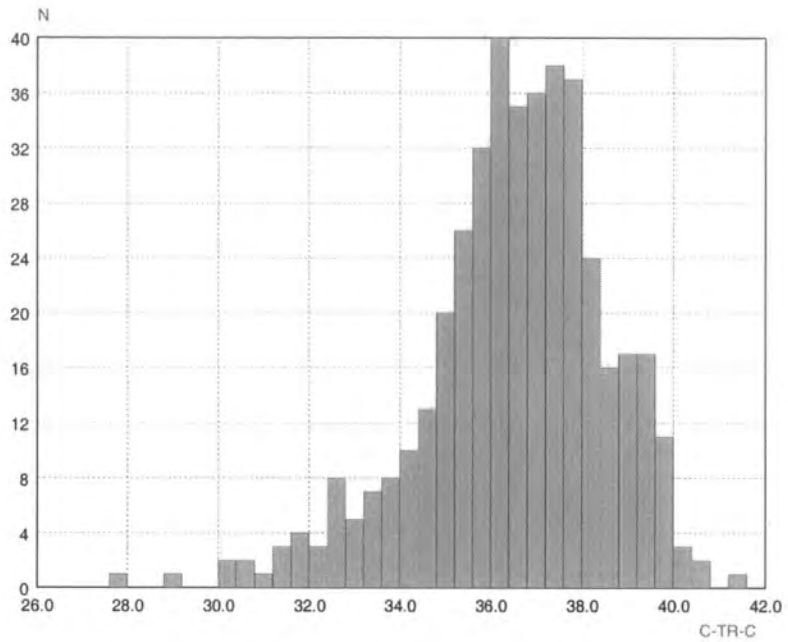
In principle, the changes of both bend back angle and C-C bond length will reflect the level of re-organization in C≡C bond or C=C bond order when they coordinate with metal atom. The larger these values, the greater the influence of the coordination π-bond. Furthermore, the variations of C-C bond are proportional to that of the bend back angle. The relationships in Figure 5.7 (a) and (b) are consistent with the theoretical prediction. It is observed that the C-C distance relates approximately linearly to the bend back angle (correlation coefficients for alkyne complexes 0.68; for alkene 0.57), that is, the C-C bond lengths increase with the larger bend back angles. In the case of alkyne, this relationship is more obvious than that in alkene, as shown by larger correlation coefficient. Therefore it can be said that there exists a stronger interaction between metal and C≡C bonds compared to C=C bonds.

5.3.2 Interactions Between Metal and C-C Bond

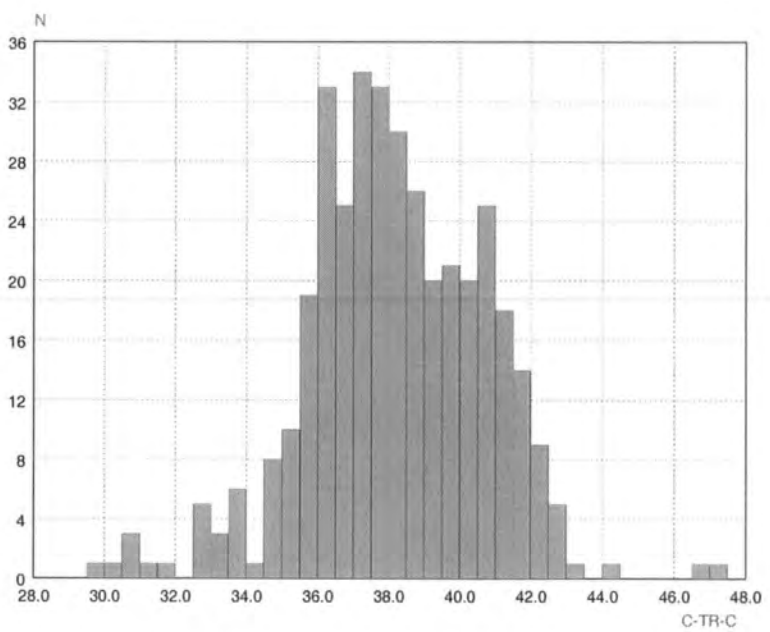
Firstly, the interactions between metal atom (TR) and the C-C bond in these kinds of complexes were examined by the basic geometrical parameters, angle C-TR-C and distances TR-C. In both alkyne and alkene complexes, there are common features for these angles and distances. Figure 5.8 (a) and (b) give distributions of angle C-TR-C as histograms in alkyne and alkene complexes, respectively. The distribution for alkyne is nearly symmetrical around 36.5°, and 38.2° for alkene.

The distances between metal and two carbon atoms are mostly equal and have a direct linear relationship, shown in Figure 5.9 (a) and (b). Two factors may affect this relationship, deviations could occur in asymmetrical substituents, such as $Cl_2C{=}CCN_2$ (CCYEPT, 2.096, 2.003Å), or substituents with strain (YAWCII, 2.330, 2.575Å) so that the two bond lengths show a significant difference.

Nevertheless, these comparisons are based on the observed interatomic distances. Although the ligands are the same, i.e. alkyne or alkene, the metal atoms may be different for each fragment. Thus the bonds are not of the same scale, for example, for a metal atom with larger atomic radius it normally has a larger observed value. This will have a considerable impact on the comparison of these values for different kinds of metal atoms. For this reason, a standard or a scale should be considered so that the systematical analysis can be carried out. In order to put all interatomic distances between metal and carbon atoms on a common scale, a normalized distance [DMC, in



(a)



(b)

Figure 5.8 Distribution of C-TR-C angles in metal-alkyne (a) and metal-alkene (b) complexes

Equation (5.1)] is generated by subtracting the covalent radius of metal atom from the observed bond length. Therefore a "net" effect of interaction between metals and ligands can be investigated.

$$DMC = \frac{d_{M-C} - r_M}{r_M} \qquad (5.1)$$

where d_{M-C} is an average value of two M--C distances and r_M is the standard value for covalent radius of metal atom, which is the same as those given in the CSD system (CSD, 1995).

The results of comparisons on the normalized distances of different metals with the carbon atom are given in Table 5.3, which shows a main range of DMC values for different metals.

Table 5.3 DMC values for different metals

Metal	Alkyne	Alkene
Ni	0.22 — 0.28	0.30 — 0.39
Cu	0.27 — 0.34	0.31 — 0.38
Pt	0.34 — 0.53	0.33 — 0.51
Pd	*	0.36 — 0.48
Mo	0.34 — 0.48	0.50 — 0.72
Nb	0.37 — 0.54	*
Co	0.39 — 0.51	0.48 — 0.59
Rh	0.40 — 0.44	0.40 — 0.60
Ta	0.41 — 0.49	0.54 — 0.69
Cr	0.43 — 0.45	0.60 — 0.62
V	0.43 — 0.57	0.63 — 0.66
W	0.44 — 0.61	0.56 — 0.79
Re	0.46 — 0.63	0.62 — 0.75
Fe	0.47 — 0.60	0.54 — 0.68
Os	0.48 — 0.68	0.58 — 0.66
Ir	0.53 — 0.60	0.57 — 0.65

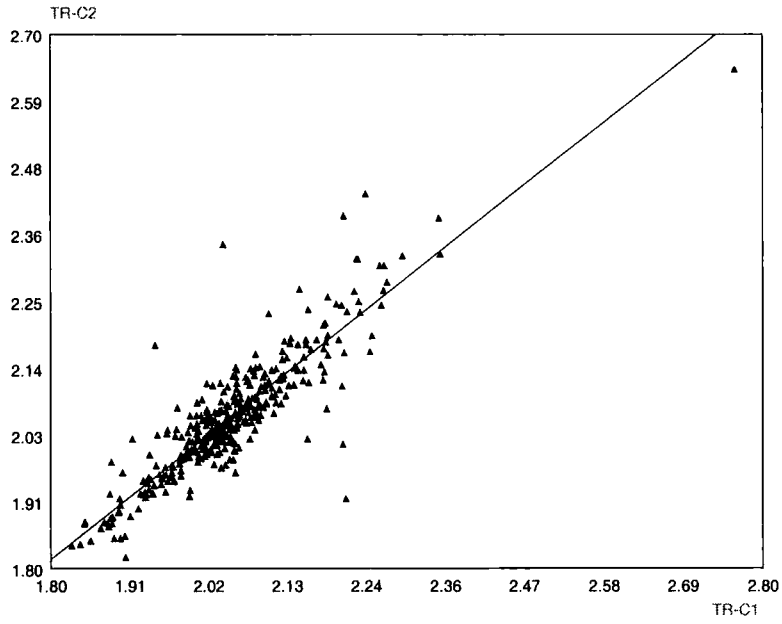
* - Not enough structural data

It can be seen that metals in group 10, Ni, Pt, Pd and Group 11 metal Cu have the shortest distances for both kinds of complexes, which imply that a stronger interaction exists within these metal complexes. From the π back donation theory stated above, it is easy to explain this result. These metals are all rich in dπ-electrons, d^8- d^{10} and Cu-alkene complexes always appear as a single charged cation. The metals not only provide the energetically compatible empty orbitals but also have more chance for back donation of *d*-electrons to ligands. The π-bond plays a more important part in these complexes. They often form the complexes in oxidation states Ni(II), Pt(II), Pd(II), Pt(0), Pd(0) and Cu(I).

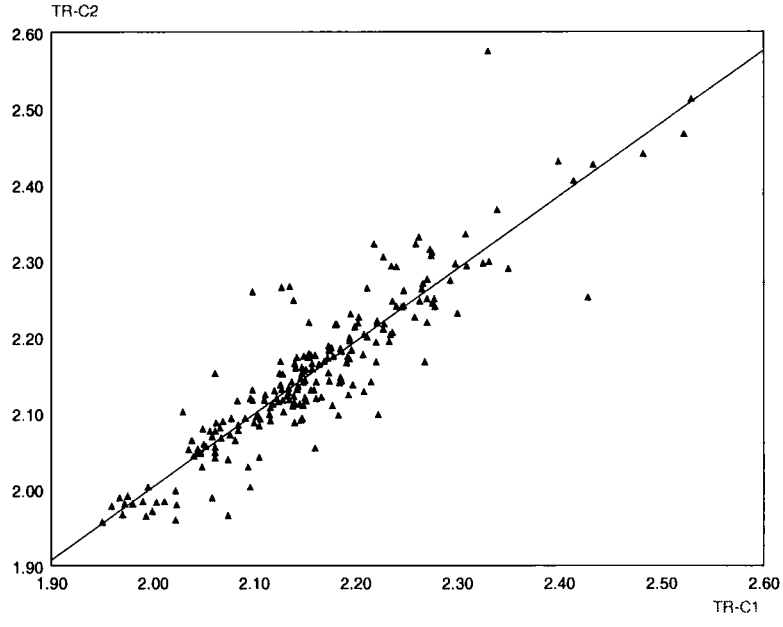
It was considered (Chetcuti, 1995) that nickel is most preferential for simple π-mono-olefin complexes. In the data set of this study the shortest DMC value is always for nickel complexes; platinum(II) can form a very large range of stable mono-alkene or mono-alkyne complexes which have a very similar range of DMC values. Especially due to their tremendous variety and their relative ease of formation in alkene complexes, they have come to be widely regarded as the characteristic representatives of transition metal π-bond complexes in their reactivity and their bonding behaviour. In the data set, the number of palladium alkyne complexes is much less than other two in group 10. In this study, only mono- alkene or alkyne are considered, while alkynes are very rapidly polymerized by Pd(II) (Hartley, 1969) and palladium (II) also oxidizes alkynes to carbonyl compounds. So palladium-alkyne complexes always exist in multi-alkyne forms. In this data set only two mono- complexes were "hit". In the case of alkene, dimeric halogen-bridged complexes are often formed.

For the remaining metals in Table 5.3, the DMC values are more variable than those of metals in group 6. The statistical results listed in Table 5.3 represent ranges that most complexes fall in, with some overlaps. However, a trend can be seen from these results in that most metals have smaller DMC values in alkyne than in alkene complexes and they are distributed in two separate ranges. Again this confirms that alkynes interact more strongly with metal complexes than do alkenes.

Figure 5.10 (a) and (b) gives the distributions of the DMC values in alkyne and alkene complex data sets. In both distributions, the metals (Ni, Pt, Pd Rh) in groups 9 and 10, and Cu in low oxidation states have stronger binding with alkyne or alkene ligand, i.e. smaller DMC values. There are two peaks in the distribution of alkyne complex data, which illustrate high occurrence of Ni, Cu and Mo, W etc. as



(a)



(b)

Figure 5.9 TR-C_2 *vs.* TR-C_1 in metal-alkyne (a) and metal-alkene (b) complexes

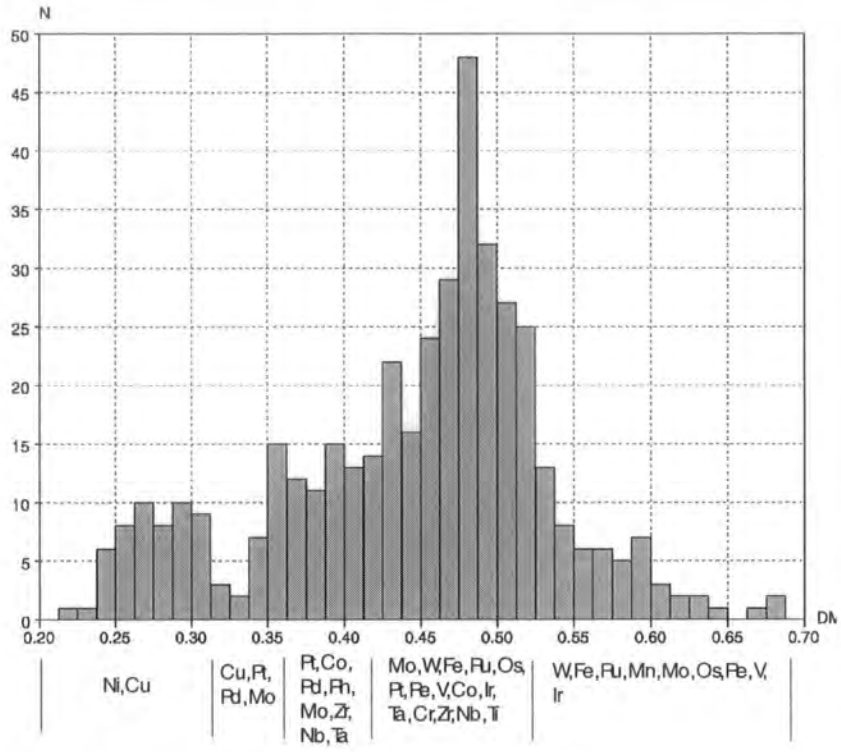
representative metals. In the alkene complex data, three major peaks represent distributions of metal Ni, Cu; Pt, Pd, Rh etc and Fe, Ir, Mn etc. complexes, as expected. This could be explained from the nature of these metals, such as, oxidation states and electron configurations (Greenwood and Earnshaw, 1994).

In addition, it is quite dependent on other coordinated ligands, especially when some other π-acid ligands compete for the $d\pi$ orbitals of the metals, the binding strength between the metal and alkyne or alkene will be decreased. Thus, a metal may have wide range of DMC values. The overlap for some metals in the range of DMC values can be seen in Figure 5.10.

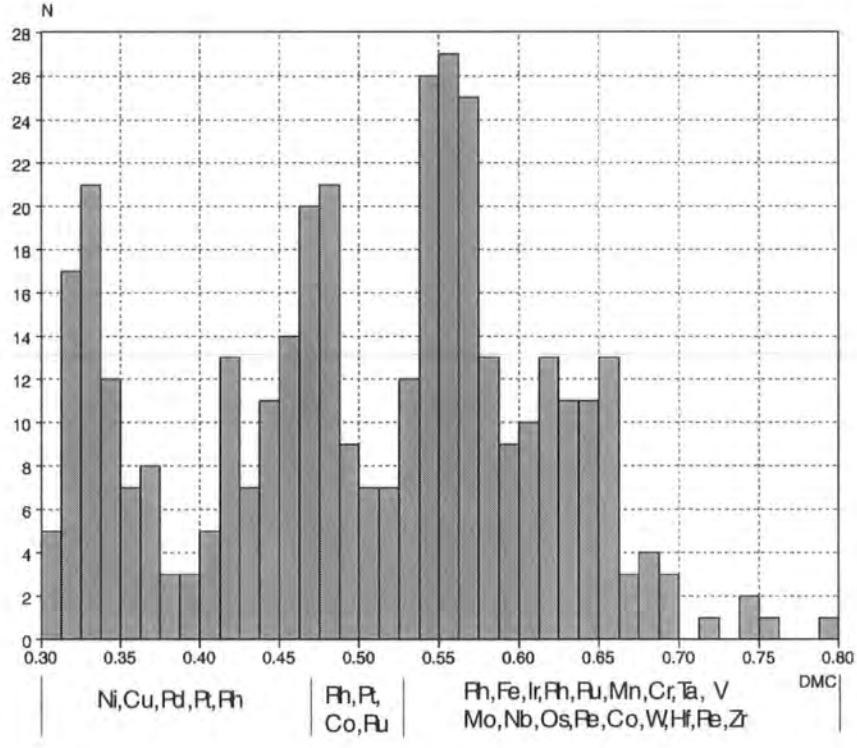
5.3.3 Influences of Substituents

Based on the σ-π bond model, metal-alkyne or metal-alkene interactions could also be affected by the substituents on the carbon atoms. It is obvious that electron-withdrawing functional groups will enhance *d*-electron back donation to π^*-orbitals; on the other hand, introduction of electron-donating groups (e.g., -NR_2, -OR) increases the donor character of the alkene and alkyne, so that this interaction becomes stronger.

Some extreme examples can easily be found in the literature (Hartley, 1972; Jutzi, Siemeling, Müller and Bögge, 1989; Hoffmann, Perez-Moya, Steigelmann and Riede, 1992). The CSD can provide many more examples, for instance, in platinum-olefin complexes, electron-negative function group substituents -F (MPZBEP), -CF_3 (OFBUPT), -C=O (HEHGIK) and -CN (FNCPPT) etc. all have shorter DMC values. π-back bonding becomes more important with increasing electronegativity of the unsaturated molecule. However, the substituent is only one of the factors influencing the interaction. A few other coordination environments, such as, other ligands, except for alkyne or alkene, coordinated with metals, metal oxidation states etc., sometimes also play a vital role. Therefore, in order to investigate the influence of substituents on this interaction more systematically or to find a correlation with different substituents, these other coordination environments should be kept as consistent as possible. For example, a hydrocarbon alkene complex (GECHEB) appears to have stronger interaction than cyano substituents (TCEPTP10) owing to the different coordination environments, in which factors other than substituents have a larger influence. Since the restriction of substituent types and tremendous variety of ligands in these kinds of complexes, there is not yet a suitable systematic method to correlate this influence from



(a)

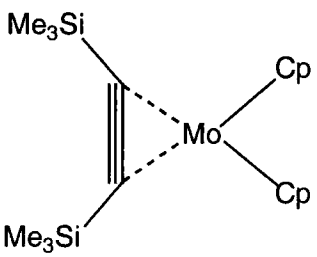


(b)

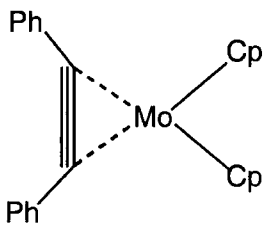
Figure 5.10 Distributions of DMC values in alkyne (a) and alkene (b) complexes

electron-withdrawing to electron-donation on the interaction between metal and C-C bonds in these data sets. Meanwhile, the lengthening of the bond C-C in the complexes is also rather insensitive to the nature of the substituents on unsaturated carbons. The comparison may still be carried out on individual cases.

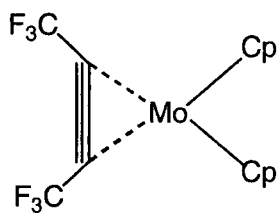
For example, molybdenum cyclopentadienyl (Cp) alkyne complexes (shown in the diagram below) with different substituents on alkene molecules, the increasing order of substituents in electronegativities is: $-SiMe_3 < -Ph < -CF_3$. The geometry parameters



SORTRAU



BALMIK



BAMMOR

corresponding to the re-order of C≡C bond of alkyne, C-C bond length; interaction between the metal and C≡C, M-C distance; and bent back angle, R-C-C are listed in Table 5.4.

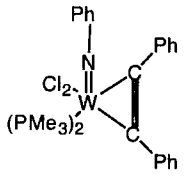
It can be seen that the stronger electron attracting group enhances the interaction between the metal and alkyne, i.e. shorter M-C distance, and the triple bond of alkyne is lengthened and there is a larger bend back angle.

Table 5.4 Influence of different substituents to the geometry parameters

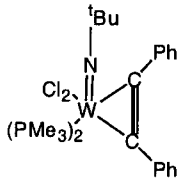
Refcode	C≡C length(Å)	M-C distance(Å)	R-C≡C angle(°)
SOTRAU	1.248	2.234	149.3
BALMIK	1.269	2.143	146.9
BAMMOR	1.282	2.128	138.2

With the same alkyne derivatives under different coordination environments on other coordinated ligands, the situations are more complex. For 4-coordinated tetrahedral complexes with alkene and alkyne π-bonding ligands, simple rules from structural patterns were suggested to distinguish the orientations of the ligands to the *d* orbitals responsible for forming π bonds and σ bonds (Gibson, 1994a,b).

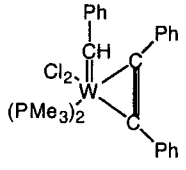
A series of diphenylacetylene tungsten complexes are chosen here, whose common coordinated ligands are dichloro-bis(trimethylphosphine) and the other ligand is variable, shown as follows:



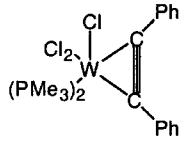
JAHNOV



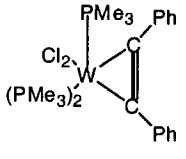
WIJKUW



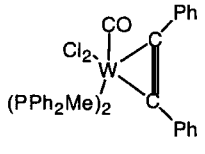
DUPMIK



YUKZOT



WIJKOP



WIJLEG

All these complexes can be seen as 6-coordinated octahedra, in which the alkyne molecule, as a ligand, locates at one of vertices of the octahedron. There are two possible orientations for alkyne to compete metal-ligand π bonds with respect to other ligands coordinated to a transition metal. This is identified as (a) C≡C bond is perpendicular to the coordinate axis and (b) C≡C lies along the coordinate axis [shown in Figure 5.11(a) and (b)]. It is considered that the orientations of the π-bonding ligands for competitive effects in these complexes are relatively easier to understand than those in tetrahedral complexes because the alignment of the ligands along the coordinate axes, which gives a clearer symmetry indication for the interaction between π-orbitals of the ligands and the *d* orbitals (the t_{2g} and e_g sets) of the metal.

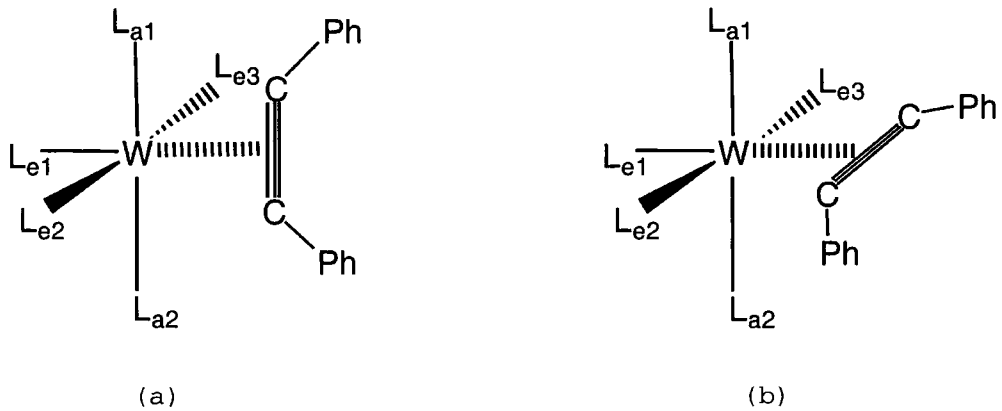


Figure 5.11 Two orientations of alkyne in octahedral complexes. (a) C≡C bond is parallel to L_{a1}-L_{a2} axis; (b) C≡C bond is perpendicular to the L_{a1}-L_{a2} axis.

The phenylimido and *t*-butylimido in JAHNOV and WIJKUW were considered as 4-electron donor ligands to the tungsten (Clark *et al*, 1989). Geometry in these complexes showed that π-acid imido ligands [placed at the L_{a1} as in Figure 5.11] are *cis* to the alkyne and the C≡C bond is along the *trans* P1-W-P2 [P1 sites at L_{e2} and P2 at L_{e3} as shown in Figure 5.11(b)] axis on the equatorial plane of an octahedron. The W-N bond is perpendicular to the W-C≡C plane. The role of imido π donation competing effectively with alkyne $\pi_\perp$ donation may decrease alkyne $\pi_\perp$ to $d\pi$ orbital and, in this orientation, the alkyne is suitably described as a 2-electron donor. The M-C distances are all over *2.1*Å. Compared with imido, the carbene ligand (DUPMIK) has less competition for the vacant $d\pi$, and interaction between tungsten and alkyne becomes stronger. The W-C distances fall into the range for 4-electron donation. The orientation of the alkyne molecule is twisted from the equatorial plane constructed by L_{e1}, L_{e2} and L_{e3}. YUKZOT, WIJKOP and WIJLEG are examples of typical tightly bound 4-electron donor alkynes, with stronger W-C interactions. In these complexes, the orientation of the alkyne molecule has C≡C rotated 90° from the equatorial plane and parallel to the W-C≡O axis (WIJLEG), W-P (WIJKOP) or W-Cl (YUKZOT) axis [i.e. M-L_{a1} in Figure 5.11(a)]. Thus, it is beneficial for overlap of the $\pi_\parallel$, $\pi_\perp$ and *d* orbitals. Figure 5.12 illustrates the relationships of W-C interactions and C-C delocalization in the alkyne ligands in this series of complexes.

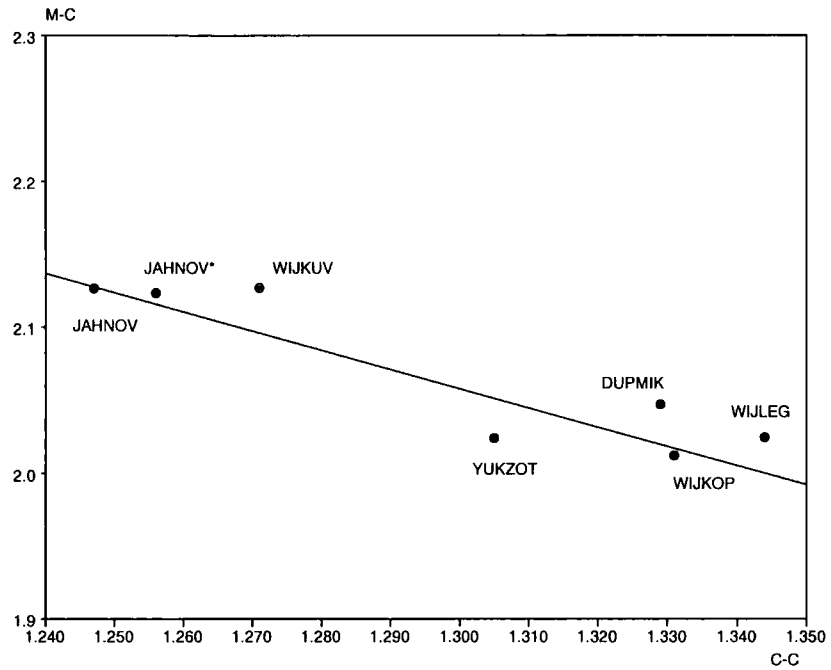
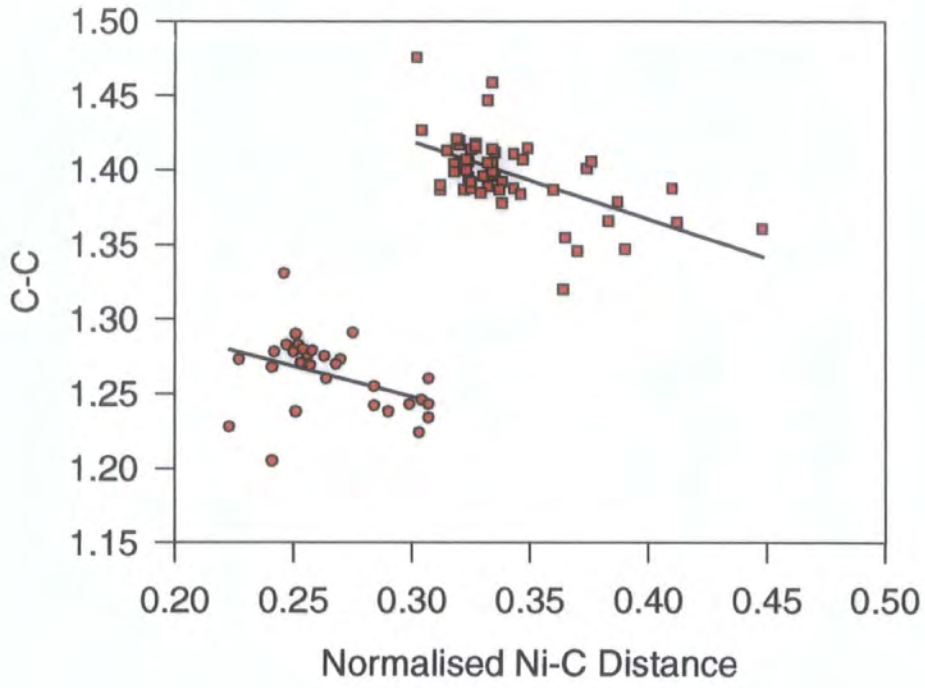


Figure 5.12 Correlation of M-C and C-C in $WL(PMe_3)_2Cl_2(PhC{\equiv}CPh)$ Complexes. *Two molecules in an asymmetric unit.

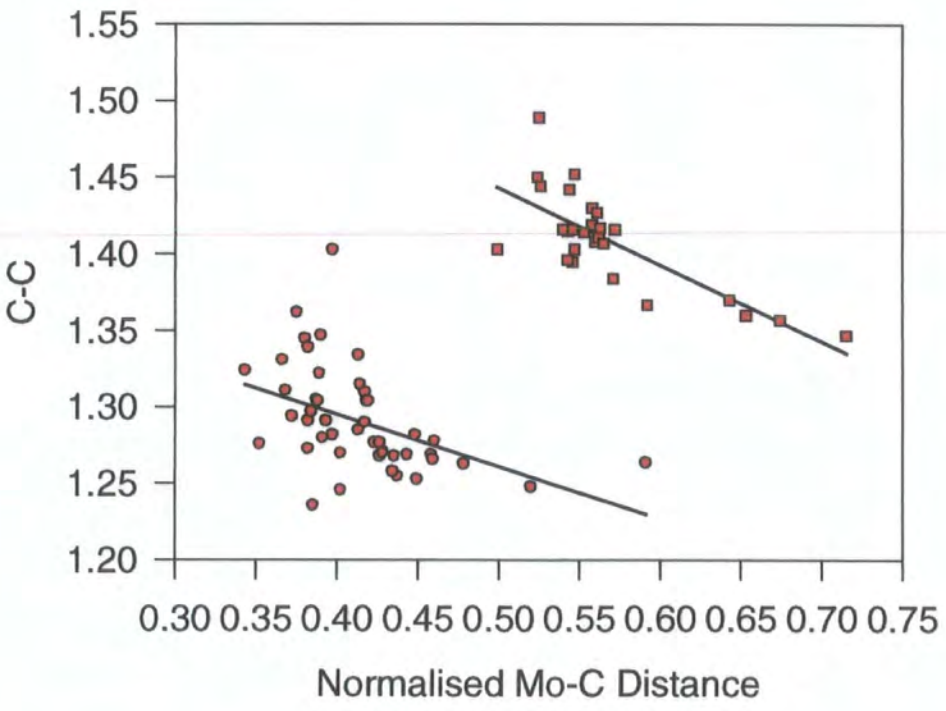
5.3.4 From Triple Bond to Double Bond and to Single Bond

The results above showed that both DMC values and C-C bond lengths reflect the interaction between the metal and C-C bonds. Can we then derive a correlation for these two parameters? If the answer is obtained, the insight of the structures of complexes will be well understood from these specific parameters. It would be beneficial to derive simple rules, from which individual structures can be rationlized and predicted. In the surveys on a large number of complexes from the CSD, and comparison of the results from alkyne- and alkene-complexes, an approximately linear relationship for these two important parameters in both sets of complexes was given, which shows a changing trend of C-C bond in the strength of interactions between the metal atom and unsaturated C-C bonds.

The trend cited in metal M-C bond distances has an inverse correlation with C-C bond lengths, that is, the C-C bond length is lengthened with stronger metal-ligand

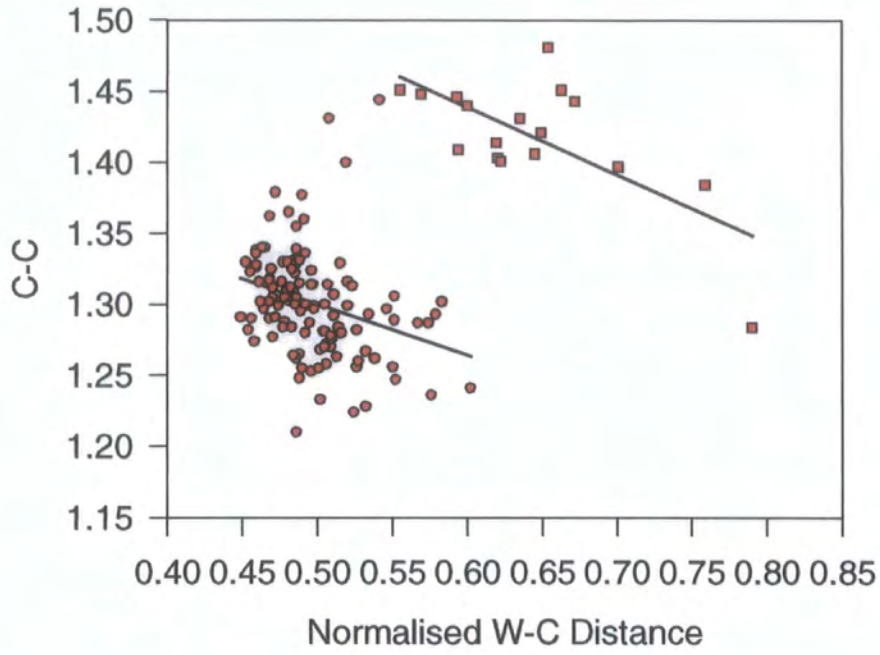


(a)

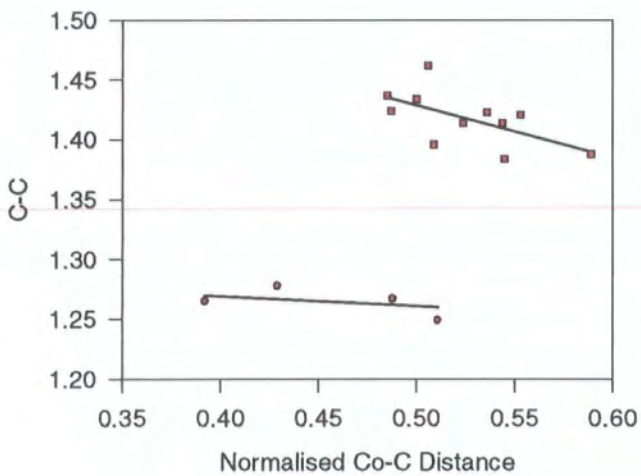


(b)

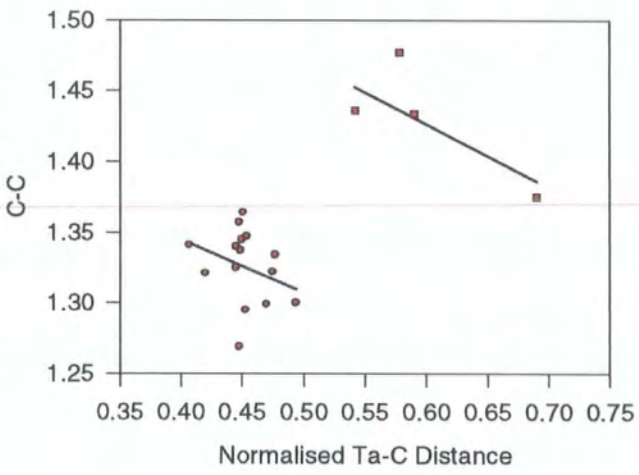
Figure 5.13 C-C vs. M-C Alkyne (▫) and alkene (o) complexes (a) Ni, (b) Mo



(c)

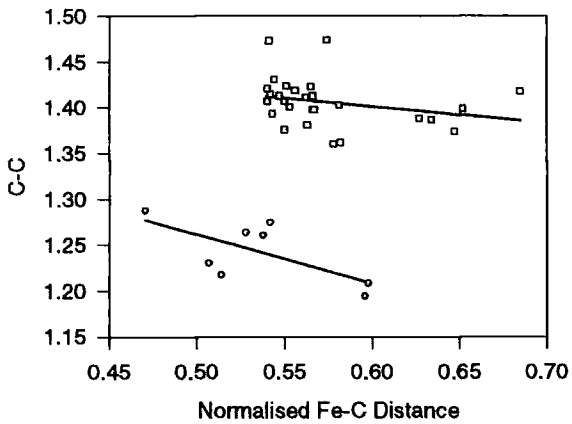


(d)

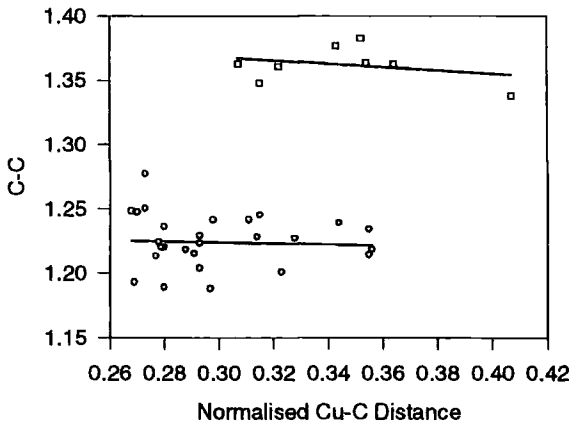


(e)

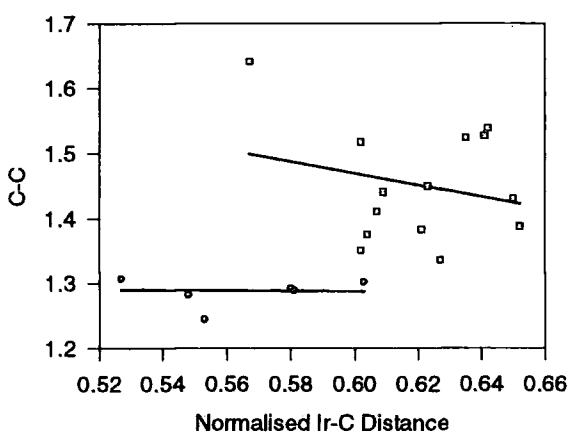
Figure 5.13 *contd.* (c) W, (d)Co, (e)Ta,



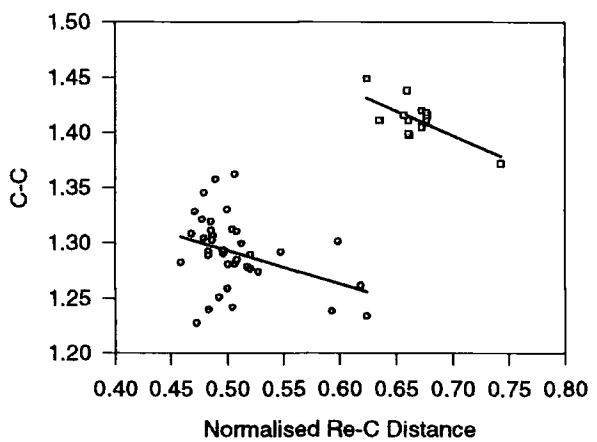
(f)



(g)



(h)



(i)

Figure 5.13 *contd.* (f)Fe, (g)Cu, (h)Ir, and (i) Re

interactions (with shorter DMC). This phenomenon has been observed in different metal complexes in which trends for bonding strength between the metals and ligands (related to M-C distance) and bond order (related to C-C distance) are given (see Figure 5.13). The linear relationships provided are very approximate, they only show a changing

trend between the two specific structural parameters. Probably, a better model to fit these observed experimental data are needed (Schultz and Messmer, 1993; Karadakov, Gerratt, Cooper and Raimondi, 1993).

It is interesting that most metals except Pt have their own interaction ranges for alkyne and alkene ligands. If the bond length (DMC) is used to assess bond strength, the conclusion is given that alkyne interacts more strongly with metal complexes than do similarly substituted alkenes. This is in accord with theoretical predictions and as expected (Nelson, Wheelock, Cusachs and Jonassen, 1969; Jensen, Bove, Westberg and Ystenes, 1995; Kitanra, Sakaki and Morokuma, 1981). The strongest interaction with alkene is nearly equal to or greater than the weakest one with alkyne for most metal atoms.

According to the model proposed in the previous section, with lengthening of C-C bond distance by the σ-π interaction, the triple bond (C≡C) in alkyne is considered as approaching a double bond (metallocyclopropene), corresponding to the strongest interaction in alkyne case in which the C-C bond length is closer to free C=C bond length. Similarly, the double bond (C=C) in alkene moves towards a single bond (metallocyclopropane). A scheme from triple bond to double bond and to single bond with the interaction strength between metal carbon bonds could be qualitatively summarized from the results of different metal complexes, shown as in Figure 5.14. As a consequence, the change of C-C bond order from 1 to 2 and to 3, corresponding to sp^1, sp^2 and sp^3 hybridization changes of C atoms, respectively, can be seen as the triple bond C≡C in alkyne is lengthened with the increase of interaction with metal, the resultant complex is considered as metallocyclopropene that has a C-C bond length closer to the free ethylene and is midway between sp^1 to sp^2 hybridization. When the double C=C bond has two substituents on each carbon atom, coordination with metal makes the bond lengths slightly longer. But the strength of interaction with metal is much weaker than that in metallocyclopropene. It is considered that the π bond dominates the coordination. Again, with the increase of interaction in this complex, a metallocyclopropane-like complex is formed, in which the carbon atoms are closer to sp^3 hybrids.

Finally, the increases in the bond lengths from a triple bond to a double bond and to a single bond for different metals are listed in Table 5.5.

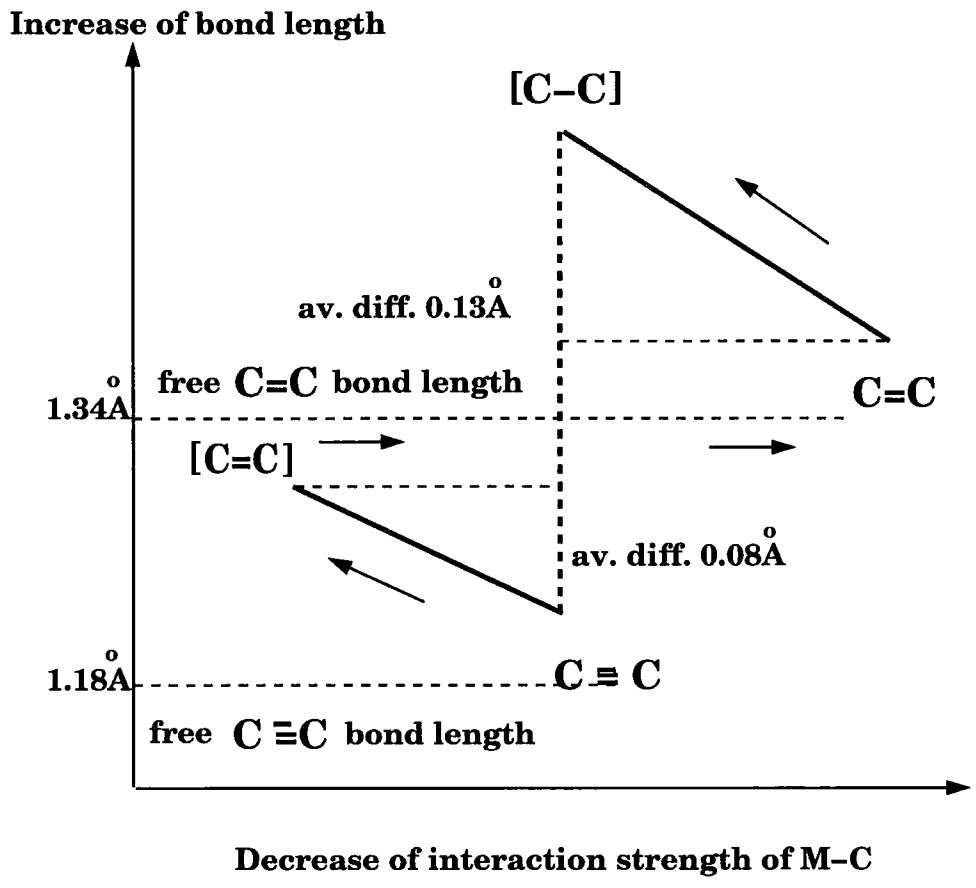


Figure 5.14 Scheme for changes in bond order

Table 5.5 Increase of bond length from triple to single bond

Metal	TB(Å)	BS(Å)
Ni	0.11	0.16
Cu	0.09	0.05
Pt	0.10	0.16
Mo	0.09	0.14
Co	0.03	0.08
Ta	0.10	0.10
W	0.10	0.17
Re	0.10	0.08
Fe	0.09	0.11
Ru	0.04	0.15
Ir	0.07	0.20

TB -- The biggest diff. bond length between C≡C & [C=C]
BS -- The biggest diff. bond length between C=C & [C-C]

In general, the C-C bond of alkynes is lengthened less upon coordination than that of alkenes. This might be interpreted as an indication that alkynes interact to a lesser degree (Ittel and Ibers, 1976). But the smaller lengthening of an alkyne molecule relative to an alkene molecule can still indicate a comparable change in bond order.

5.4 Conclusion

The broad geometrical features of transition metals alkyne and alkene complexes have now been well characterized by a systematic analysis of a large number of available structural data in the CSD. The difference of this study with other previous reviews is that the results are based on all possible known related structural studies rather than being restricted in a small samples by using the facilities provided by the CSD, several direct correlation relationships and mean values that represent geometrical characteristics of these kinds of complexes are given for the first time. From these results, the geometrical features and changes with different coordination environments become much clearer. This study is based on surveys of overall structural data, although an individual comparisons are easy to pick up. Therefore, the results reflect an average trend for these types of complexes.

For the structural data, it is clear that with greater precision, the greater the contribution that these structural parameters make to the systematic analysis. From experiment, accurate locations of atomic positions by diffraction methods are important, especially the location of the H atoms in alkene complexes that directly affect the measurement of the bend back angle. Generally, neutron diffraction could provide much greater precision for the location of both the C and H atoms of an unsaturated molecule in the presence of a transition metal. As yet, very few such diffraction results can be included in this study.

The other important factor to affect the geometrical features of alkyne and alkene molecules in complexes is the bound ML_x system in which first M and then L is varied. Exploration of changes of these coordination factors with the unsaturated molecules through a systematic approach will result in valuable information. A more suitable method for defining parameters is needed to sort out the variety of definitions for the related parameters in the CSD. In addition, for different substituted alkyne and alkene molecules, ideally a systematic comparison should be carried out on a system in which

alkynes or alkenes are bound to the same ML_x system. Clearly such systematic studies are dependent on new syntheses and structural data.

These observed results are still based on the $\sigma-\pi$ bonding model. If the overall understanding of the bonding of unsaturated molecules to transition metals is to be improved, additional experimental and theoretical work needs to be done. Before improvement of theoretical models from the calculations become available for this system, these observations are expected to provide a useful basis for further developments.

References

Albright, T. A., Hoffmann, R., Thibeault, J. C. and Thorn, D. L., *J. Am. Chem. Soc.*, **101**(14), 3801, (1979).

Bruce, M. L. in *Comprehensive Organometallic Chemistry II, Eds.* E. W. Abel, F. G. A. Stone and G. Wilkinson, **13**, Pergamon Press, Oxford, UK, (1995).

Chetcuti, M. J. in *Comprehensive Organometallic Chemistry II, Eds.* E. W. Abel, F. G. A. Stone and G. Wilkinson, **9**, Pergamon Press, Oxford, UK, (1995).

Chatt, J. and Duncanson, L. A., *J. Chem. Soc.*, 2939, (1953).

Clark, G. R., Nielson, A. J., Rickard, C. E. F. and Ware, D. C., *J. Chem. Soc., Chem. Comm.*, 343, (1989).

CSD, Cambridge Structural Database, CSD User Manuals, CCDC, (1995).

Dunitz, J. D., *X-ray Analysis and the Structure of Organic Molecules*, p330, Cornell University Press, New York, USA, (1979).

Eller, P. G., Ryan, R. R. and Schaeffer, R. O., *Cryst. Struct. Comm.*, **6**, 163, (1977).

Frenking, G. and Pidun, U., *J. Chem. Soc., Dalton Trans.*, 1653, (1997).

Gibson, V. C., *J. Chem. Soc., Dalton Trans.*, 1607, (1994a).

Gibson, V. C., *Angew. . Chem., Int. Ed. Engl.*, **33**, 1565, (1994b).

Greenwood, N., N. and Earnshaw, A., *Chemistry of the Elements*, Pergamon Press, Oxford, UK, (1994).

Hartley, F. R., *Chem. Rev.*, **69**, 799, (1969).

Hartley, F. R., *Angew. Chem., Int. Ed. Engl.*, **11**, 596, (1972).

Herberhold, M., *Metal π-Complexes,* Vol. **2**, 125, Elsevier Publishing Company, Netherlands, (1972).

Hoffmann, P., Perez-Moya, L. A., Steigelmann, O. and Riede, J., *Organometallics*, **11**, 1167, (1992).

Howard, J. A. K., Mitrpeachon, P. and Roy, A., *J. Organomet. Chem.*, **235**, 375, (1982).

Howard, J. A. K., Spencer, J. L. and Mason, S. A., *Proc. R. Soc. Lond.*, A**386**, 145, (1983).

Ittel, S. D. and Ibers, J. A., *Adv. Organomet. Chem.*, **14**, 33, (1976).

Jensen, V. R., Borve, K. J., Westberg, N. and Ystenes, M., *Organometallics*, **14**, 4349, (1995)

Jutzi, P. Siemeling, U., Müller, A. and Bögge, H., *Organometallics*, **8**, 1744, (1989).

Kitanra, K., Sakaki, S. and Morokuma, K., *Inorg. Chem.*, **20**, 2292, (1981).

Karadakov, P. B., Gerratt, J., Cooper, D. L. and Raimondi, M., *J. Am. Chem. Soc.*, **115**, 6863, (1993).

Love, R. A., Koetzle, T. F., Williams, G. J. B., Andrews, L. C. and Bau, R., *Inorg. Chem.*, **14**, 2653, (1975).

Mingos, D. M. P., in *Comprehensive Organometallic Chemistry, Eds.* G. Wilkinson, F. G. A. Stone and E. W. Abel, **3**, Pergamon Press, Oxford, UK, (1982).

Nelson, A. J., Bayd, P. D. W., Clark, G. R., Hunt, T. A., Melson, J. B. Richard, C. E. F. and Schwerdtfeger, P., *Polyhed.*, **11**, 1419, (1992).

Nelson, J. H., Wheelock, K. S., Cusachs, L. C. and Jonassen, H. B., *J. Am. Chem. Soc.*, **91**, 7005, (1969).

Pidum, U. and Frenking, G., *Organometallics*, **14**, 5325, (1995).

Rösch, N., Messmer, R. P. and Johnson, K. H., *J. Am. Chem. Soc.*, **96**, 3855, (1974).

Schilling, B. E. R., Hoffmann, R. and Lichtenberger, D. L., *J. Am. Chem. Soc.*, **101**, 585, (1979).

Schultz, P. A. and Messmer, R. P., *J. Am. Chem. Soc.*, **115**, 10925, (1993).

Templeton, J. L., *Adv. in Organomet. Chem.*, **29**, 1, (1989).

VISTA –*Version 2.0 User Guide*, Cambridge Structural Database System, Cambridge Crystallographic Data Centre, October, 1995.

Widmark, P. O., Roos, B. O. and Siegbahu, P. E. M., *J. Phy. Chem.*, **89**, 2180, (1985).

Chapter 6

Crystal Structures of the Metal Complexes

6.1 A Brief Summary of Crystal Structure Determination by X-ray Diffraction Techniques

To understand a chemical molecular structure and characterize some special structural properties and functions from a series of molecules by structure correlation, it is important to have the reliable and accurate 3D atomic coordinate data. This has been achieved by X-ray diffraction methods. A detailed picture of a molecule presented in 3D is obtained by interpreting the diffraction of X-rays from molecules making up a crystal.

For a crystallographer, a basic knowledge of X-ray diffraction theory is necessary although the modern instruments and computers can provide a fast and easy structure determinations without too much interaction from the user. But the knowledge and experience are always the essential assurance to gain correct and accurate structure data.

The fundamental theory of X-rays crystallography has been developed for more that half century. Useful standard accounts can be found in some recent published textbooks (Giacovazzo *et al,* 1992; Glusker *et al*, 1994; McKie & McKie, 1986; Woolfson, 1997). This Chapter will give a brief summary of some of the more important points in obtaining an accurate crystal structure determination. Some experimental results on the structure determinations of ten transition metal coordination complexes are given. Particularly, the geometry of coordination spheres in these complexes, which relate to the interests in this project, are described.

6.1.1 X-ray Diffraction by Crystals

X-rays were found to have a wavelength range that is compatible to the distances between adjacent atoms in crystalline solids. For crystallographic aims, a source of X-rays is usually applied in the wavelength range *0.5-2.5*Å. When an X-ray beam passes through a crystal, the incident beam is diffracted by the electrons in the atoms. The regular arrangement of the atoms in three-dimensional space is repeated throughout the crystal. Thus, the diffraction caused by the atoms will be reinforced or weakened in certain directions. Once the diffraction pattern containing information about the

arrangement of the atoms, has been recorded, by analysis, both the crystal and molecular structures may be derived.

An ideal crystal consists of some structural unit that is repeated in a perfectly regular way throughout the whole space occupied by the crystal. If the structural unit is replaced by a point, the array of the points in space that is called a lattice, reflects the geometrical relations between the repeat units. The unit cell of the crystal structure forms a parallelepiped, which is described by three non-coplanar lattice basis vectors, **a,** **b** and **c**[1] and three inter-axial angles α, β, and γ. The volume V of the unit cell is therefore given by $V = \mathbf{a}\cdot\mathbf{b} \times \mathbf{c}$. For the primitive unit cell, in which a single lattice point is repeated, any point in the lattice can be defined by a position vector $\mathbf{r} = p\mathbf{a} + q\mathbf{b} + r\mathbf{c}$, where p, q and r are integer variables. However, for a non-primitive unit cell in which more than one lattice point is repeated, certain fractional values of p, q and r may be allowed. The symmetry contained in any lattice may be characterized as belonging to one of *14* types of lattice, known as the Bravais Lattices.

Every point in a crystal structure may be related to every other point by a set of symmetry elements, which include rotations, reflections, inversions and other symmetry operations. It has been proved that there are *32* different symmetry operations that can act on a point in three-dimensional space. These *32* arrangements result in *32* point groups. Every crystalline structure must belong to one of these point groups.

If symmetry elements, glide planes and screw rotations are also considered in three dimensions, then *230* space group can be derived by combining these symmetry operations with the *32* crystal classes. The symmetries present in all crystal structure analyses are based on these *230* space groups. The symmetry is an important consideration in crystallography.

For a group of defined unit cell vectors, **a, b** and **c**, vectors **a***, **b*** and **c*** can be derived, which are reciprocal to the first set in the sense that:

$$\begin{array}{lll} \mathbf{a}\cdot\mathbf{a}^* = 1 & \mathbf{a}\cdot\mathbf{b}^* = 0 & \mathbf{a}\cdot\mathbf{c}^* = 0 \\ \mathbf{b}\cdot\mathbf{a}^* = 0 & \mathbf{b}\cdot\mathbf{b}^* = 1 & \mathbf{b}\cdot\mathbf{c}^* = 0 \\ \mathbf{c}\cdot\mathbf{a}^* = 0 & \mathbf{c}\cdot\mathbf{b}^* = 0 & \mathbf{c}\cdot\mathbf{c}^* = 1. \end{array}$$

These reciprocal vectors are given by the following relationships:

[1] the vectors are represented by bold letters

$$\mathbf{a}^* = \frac{\mathbf{b} \times \mathbf{c}}{\mathbf{V}}; \quad \mathbf{b}^* = \frac{\mathbf{c} \times \mathbf{a}}{\mathbf{V}}; \quad \mathbf{c}^* = \frac{\mathbf{a} \times \mathbf{b}}{\mathbf{V}}.$$

where V is the volume of the unit cell as given previously. **a**, **b** and **c** are basis vectors having the dimensions of length, and **a***, **b*** and **c*** are reciprocal basis vectors having the dimensions of reciprocal length. In the X-ray diffraction by a crystal, each atom within the crystal acts as a scattering centre, so that the waves scattered by different atoms interfere with one another. Constructive interference occurs between the waves scattered by all the unit cell in the crystal only, when the scattering vector has discrete values **h** such that:

$$\mathbf{h}(hkl) = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* \tag{6.1}$$

where h, k and l are integer. Therefore, the diffraction pattern of a crystal consists of series of reflections, with each reflection being specified by a particular scattering vector $\mathbf{h}(hkl)$.

In the practical applications, it is of much concern how these diffraction patterns are recorded, interpreted and thereby the structure information involved in these diffraction patterns is revealed. These are also the major processes by which we determine a crystal structure using by X-ray diffraction techniques. It can be summarized very approximately as three main steps, that is data collection, structure solution and structure refinements and representation.

The experiments on structure determination during this study were performed on X-ray diffractometers Rigaku AFC6S and Siemens Smart/CCD. The former one uses the conventional four-circle techniques for the data collection and Siemens one uses new developed techniques on the area detector so that faster overall process is achieved.

6.1.2 Data Collection

To ensure a high quality of diffraction crystal data, the first step is to have a good single crystal. A rapid and effective method of deciding if a crystal is likely to be suitable for data collection is to examine the crystal under a microscope. In the magnification field of a microscope, crystals can be investigated as to whether they have a good shape. Selection is always preferred for those that are not curved and no small crystals attached to the surface. Meanwhile, the correct size of crystal is chosen (usually

between *0.1-0.5* mm and < *1.0* mm in each of three dimensions). A further check for a single crystal is to use polarized light, if the sample is transparent, a single crystal will extinguish the polarized light every *90°* turn of the microscope stage.

A good single crystal should diffract sufficient intensity too. The reflection points should have symmetric profile and not split or excessively broadened. And a sensible unit cell can be found by indexing of the first sets of reflections.

For an air/water sensitive sample, the crystal selection should be treated under N_2 gas protection and immersed in an oil. Most of the crystal structures presented in the following section were treated in this way. The crystals are normally stuck to a thin glass fibre by glue and mounted on to a goniometer head on the ϕ circle of the diffractometer and optically aligned in the centre of the diffractometer. After these, the crystal is ready to collect diffraction data.

The diffractometer is designed to collect reflection intensity of crystal on the basis of the reciprocal relationship between a crystal lattice and its diffraction pattern, that is the Bragg equation:

$$sin\theta = \frac{\lambda}{2d_{hkl}}; \tag{6.2}$$

θ is the scattering angle, which represents the position of reflection spot on the diffraction pattern and d_{hkl} corresponds to the spacing of the *hkl* planes of the lattice.

The reciprocal lattice can be derived so long as the crystal lattice (**a, b** and **c**) is known. Each point on the diffraction pattern can be expressed in reflection indices *h, k, l,* which are three unique integers as given in Eq. (6.1).

In the traditional way, the unit cell of the crystal has to be determined first. Thus reciprocal cell parameters can be connected to diffractometer geometry. An orientation matrix is generated and then used in the control the diffractometer to find each reflection correctly because the diffractometer angles for any reflection can be calculated from this matrix. Therefore, having the correct cell parameters at the beginning is a key point to guarantee the quality of whole data collection for the four-circle diffractometer.

The cell parameters and orientation matrix are usually derived from the selected observed reflections. Different diffractometers have different strategies to do this. On Rigaku AFC6S diffractometer, a zig-zag search is used to ensure that maximum amount

of reciprocal space is scanned. Once a reflection is found, an ±ω centering routine is applied to centre diffracted beam to the Bragg reflection position and provide intensity profile of the reflection. After a set of such reflections have been found, the observed vectors $\boldsymbol{x}_i \pm \boldsymbol{x}_j$ which all correspond to reciprocal lattice points are included in a list. The three shortest non-coplanar vectors are selected as the basic reciprocal axes and the nine coordinates formed by the components of reciprocal cell axes on each of the diffractometer geometrical axes, are the nine elements of orientation matrix. Then *hkl* indices can be generated for all reflections from the determined reciprocal cell. A correction is necessary if any simple fractional indices are found. Finally the conventional cell and metric symmetry can be determined automatically after least-square refinement on the basis of all "genuine" reflections.

The orientation matrix determined by SMART/CCD follows the same convention as the real-space method. In the conventional scintillation-counter program (Such as Siemens P4), this method firstly assigns three shortest non-coplanar $\boldsymbol{x}$ vectors indices *100*, *010* and *001* to form a preliminary unit cell and then orientation matrix. From this sub-cell (**a'**, **b'** and **c'**), which is not usually correct, a set of vectors $\mathbf{t} = u\mathbf{a}' + v\mathbf{b}' + w\mathbf{c}'$ (u,v,w integral) can be generated with the maximum length limits, followed by computerizing test solution to find a true unit cell. The criterion for this is that $\mathbf{t}\cdot\mathbf{x}$ must be a integer with as low as possible residual for every reflection in the set.

An area detector can provide a much larger set of initial reflections. All differences $\boldsymbol{x}_i - \boldsymbol{x}_j$ between different vectors in the list are generated and repeats of many of these differences can be observed from the list of reflections. And averaging over the vectors improves the accuracy. However, unlike in the scintillation-counter, the matrix itself is independent of the setting angles of the frames used to determine it. Since the strategy in data collection is different for an area detector, the orientation matrix is not an essential step before data collection beginning for this machine. But a good reason to do this is to check whether or not a crystal is a real single crystal.

Nevertheless, an orientation matrix plays so important role in AFC6S-kind diffractometer, the initial matrix and cell should be well determined. The reflections with higher Bragg angles are usually required so that the relative errors in setting angles are reduced (Arndt & Willis, 1966).

Data collection on the diffractometer is a fully-automated procedure. So correctly setting various parameters to control it is basic assurance to obtain intensity data to maximum possible accuracy and precision in the minimum time and without errors.

There are several main points which need considering before setting up controlling parameters and data collection commences for AFC6S machine.

(a) There are normally two scan types, ω scan is used when a larger size of unit cell or wide reflection profiles are present. This scan can reduce the overlap between neighboring reflections, while the use of ω/2θ scan is beneficial for crystals having very low mosaic spread.

(b) The detector apperture should be set appropriate to avoid being too wide which increases background and reduces the precision, or too narrow which will cause the reflection to be truncated. The same considerations are also used for the choice of scan width.

(c) Collection speed is set to maximise/optimise the ratio of enough "observed" reflections

(d) Decide the number of reflections to measure. This depends greatly on the maximum 2θ limit. Obviously, the more reflections to be measured, the more time is required. For a Mo (K_α) radiation, the usual 2θ upper limit is $50°$, but not all crystals diffract to this angle.

Some other factors include: selection of intensity standard reflections and care with weak reflections.

Instead of a serial measurement of reflections in certain order, modern area detector diffractometers can measure a large number of reflections simultaneously. Unlike the traditional techniques, in which each reflection is recorded one at a time, an area detector records the whole of the intercepted diffraction pattern and not just the Bragg reflections. Thus, the whole of reciprocal space is observed. A major advantage of this machine is that the data collection process is greatly shortened. For a medium size crystal structure, data collection time could reduce from a couple days using a four-circle diffractometer to less than 10 hours for an area detector. Since the ability to record a large fraction of the reflections which simultaneously cross the Ewald sphere, the whole hemisphere data can be recorded for any structure and there is no need to consider what Laue symmetry it is, which is normally used in the traditional method. This also provides an advantage to average on reflections when the equivalent reflections are merged on the real symmetry, thus improving the accuracy. The parameters to control the data collection are mainly concerned with maximum 2θ angle, detector swing angle, at which the frames in the scan series are to be collected. ω

angle indicates the start of the frame series; frame width, that is, scan width of each frame to be acquired in degree; scan time – time in second spent on each frame. Once data collection starts, a full data set can be collected regardless of the crystal orientation.

Data reduction

From a diffractometer, raw integrated intensities of reflections are collected. For the further use of structure determination, relative structure factor amplitudes $|F_o(\mathbf{h})|$ or amplitude squares $|F_o(\mathbf{h})^2|$ need to be extracted. In this procedure, several correction factors are necessary, which may arise from instruments and systematic errors into the diffraction process.

(a) Lorentz and polarization corrections. These two factors are dependent on the experimental conditions of the intensity measurement. They correspond to corrections for instrument-geometry and X-rays beam polarisation when measuring each reflection. These corrections must be applied to every set of data. Because they both depend systematically on the Bragg angle θ, they can be set into data reduction programs, no operator input or interaction is actually needed.

(b) When X-rays are incident on a crystal, the diffracted beam may have some decay due to the absorption of the crystal. The absorption is dependent on the contents of the unit cell and on the crystal morphology. Hence the absorption correction may be required to enhance the quality of the data. If the incident beam has an intensity of I_0, the intensity of the diffracted X-ray (I) is :

$$I = I_0 e^{-\mu} . \qquad (6.3)$$

where μ is the linear absorption coefficient (in mm^{-1}) corresponding to the cell contents and the radiation used. t is the average crystal dimension (in mm). The absorption correction usually is larger with Cu radiation than Mo X-radiation.

The common method used for the absorption correction is azimuthal (ψ) scan, where the crystal is rotated about the scattering vectors and a reflection will still remain in the diffraction position. The absorption effects may be measured from the variation in intensity observed.

Usually, for this test, a few of reflections with χ angle at $\sim 90°$ is chosen because the movement of diffractometer at this angle corresponds to rotation about ϕ, while at

other χ values to a more complicated condition of circle movements. Once the absorption factors have been determined, it will be applied to all reflections.

Some other possible corrections include decay correction, extinction and thermal diffuse corrections etc. They may be applied to individual crystal structures depending on the quality of crystal.

Using an area detector, raw data is recorded in sequential frames. An integration method that is based on 3D profile algorithms is used (SAINT Software Reference Menu, 1995) to perform all the necessary steps to reduce the data. Finally, the same format as those provided by the conventional methods for all reflections in *hkl*s, *LP*-correction intensities and standard deviation can be obtained. An accurate orientation matrix is required at this stage. The reflections to determine the unit cell are chosen from all collected reflections. In addition, the position and shapes of reflection spots on the detector need to be considered and the correct X, Y beam centre and the sample-detector distance should be given.

6.1.3 Structure Determination

A set of data with *hkl* indices, intensities [$|F_o(\mathbf{h})|$ or $|F_o(\mathbf{h})^2|$] and standard deviations (σ) is generated as described above The next step is how to analyse the data and thereby reveal the structure, i.e. the atomic positions **r** in the unit cell.

The unit-cell contents can be described in terms of electron density $\rho(\mathbf{r})$ and expressed as Fourier transform of structure factors $F(\mathbf{h})$:

$$\rho(\mathbf{r}) = \frac{1}{\mathbf{V}} \sum_{\mathbf{h}} \mathbf{F(h)} \exp(-2\pi i \mathbf{h} \cdot \mathbf{r}) \qquad (6.4)$$

Atom coordinates will be located from the peak positions of $\rho(\mathbf{r})$ values. Therefore, if structure factors $F(\mathbf{h})$ is known, $\rho(\mathbf{r})$ can be calculated to draw up an electron density map in the unit cell. $F(\mathbf{h})$ is a complex quantity and can be written as a combination of a magnitude $|F(\mathbf{h})|$ and a phase represented by $\phi(\mathbf{h})$:

$$\mathbf{F(h)} = | \mathbf{F(h)} | \, \mathbf{e}^{i\phi(\mathbf{h})} \qquad (6.5)$$

As given above, the magnitude part has been obtained from the dada collection, but no direct information for phase derives from the measurement of the data. Thus, to solve a

structure, it requires the solution of this "phase problem". Many contributions have been made in solving this problem in many decades (Patterson, 1934; Harker & Kasper, 1948; Sayre, 1952; Woolfson, 1958; Karle & Hauptman, 1956; Karle & Karle, 1964: Debaerdemaeker *et al*, 1988). Now there are two major methods in use in the routine performance of small molecule structure solution.

Patterson method

Patterson function uses $|F(\mathbf{h})|^2$ instead of $F(\mathbf{h})$ in the electron density equation [Eq. (6.4)], thus:

$$P(\mathbf{u}) = \frac{1}{V}\sum_{\mathbf{h}} |\mathbf{F}(\mathbf{h})|^2 \exp(2\pi i \mathbf{h}\cdot\mathbf{u}) \qquad (6.6)$$

which will show a map of interatomic vectors **u** and phase angles all disappear in this expression.

Since $|F(\mathbf{h})|^2$ depends on the scattering power while the scatter power are proportional to the atomic number, the interactions between heavy atoms will appear as distinguishable peaks in the resulting Patterson maps. This provides the possibility to derive the heavy atom positions from those strong vector peaks according to the symmetry elements in the unit cell. The remaining atoms may be further derived from these defined heavy atom positions by Fourier methods. This usually works well for the structures containing heavy atoms but for "equal" atom structures, e.g. most organic molecules, the maps can not be deconvoluted because no definite peaks can be located to any special atoms.

Direct Methods

Direct methods are more powerful for various molecular structure solutions. They seek to estimate the phases directly from the observed structure amplitudes by mathematical processes (Woolfson, 1977).

Structure factor amplitudes and phases may have special relationships. Such relationships may be used and reveal the electron density map. When some mathematical constraints are applied on the electron density function $\rho(\mathbf{r})$, the phase

values may be determined from the derived relationships based on the corresponding probability theory (Cochran & Woolfson, 1955).

In the practical application, normalized structure factors E(**h**) are used instead of the structure factors F(**h**) to remove the effects of the atomic shape and thermal motion. Thus, atoms in the structure are considered as points, which will enhance one of the constraints used in this method, i.e. all the atoms in the unit cell are discrete.

The other two useful constraints are (a) the electron densities in the unit cell has never a negative value and (b) maximum value of $\int \rho^3(\mathbf{x})\mathbf{dV}$ over the whole of the unit cell. From these, the phase relationships and tangent formula (Karle & Hauptman, 1956) could be derived. This latter allows the users to determine the phase values by setting up phase relationships from certain combinations of indices of reflections. Those strong reflections E(**h**) with acceptable reliability are used for phase determination.

Random values (Yao, 1981) are assigned to all of the phases to be determined and the phase refinement procedure is repeated until a set of satisfied solution is reached.

Finally, the correct set of phases is identified from the large number of incorrect phase sets by the designed figures of merit, which are quick tests for the quality of the phases.

In the Patterson method, only the heavy atoms are located, while in direct methods, most of structure can usually be derived. From these known atomic positions, further Fourier refinement cycles are needed. The missing atoms are found and the known atoms are refined through the cycles, resulting in the complete structure.

6.1.4 Structure Refinements

A solved structure means having located all the atoms in approximate positions and assigned initial amplitudes for the thermal vibration for each atom. The following step is to refine these parameters to more accurate values so that the difference between observed structure factor amplitudes $|F_o(\mathbf{h})|$ or $|F_o(\mathbf{h})^2|$ and calculated $|F_c(\mathbf{h})|$ or $|F_c(\mathbf{h})^2|$ is a minimum, i.e.

$$R = \left[\frac{\sum_{h} \left(|F_o(h)| - |F_c(h)| \right)^2}{\sum_{h} |F_o(h)|^2} \right]^{\frac{1}{2}} \qquad (6.7)$$

or replace | F | by | F^2 | in the expression. This is an expression to test the reliability of the structure by the comparison of the calculated structure factors with those deduced from the experimental data. It is obvious that the smaller R values represent better agreement between the observed and calculated structure factors, thus indicate a satisfactory and reliable structure.

The standard mathematical least-squares refinement techniques are used in the crystal structure refinements and a cyclic process of refinement is carried out until convergence is achieved.

The advantage of using F^2 instead of F's is that all weak reflections can be involved in the refinements. These weak reflections also contain somehow important structural information. In all of the structures detailed below, SHELXL-93 (Sheldrick, 1993) program, which is a least-square refinement program based on F^2, was used.

Temperature factors of atoms are additional parameters required in the refinements, because a temperature factor term $exp(-Bsin^2\theta/\lambda^2)$ should be included in the calculated structure factors when comparing them with observed structure factors. These parameters can also be good indicators of the quality of the model. The atoms in a crystal vibrate about their mean positions due to their having thermal energies. This thermal motion makes the electron density peak more diffuse and thus reduces the scattering power at higher angles so that a more rapid decline of the atomic scattering factors $f'(\theta)$ is observed. Since the square of mean vibration amplitude increases proportionally to the temperature, $f'(\theta)$ can be written in the form:

$$f'(\theta) = f \exp[-B(\sin^2 \theta / \lambda^2)] \qquad (6.8)$$

where θ is the Bragg angle of a particular reflection, f is the value of the atomic scattering factor at zero scattering angle, and B is the temperature factor. B is equal to $8\pi^2 \overline{U}^2$, where $\overline{U}^2$ represents the mean square displacement of an atom from its equilibrium position which is assumed to be the same in all given directions. In practice,

the thermal vibrations of an atom in a crystal are not isotropic and one can describe them in terms of an ellipsoid of vibration in reciprocal space. Thus it can be expressed as:

$$-\frac{1}{4}(B_{11}h^2\,a^{*2}+B_{22}k^2b^{*2}+B_{33}l^2c^{*2}+2B_{23}klb^*c^*+2B_{13}hla^*c^*+2B_{12}hka^*b^*) \quad \text{or}$$

$$-2\pi^2(U_{11}h^2\,a^{*2}+U_{22}k^2b^{*2}+U_{33}l^2c^{*2}+2U_{23}klb^*c^*+2U_{13}hla^*c^*+2U_{12}hka^*b^*) \quad (6.9)$$

in which a^*, b^* and c^* are reciprocal cell axes as given previously, and two formula are equivalent to isotropic B or U, respectively.

Very large U_{ij} value may indicate a disordered atom, or it may just be due to imprecise (high esd's) U_{ij} parameters.

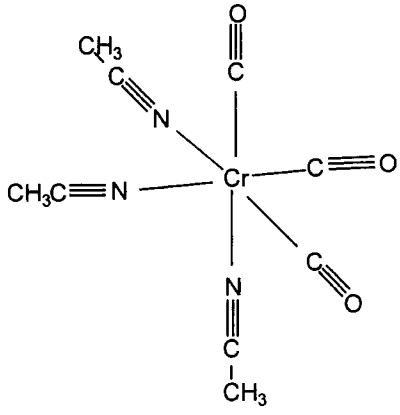
Finally, for a complete structure, three coordinates (x,y,z) for each atom are listed in fractional forms with estimated standard deviation values in parentheses as well as displacements. From the coordinates, all corresponding geometrical parameters, e.g. bond lengths and angles etc. can be derived, together with a measure of their reliability. Three-dimensional structure diagrams can be produced. These also enable one to use the CSD for structure correlation studies.

6.2 Determined Structures

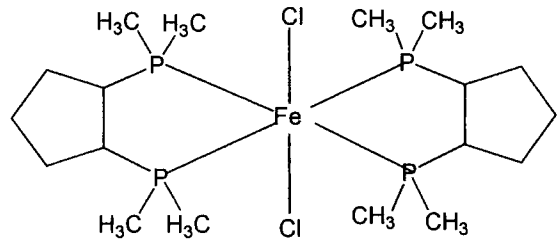
The crystal structures of ten metal complexes* in low coordination numbers have been determined by X-ray diffraction methods. The line diagrams are shown in Figure 6.1. The details of the experiments and structures will be given in the following sections. The coordination geometry for these individual compounds will be examined from the resultant geometry parameters. The atomic coordinates for each structure will be listed in Appendix I.

Interest is mainly focused on the geometry of the metal coordination spheres. There are also other interesting structural features, which are not discussed in detail here.

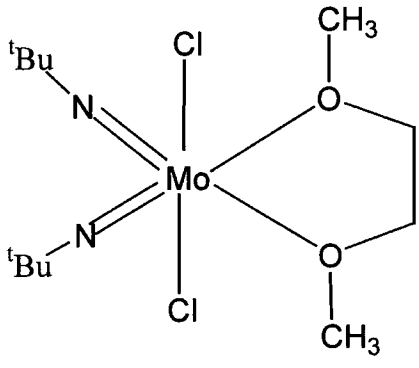
* I would like to take this opportunity to thank the persons who supplied these crystal samples. The names have been given in Acknowledgements.



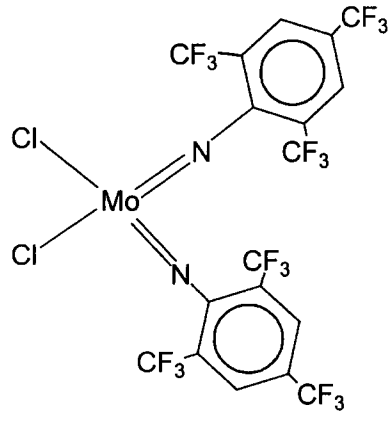
I



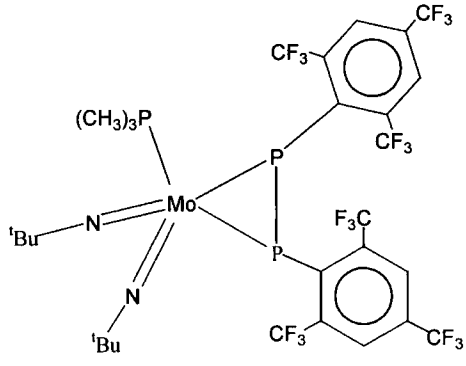
II



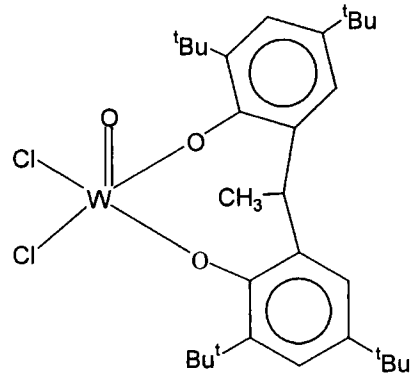
III



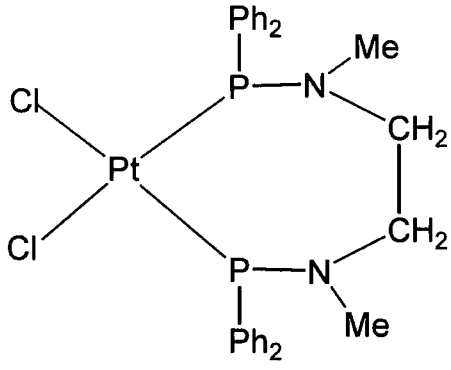
IV



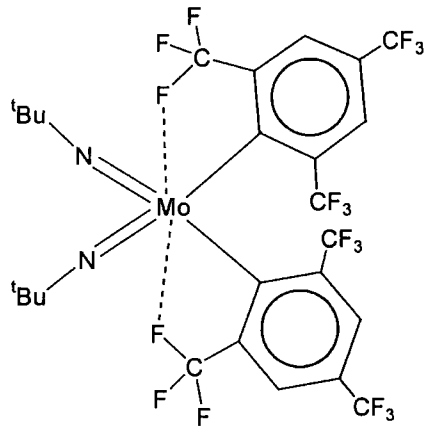
V



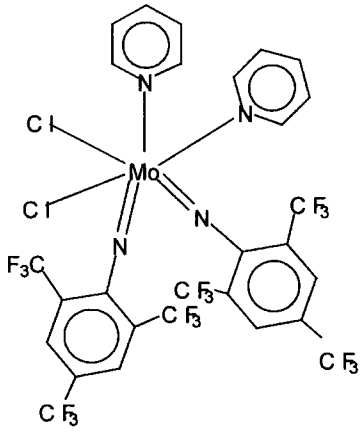
VI



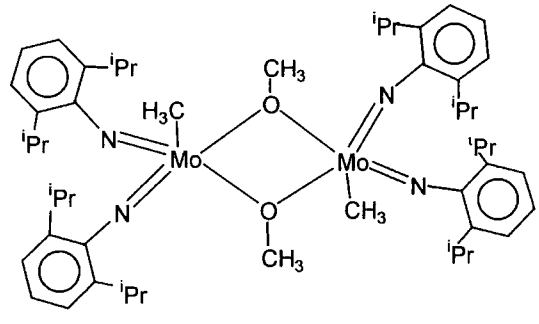
VII



VIII



IX



X

Figure 6.1 Molecular connectivity diagrams used in this chapter

I. Tri-carbony-trimethylcyanide-chromium $[Cr(CO)_3(CH_3CN)_3 \cdot CH_3CN]$

This structure shows a 6-coordination sphere of chromium. Six unidentate ligands (three CH_3CN and three CO) are involved in the coordination sphere. There are no chelate effects so it has effectively an octahedral geometry (shown as Figure 6.2). The key parameters for this geometry are listed in Table 6.2. They all have close values to the standard ones in an octahedron, that is, valence angles between the metal (M) and ligands (L), L-M-L, 90°, 180°.

Empirical formula	$C_{11}H_{12}N_4O_3Cr$
Formula weight	300.2
Temperature	150(2)K
Wavelength	0.71073Å
Crystal system	Orthorhombic
Space group	$Pca2_1$
Unit cell dimensions	a = 12.4622(1)Å α = 90.0°
	b = 8.2455(1)Å β = 90.0°
	c = 14.7451(2)Å γ = 90.0°
Volume	1515.16(3) $Å^3$
Z	4
Crystal colour	Yellow
Absorption Coefficient	0.939 mm^{-1}
F(000)	660
Crystal size	0.48 × 0.24 × 0.20 mm
θ range for data collection	2.47 → 27.45°
Index range	$-16 \leq h \leq 15, -9 \leq k \leq 10, -19 \leq l \leq 18$
Experiment device	Siemens Smart CCD
Reflections collected	10291
Independent reflections	3447 [R_{int} = 0.0352]
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	3424/0/208
Goodness-of-fit on F^2	1.191
Final R indices [I > 2σ(I)]	R1 = 0.0285, wR2 = 0.0610
R indices (all data)	R1 = 0.0424, wR2 = 0.0763
Largest diff. Peak and hole	0.205 and -0.273 $eÅ^{-3}$

There is a cyanide molecule as a solvent molecule and as shown in the packing in the unit cell, it forms a weak hydrogen bond with the oxygen atom of one of the coordinated carbonyl ligands [C≡O(2)].

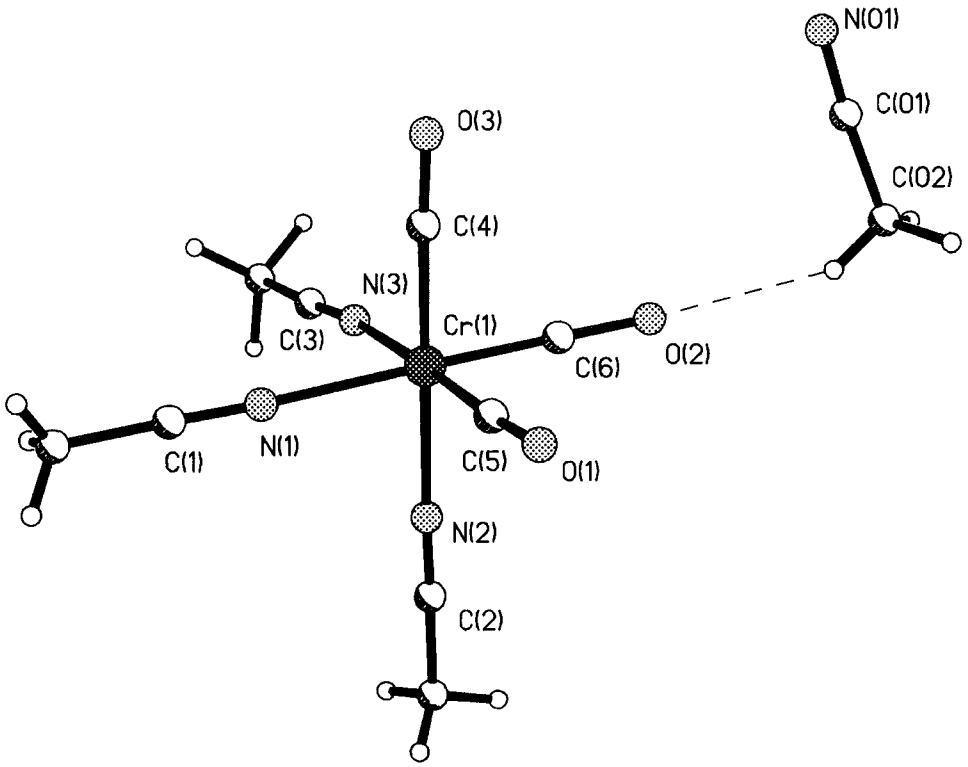


Figure 6.2 Structure of $Cr(CO)_3(CH_3CN)_3 \cdot CH_3CN$

Table 6.1 Bond lengths (Å) and bond angles (°) for the coordination sphere in **I**

Cr(1)-C(4) 1.827(2)	Cr(1)-C(5) 1.848(6)	Cr(1)- C(6) 1.802(7)
Cr(1)-N(1) 2.110(5)	Cr(1)-N(2) 2.106(2)	Cr(1)- N(3) 2.100(6)
C(4)-Cr(1)-C(5) 88.5(3)	C(4)-Cr(1)-C(6) 83.3(4)	C(4)-Cr(1)-N(1) 95.9(3)
C(4)-Cr(1)-N(2)178.2(4)	C(4)-Cr(1)-N(3) 93.2(3)	C(5)-Cr(1)-C(6) 85.1(1)
C(5)-Cr(1)-N(1) 97.1(3)	C(5)-Cr(1)-N(2) 93.1(3)	C(5)-Cr(1)-N(3)177.9(3)
C(6)-Cr(1)-N(1)176.9(3)	C(6)-Cr(1)-N(2) 93.1(3)	C(6)-Cr(1)-N(3) 96.2(3)
N(1)-Cr(1)-N(2) 84.6(3)	N(1)-Cr(1)-N(3) 81.5(1)	N(2)-Cr(1)-N(3) 85.2(3)

II. Trans-bis[1,2-bis(dimethylphosphino)cyclopentane]-dichloroiron (II)

This is another six-coordination complex, metal iron(II) coordinates with four phosphine atoms of 1,2-bis(dimethylphosphino) cyclopentane on the equatorial plane

Empirical formula	$C_{18}H_{40}Cl_2P_2Fe$
Formula weight	507.1
Temperature	293(2)K
Wavelength	0.71073Å
Crystal system	Monoclinic
Space group	$P2_1/n$
Unit cell dimensions	a = 7.996(2)Å α = 90.0°
	b = 16.846(3)Å β = 97.11(3)°
	c = 9.076(2)Å γ = 90.0°
Volume	1213.1(5) $Å^3$
Z	2
Crystal colour	Green
Absorption Coefficient	1.108 mm^{-1}
F(000)	536
Crystal size	0.75 × 0.55 × 0.30 mm
θ range for data collection	2.56 → 25.00°
Index range	$-9 \le h \le 9, -20 \le k \le 20, -10 \le l \le 1$
Experiment device	Rigaku AFC6S
Reflections collected	5057
Independent reflections	2143 [R_{int} = 0.0602]
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	2142/0/140
Goodness-of-fit on F^2	1.141
Final R indices [I > 2σ(I)]	R1 = 0.0356, wR2 = 0.0883
R indices (all data)	R1 = 0.0555, wR2 = 0.0958
Largest diff. Peak and hole	0.365 and -0.453 $eÅ^{-3}$

and two chlorine atoms in the axial positions. The coordination sphere is almost a perfect octahedron. The angles between two P atoms with the metal on chelate ring,

P(1)-Fe -P(2) 86.5°, is little smaller than the standard value of 90° . While angle of P(1)-Fe-P(2)$^{\#*}$ is 3.5° greater than the standard value.

The complex crystallizes in $P2_1/n$ and the metal atom sits at a special position of this space group. Half the molecule can be located in asymmetric unit and the other half can be found by the inversion centre in this symmetry. Two carbon atoms, C(1) and C(5), on the cyclopentane were found to be disordered. Each occupies 51% and 49%, respectively. This disorder has been shown in Figure 6.3 by dotted lines. The cyclopentane rings have a chair conformation.

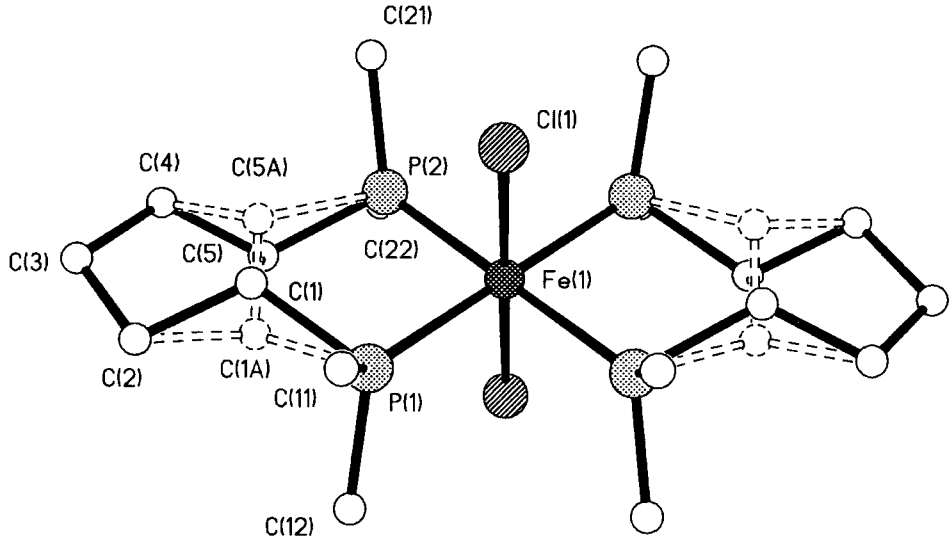


Figure 6.3 A plot of complex Fe[1,2-bis(dimethylphosphino)cyclopentane]$_2$Cl$_2$ with disorder model

The major geometry parameters are listed in Table 6.3.

Table 6.2 Bond lengths (Å) and bond angles (°) for complex **II**

Fe(1)-P(1)	2.257(1)	Fe(1)-P(2)	2.258(1)
Fe(1)- Cl(6)	2.344(1)	P(1)-Fe(1)-P(1)*	180.0
P(1)-Fe(1)-P(2)	86.50(3)	P(1)-Fe(1)-P(2)*	93.50(3)
P(1)-Fe(1)-Cl(1)	89.88(3)	P(2)-Fe(1)-Cl(1)	89.91(3)
P(2)-Fe(1)-P(2)*	180.0	Cl(1)-Fe(1)-P(1)*	90.12(3)
Cl(1)-Fe(1)-P(2)*	90.12(3)	Cl(1)-Fe(1)-Cl(1)*	180.0

* Symmetry transformations: -x, -y, -z+1

III Dichloro-(dimethoxyethane)-bis(t-butylimido)-molybdenum(VI)

This complex contains a bidentate ligand, dimethoxyethane, with two oxygen atoms coordinated to the metal Mo. The coordination sphere is a distorted octahedron, two chlorine atoms locate at the axial positions and the bidendate ligand and two t-butylimido on the plane. Because of the chelate ring between the Mo and dimethoxyethane ligand and bulkier t-butylimido ligands, the axial atoms (Cl) incline

Empirical formula	$C_{12}H_{28}N_2O_2Cl_2Mo$
Formula weight	399.22
Temperature	150(2)K
Wavelength	0.71073Å
Crystal system	Orthorhombic
Space group	Pbca
Unit cell dimensions	a = 9.864(1)Å α = 90.0º
	b = 12.495(1)Å β = 90.0º
	c = 29.853(4)Å γ = 90.0º
Volume	3679.4(1) $Å^3$
Z	8
Crystal colour	Yellow
Absorption Coefficient	1.004 mm^{-1}
F(000)	660
Crystal size	0.40 × 0.22 × 0.20 mm
θ range for data collection	2.47 → 26.15°
Index range	$-11 \le h \le 12$, $-15 \le k \le 15$, $36 \le l \le 19$
Experiment device	Siemens Smart CCD
Reflections collected	14709
Independent reflections	3340 [R_{int} = 0.0673]
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	3340/0/257
Goodness-of-fit on F^2	1.211
Final R indices [I > 2σ(I)]	R1 = 0.0453, wR2 = 0.1091
R indices (all data)	R1 = 0.0508, wR2 = 0.1132
Largest diff. Peak and hole	0.510 and -0.509 $eÅ^{-3}$

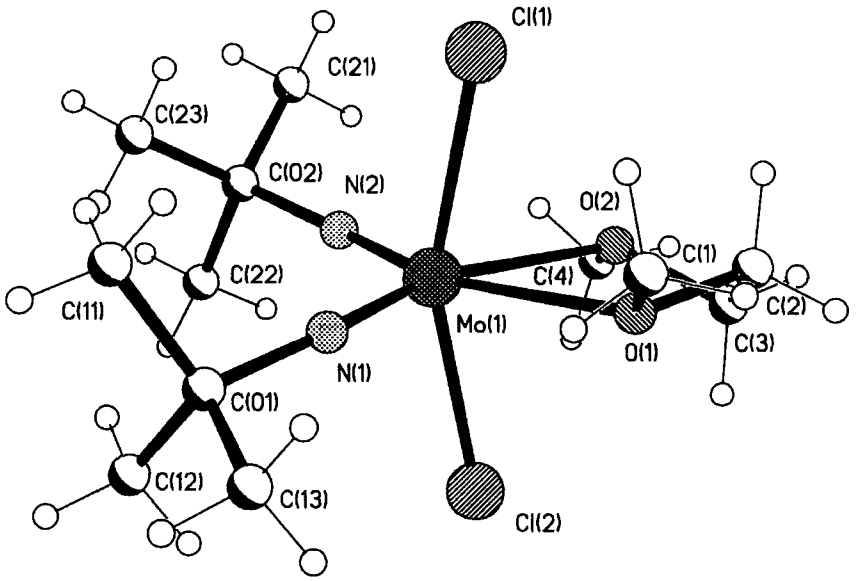


Figure 6.4 View of structure III

to the chelate ring, the angle of Cl(1)-Mo-Cl(2) is obviously smaller than the standard value 180° [158.96(4) °]. The angle between bidendate atoms O(1) and O(2) is also contracted to 68.9(1) °, considerably less than 90°. All these factors make this coordination geometry rather distorted from a regular octahedron.

Table 6.3 Selected bond lengths(Å) and angles(°) for structure III

Mo(1)-N(1)	1.726(3)	Mo(1)-N(2)	1.722(3)
Mo(1)-O(2)	2.380(3)	Mo(1)-Cl(1)	2.409(1)
Mo(1)-O(1)	2.397(3)	Mo(1)-Cl(2)	2.427(1)
Cl(1)-Mo(1)-Cl(2)	158.96(4)	Cl(1)-Mo(1)-N(1)	99.2(1)
Cl(1)-Mo(1)-N(2)	93.6(1)	Cl(1)-Mo(1)-O(1)	79.4(1)
Cl(1)-Mo(1)-O(2)	83.6(1)	Cl(2)-Mo(1)-N(1)	92.9(1)
Cl(2)-Mo(1)-N(2)	99.2(3)	Cl(2)-Mo(1)-O(1)	83.1(1)
Cl(2)-Mo(1)-O(2)	79.4(1)	N(1)-Mo(1)-O(1)	92.3(1)
N(1)-Mo(1)-O(2)	160.3(1)	N(1)-Mo(1)-N(2)	107.1(1)
N(2)-Mo(1)-O(1)	160.1(1)	N(2)-Mo(1)-O(2)	92.1(1)
O(1)-Mo(1)-O(2)	68.9(1)		

IV Dichloro-bis(2,4,6-trifluoromethyl-phenylimido)-molybdenum(VI)

Complex IV crystallises in the space group $P\bar{1}$. The coordination geometry of the Mo atom can be described as tetrahedral with two chlorine atoms and two phenylimido ligands located at the *cis*-position. The four M-L bond lengths all fall into the normal values for these ligands and the angle between two chlorine atoms

Empirical formula	$C_{18}H_4N_2F_{18}Cl_2Mo$	
Formula weight	757.07	
Temperature	150(2)K	
Wavelength	0.71073Å	
Crystal system	Triclinic	
Space group	$P\bar{1}$	
Unit cell dimensions	a = 8.917(2)Å	α = 96.78(3)°
	b = 9.078(2)Å	β = 104.08(3)°
	c = 16.050(3)Å	γ = 103.63(3)°
Volume	1203.4(4) Å^3	
Z	2	
Crystal colour	Dark red	
Absorption Coefficient	0.919 mm^{-1}	
F(000)	728	
Crystal size	0.44 × 0.32 × 0.24 mm	
θ range for data collection	2.35 → 23.26°	
Index range	$-9 \le h \le 9$, $-10 \le k \le 9$, $-17 \le l \le 14$	
Experiment device	Siemens Smart CCD	
Reflections collected	4622	
Independent reflections	3308 [R_{int} = 0.0319]	
Refinement method	Full-matrix least-squares on F^2	
Data/restraints/parameters	3308/0/382	
Goodness-of-fit on F^2	1.100	
Final R indices [I > 2σ(I)]	R1 = 0.0371, wR2 = 0.0955	
R indices (all data)	R1 = 0.0378, wR2 = 0.0962	
Largest diff. Peak and hole	0.931 and -0.755 eÅ^{-3}	

Cl(1)-Mo-Cl(2),115.5(1)°, is a little stretched from the standard angle for a perfect tetrahedron. But overall, the structure is a typical tetrahedral pattern.

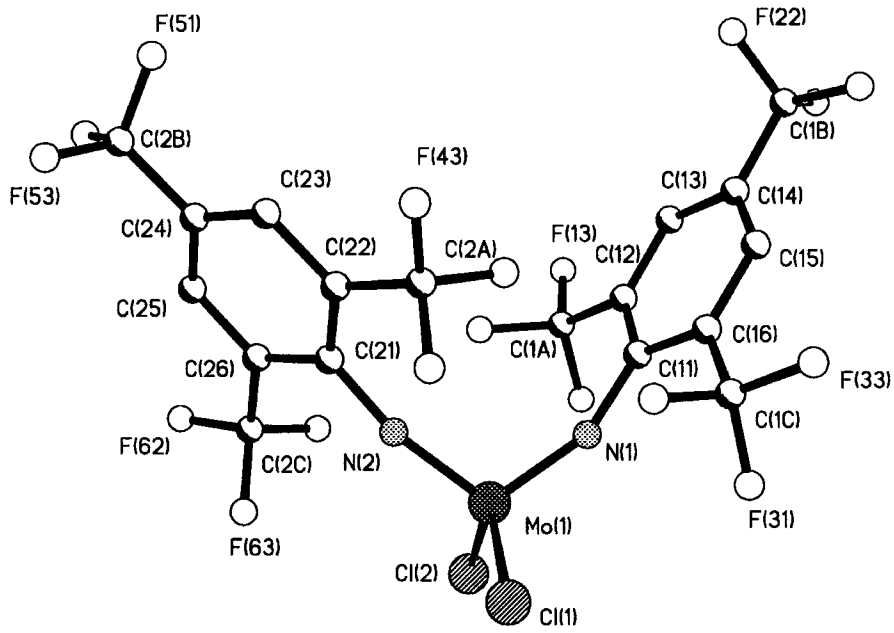


Figure 6.5 A view of the molecule $Mo[NC_6H_2(CF_3)_3]_2Cl_2$

Table 6.4 Selected bond lengths and angles for complex $Mo[NC_6H_2(CF_3)_3]_2Cl_2$

Mo(1)-Cl(1)	2.252(1)	Mo(1)-Cl(2)	2.259(1)
Mo(1)-N(2)	1.736(3)	Mo(1)-N(1)	1.745(3)
Cl(1)-Mo(1)-Cl(2)	115.5(1)	Cl(1)-Mo(1)-N(1)	107.6(1)
Cl(1)-Mo(1)-N(2)	107.3(1)	Cl(2)-Mo(1)-N(1)	108.9(1)
Cl(2)-Mo(1)-N(2)	108.0(1)	N(1)-Mo(1)-N(2)	109.3(2)

V Bis(t-butylimido)-[1,2-bis(2,4,6-trifluoromethylphenyl)diphosphene]-trimethylphosphine- Molybdenum

Metal Mo in this complex has five coordinated ligands. It can be seen as a distorted square-based pyramid around the metal atom. The influence on the geometry

Empirical formula	$C_{29}H_{31}N_2P_3F_{18}Mo$	
Formula weight	938.41	
Temperature	150(2)K	
Wavelength	0.71073Å	
Crystal system	Triclinic	
Space group	$P\bar{1}$	
Unit cell dimensions	a = 12.265(6)Å	α = 66.45(4)°
	b = 12.448(8)Å	β = 88.01(4)°
	c = 14.760(7)Å	γ = 66.37(4)°
Volume	1871(2) $Å^3$	
Z	2	
Crystal colour	Green	
Absorption Coefficient	0.539 mm^{-1}	
F(000)	936	
Crystal size	0.80 × 0.75 × 0.40 mm	
θ range for data collection	1.56 → 25.00°	
Index range	$-14 \le h \le 13, -14 \le k \le 0, -17 \le l \le 16$	
Experiment device	Rigaku AFC6S	
Reflections collected	6583	
Independent reflections	6583	
Refinement method	Full-matrix least-squares on F^2	
Data/restraints/parameters	6578/0/505	
Goodness-of-fit on F^2	1.032	
Final R indices [I > 2σ(I)]	R1 = 0.0303, wR2 = 0.0799	
R indices (all data)	R1 = 0.0453, wR2 = 0.1044	
Largest diff. Peak and hole	0.893 and -0.633 $eÅ^{-3}$	

for this coordination sphere is from the bidentate ligand, diphosphene, with tight bite angle, P(2)-Mo-P(3) 49.82(4)°. This angle is only about half of the normal value in this geometry.

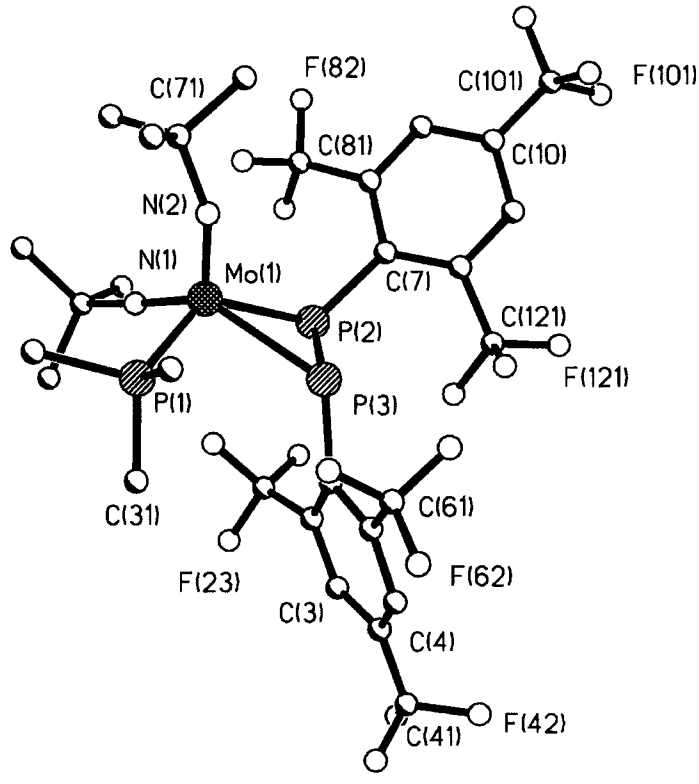


Figure 6.6 Structure diagram of complex V

On the other hand, the bulkier disphosphene ligand forces the other bond angles around the Mo to contract compared with analogue bis-imidomolybdenum alkene complex (Dyer *et al*, 1995), such that angle P(1)-Mo-N(1) is 101.9(1)° in the alkene and 95.99(9) °in the diphosphene complex, angle P(1)-Mo-N(2) reduce from 99.4(1) ° to 95.58(9) °and angle N(1)-Mo-N(2) from 123.0(2) ° to 115.90(1) °.

Table 6.5 Selected bond lengths(Å) and angles(°) for coordination sphere of V

Mo(1)-N(2)	1.763(3)	Mo(1)-N(1)	1.757(3)
Mo(1)-P(1)	2.475(2)	Mo(1)-P(2)	2.538(2)
Mo(1)-P(3)	2.588(2)	N(2)-Mo(1)-P(1)	95.6(1)
N(2)-Mo(1)-P(2)	116.0(1)	N(2)-Mo(1)-P(3)	110.3(1)
N(2)-Mo(1)-N(1)	115.9(1)	N(1)-Mo(1)-P(1)	96.0(1)
N(1)-Mo(1)-P(2)	99.5(1)	N(1)-Mo(1)-P(3)	132.8(1)
P(1)-Mo(1)-P(2)	133.17(4)	P(1)-Mo(1)-P(3)	88.2(1)
P(2)-Mo(1)-P(3)	49.82(4)		

VI Dichloro-oxo-[1,1-bis(3,5-di-t-butyl-2-hydroxyphenyl)ethane]-tungsten

This is the first crystal structure of a1,1-bis(3,5-di-t-butyl-2-hydroxyphenyl)ethane tungsten complex. Tungsten coordinates with two oxygen atoms of hydroxyphenyl and forms an 8-membered chelate ring with the ligand, but the chelate effect on the geometry of coordination sphere is not obvious. A regular square-based pyramid can be observed (shown in Figure 6.7).

Empirical formula	$C_{30}H_{44}O_3Cl_2W$
Formula weight	707.40
Temperature	150(2)K
Wavelength	0.71073Å
Crystal system	Monoclinic
Space group	$P2_1/n$
Unit cell dimensions	a = 14.582(3)Å α = 90.0º
	b = 16.582(3)Å β = 113.32(3)º
	c = 15.664(3)Å γ = 90.0º
Volume	3478.1(1) $Å^3$
Z	4
Crystal colour	Black
Absorption Coefficient	3.499 mm^{-1}
F(000)	1424
Crystal size	0.60 × 0.40 × 0.20 mm
θ range for data collection	2.74 → 25.00°
Index range	$-1 \le h \le 17$, $-1 \le k \le 19$, $-18 \le l \le 17$
Experiment device	Rigaku AFC6S
Reflections collected	7404
Independent reflections	6114 [$R_{int.}$= 0.0385]
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	6113/0/535
Goodness-of-fit on F^2	1.032
Final R indices [I > 2σ(I)]	R1 = 0.0338, wR2 = 0.0749
R indices (all data)	R1 = 0.0742, wR2 = 0.0886
Largest diff. Peak and hole	0.935 and -0.877 $eÅ^{-3}$

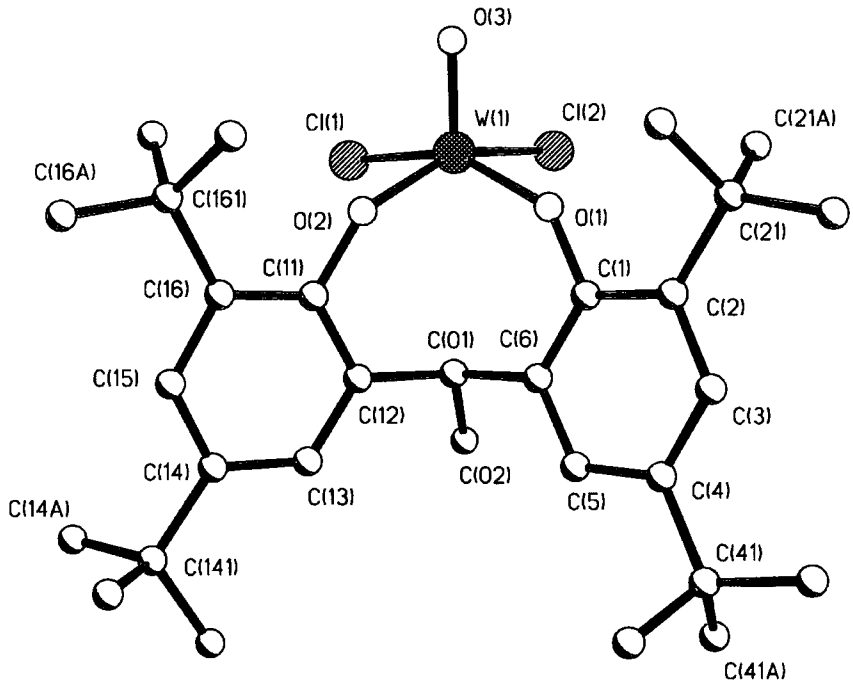


Figure 6.7 Molecular structure of complex VI

All the bond lengths of M-L fall in the normal values compared with other analogous bonds in different compounds. Oxo oxygen O(3) can be seen as at the apical position of the square-based pyramid and two chlorine and two oxygen at *cis* position respectively form the "square base". The angles between apical atom and base atoms [O(3)-W-L] are all close to the standard value 105.0° and the angles between base atoms also have good agreement in those standard values 86.0°. The largest deviation ~3° is observed from the angle O(1)-W(1)–O(2), 89.2(1) °. Therefore, this is a good example of a square-based pyramid.

Table 6.6 Geometry parameters for W-coordination sphere in complex **VI**

W(1)-Cl(1)	2.332(3)	W(1)-Cl(2)	2.311(2)	W(1)-O(1)	1.857(2)
W(1)-O(2)	1.869(2)	W(1)-O(3)	1.679(2)	Cl(1)-W(1)-Cl(2)	85.2(1)
Cl(1)-W(1)-O(1)	151.6(2)	Cl(1)-W(1)-O(2)	85.3(1)	Cl(1)-W(1)-O(3)	105.0(2)
Cl(2)-W(1)-O(1)	87.4(2)	Cl(2)-W(1)-O(2)	153.4(1)	Cl(2)-W(1)-O(3)	103.3(2)
O(1)-W(1)-O(2)	89.2(1)	O(1)-W(1)-O(3)	103.4(2)	O(2)-W(1)-O(3)	103.2(1)

VII Dichloro-[bis(diphenylphosphino)-N,N'-dimethylethylenediamine-P,P']-platinum

The coordination number in this complex is four. It shows a clear square planar geometry for the coordination sphere of the metal Pt (Figure 3.8). The chelate ring (7-membered) formed by two phosphorous atoms with the metal has a larger angle 95.46(2)° than the other side formed by two chlorines, Cl(1) – Pt(1) – Cl(2), 88.54(2)°.

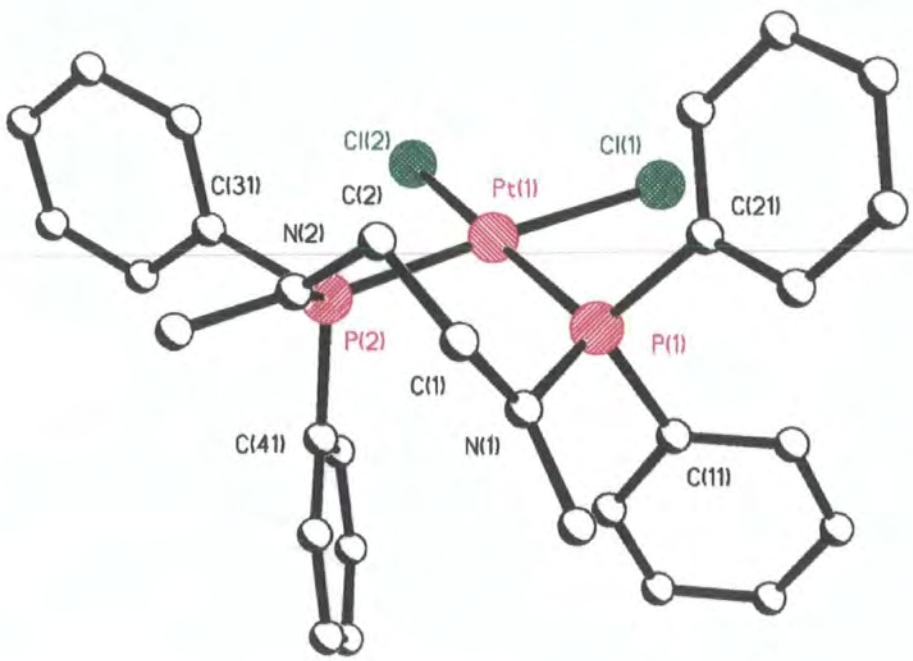
Empirical formula	$C_{28}H_{30}N_2P_2Cl_2Pt \cdot CH_2Cl_2$
Formula weight	807.40
Temperature	150(2)K
Wavelength	0.71073Å
Crystal system	Monoclinic
Space group	$P2_1/n$
Unit cell dimensions	a = 10.6635(1)Å α = 90.0º
	b = 16.8728(2)Å β = 95.10(1)º
	c = 17.0720(2)Å γ = 90.0º
Volume	3059.50(6) $Å^3$
Z	4
Crystal colour	Colourless
Absorption Coefficient	5.063mm^{-1}
F(000)	1584
Crystal size	0.44 × 0.30 × 0.26 mm
θ range for data collection	1.70 → 27.47°
Index range	$-13 \le h \le 13$, $-21 \le k \le 21$, $-22 \le l \le 15$
Experiment device	Siemens SMART
Reflections collected	21902
Independent reflections	6987 [R_{int} = 0.0215]
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	6984/0/430
Goodness-of-fit on F^2	1.209
Final R indices [I > 2σ(I)]	R1 = 0.0213, wR2 = 0.0515
R indices (all data)	R1 = 0.0268, wR2 = 0.0544

The geometry parameters are given in Table 6.7. Four ligands have good co-planarity (deviations less than 0.02(2)Å from the least-square plane) and the Pt atom is out of the plane by 0.10(2) Å.

Table 6.7 Bond lengths(Å) and angles(º) for coordination sphere of **VII**

Pt(1)-Cl(1)	2.371(1)	Ni(1)-Cl(2)	2.367(1)
Pt(1)-P(1)	2.250(1)	Ni(1)-P(2)	2.230(1)
Cl(1)-Ni(1)-Cl(2)	88.54(2)	Cl(1)-Ni(1)-P(1)	85.98(2)
Cl(1)-Ni(1)-P(2)	175.70(2)	Cl(2)-Ni(1)-P(1)	171.95(2)
Cl(2)-Ni(1)-P(2)	89.60(2)	P(1)-Ni(1)-P(2)	95.46(2)

Figure 6.8 Plot of complex Dichloro-[bis(diphenylphosphino)-N,N'-dimethylethylenediamne-P,P']-platinum



VIII Bis(t-butylimido)-bis[2,4,6-tris(trifluoromethyl)phenyl]-molybdenum

This structure has a distorted tetrahedral geometry with N-Mo-N and C-Mo-C angles of 110.6°and 138.3° respectively. The latter is considerably distorted from the standard angles of the tetrahedron due to the larger size of the fluoromethyl groups (Dillon, *et al*, 1997). Additionally, fluorines on two of the *ortho*-CF_3 groups lie in close contact with the Mo centre with Mo ··· F distances of 2.467(3) and 2.476(3) Å.The CSD

Empirical formula	$C_{26}H_{22}N_2F_{18}Mo$
Formula weight	800.40
Temperature	150(2)K
Wavelength	0.71073Å
Crystal system	Monoclinic
Space group	C2/c
Unit cell dimensions	a = 18.721(7)Å α = 90.0º b = 18.622(8)Å β = 92.66(3)º c = 17.652(6)Å γ = 90.0º
Volume	6147(4) $Å^3$
Z	8
Crystal colour	Yellow
Absorption Coefficient	0.557 mm^{-1}
F(000)	3168
Crystal size	0.40 × 0.30 × 0.20 mm
θ range for data collection	2.74 → 25.00°
Index range	$-1 \le h \le 22$, $-1 \le k \le 22$, $-20 \le l \le 20$
Experiment device	Rigaku AFC6S
Reflections collected	6388
Independent reflections	5422 [$R_{int.}$ = 0.1039]
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	5387/0/506
Goodness-of-fit on F^2	0.950
Final R indices [I > 2σ(I)]	R1 = 0.0438, wR2 = 0.1121
R indices (all data)	R1 = 0.0882, wR2 = 0.1957
Largest diff. Peak and hole	0.886 and -0.803 $eÅ^{-3}$

Table 6.8 Bond lengths(Å) and angles(°) for coordination sphere of **VIII**

Mo(1)-N(1)	1.724(4)	Mo(1)-N(2)	1.729(2)
Mo(1)-C(11)	2.252(5)	Mo(1)-C(21)	2.241(3)
Mo(1)-F(161)	2.467(3)	Mo(1)-F(221)	2.476(3)
N(1)-Mo(1)-N(2)	110.6(2)	N(1)-Mo(1)-C(21)	111.0(2)
N(1)-Mo(1)-C(11)	93.6(2)	C(11)-Mo(1)-C(21)	138.3(2)
N(1)-Mo(1)-F(161)	92.2(2)	N(1)-Mo(1)-F(221)	159.0(2)
N(2)-Mo(1)-F(161)	156.6(2)	N(2)-Mo(1)-F(221)	89.8(2)
C(11)-Mo(1)-F(161)	71.8(2)	C(21)-Mo(1)-F(161)	74.1(2)
C(11)-Mo(1)-F(221)	74.1(2)	C(21)-Mo(1)-F(221)	71.4(2)
F(161)-Mo(1)-F(221)	68.0(1)		

search on 114 Mo-F complexes shows that the mean value for Mo-F bond is 1.998Å. Therefore, the Mo···F in this structure can be considered as weak interactions. The coordination sphere, indeed, can be seen as an intermediate transformed from a tetrahedron to octahedron by a reaction pathway 4+2 → 6 (Drew, 1977). The major geometrical parameters are listed in table 6.8 and structure diagram is illustrated in Figure 6.9.

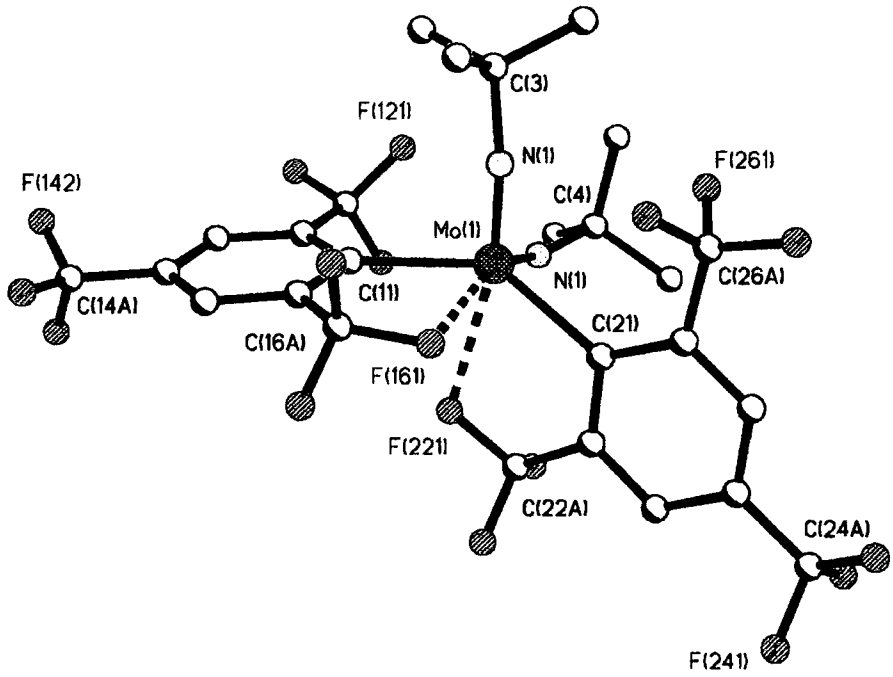


Figure 6.9 Plot of complex VIII

IX Dichloro-bis(*2,4,6*-trifluoromethylphenylimido)-bis-pyridine-molybdenum

Structure IX has also six coordinated ligands, a pair of nitrogen atoms from pyridine and the other from imido, two chlorine atoms sit at the trans-position (see Figure 6.10).The four N atoms can be seen located in the equatorial plane of octahedron

Empirical formula	$C_{28}H_{14}N_4F_{18}Cl_2Mo$	
Formula weight	915.27	
Temperature	150(2)K	
Wavelength	0.71073Å	
Crystal system	Monoclinic	
Space group	$P2_1/n$	
Unit cell dimensions	a = 8.841(1)Å	α = 90.0º
	b = 17.120(2)Å	β = 96.77(1)º
	c = 21.634(2)Å	γ = 90.0º
Volume	3251.6(11) $Å^3$	
Z	4	
Crystal colour	Black	
Absorption Coefficient	0.863 mm^{-1}	
F(000)	2030	
Crystal size	0.36 × 0.30 × 0.20 mm	
θ range for data collection	1.52 → 23.26°	
Index range	$-9 \le h \le 7, -15 \le k \le 18, -21 \le l \le 24$	
Experiment device	Siemens Smart CCD	
Reflections collected	12069	
Independent reflections	4582 [$R_{int.}$= 0.0577]	
Refinement method	Full-matrix least-squares on F^2	
Data/restraints/parameters	4582/0/521	
Goodness-of-fit on F^2	1.277	
Final R indices [I > 2σ(I)]	R1 = 0.0434, wR2 = 0.0976	
R indices (all data)	R1 = 0.0477, wR2 = 0.1001	
Largest diff. Peak and hole	0.321 and -0.621 $eÅ^{-3}$	

and Cl atoms at axial positions. Since N(1), N(2) have shorter distances with Mo atoms (double bonds) than those of N(3), N(4), two Cl atoms incline to N(3), N(4) side,

which gives the angle 157.4° for Cl(1)-Mo-Cl(2), and deviates considerably from standard angle of 180° in this geometry. The largest deviation on the plane is the angle between N(1) and N(2), N(1)-Mo-N(2) 102.4°. It may be caused by the bulky ligands.

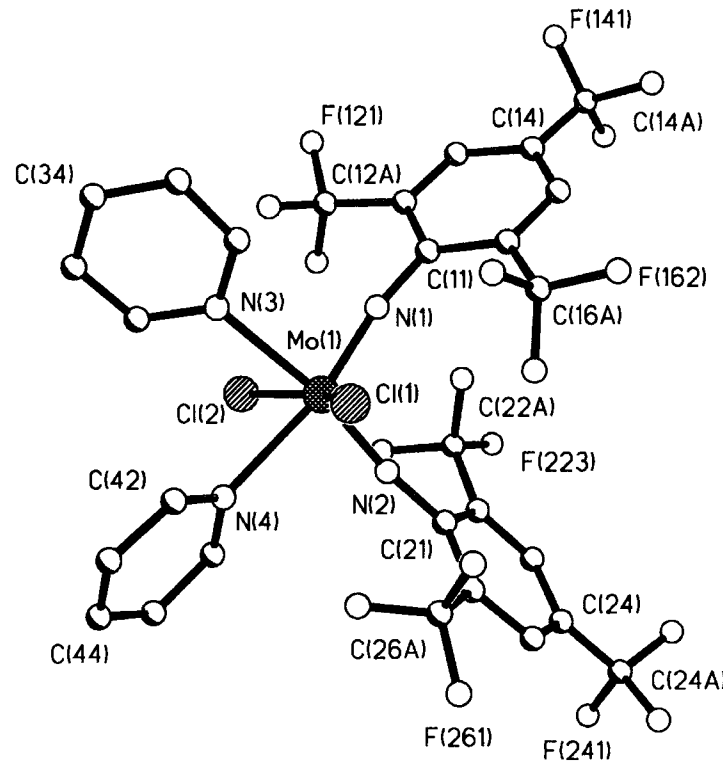


Figure 6.10 Structure of $Mo(C_5H_5N)_2[NC_6H_2(CF_3)_3]_2Cl_2$

Table 6.9 Bond lengths (Å) and angles for the coordination sphere in structure **IX**

Mo(1)-Cl(1)	2.374(1)	Mo(1)-Cl(2)	2.378(3)	Mo(1)-N(1)	1.774(4)
Mo(1)-N(2)	1.775(4)	Mo(1)-N(3)	2.354(4)	Mo(1)-N(4)	2.385(4)
Cl(1)-Mo(1)-Cl(2)	154.7(1)	Cl(1)-Mo(1)-N(1)	98.6(1)	Cl(1)-Mo(1)-N(2)	96.8(1)
Cl(1)-Mo(1)-N(3)	80.1(1)	Cl(1)-Mo(1)-N(4)	81.5(1)	Cl(2)-Mo(1)-N(1)	98.0(1)
Cl(2)-Mo(1)-N(2)	98.0(1)	Cl(2)-Mo(1)-N(3)	81.8(1)	Cl(2)-Mo(1)-N(4)	78.5(1)
N(1)-Mo(1)-N(2)	102.4(2)	N(1)-Mo(1)-N(3)	87.6(2)	N(1)-Mo(1)-N(4)	168.9(2)
N(2)-Mo(1)-N(3)	169.8(2)	N(2)-Mo(1)-N(4)	88.6(2)	N(3)-Mo(1)-N(4)	81.5(1)

X. Bis(μ^2-methoxo)-tetra-bis(2,6-di-isopropylphenyl)imido-dimethyl-di-molybdenum

The crystal structure has shown that this is a dimer of a Mo complex. Two Mo atoms are bridged by two methoxy molecules. Each Mo atom coordinates with two imido ligands and a methyl molecule. Both Mo atoms have five-coordination ligands. Each Mo-coordination sphere can be described as square pyramid, in which N(2) or N(2A) are the apical atoms, N(1), O(1), O(1A) and C(3), or N(1A), O(1), O(1A)

Empirical formula	$C_{52}H_{46}N_4O_2Mo_2$
Formula weight	950.84
Temperature	150(2)K
Wavelength	0.71073Å
Crystal system	Monoclinic
Space group	$P2_1/c$
Unit cell dimensions	a = 12.629(1)Å α = 90.0º b = 10.188(2)Å β = 97.67(1)º c = 20.529(2)Å γ = 90.0º
Volume	2617.6(9) $Å^3$
Z	2
Crystal colour	Dark green
Absorption Coefficient	0.988 mm^{-1}
F(000)	1296
Crystal size	0.34 × 0.26 × 0.20 mm
θ range for data collection	1.63 → 26.07°
Index range	$-13 \le h \le 15, -9 \le k \le 12, -25 \le l \le 23$
Experiment device	Siemens Smart CCD
Reflections collected	11213
Independent reflections	4500 [$R_{int.}$= 0.0486]
Refinement method	Full-matrix least-squares on F^2
Data/restraints/parameters	4500/0/376
Goodness-of-fit on F^2	1.280
Final R indices [I > 2σ(I)]	R1 = 0.0494, wR2 = 0.1004
R indices (all data)	R1 = 0.0692, wR2 = 0.1263
Largest diff. Peak and hole	0.660 and -0.499 $eÅ^{-3}$

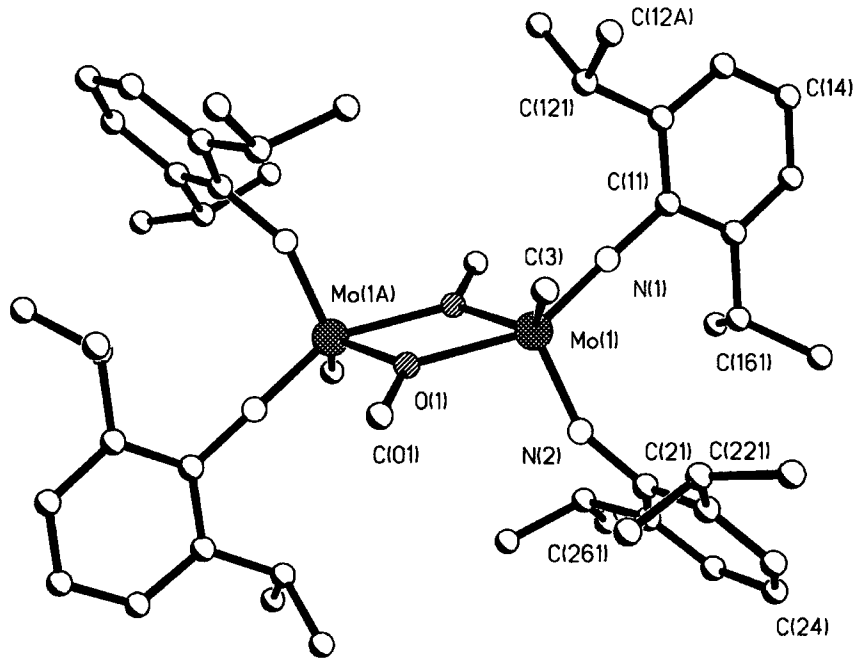


Figure 6.11 View of complex X.

and C(3A) atoms constitute the square base.

There is only half molecule contained in an asymmetric unit, the other half can be located by the inversion centre of $P2_1/c$. The distance of Mo(1)-Mo(1A) is 3.486(2)Å, which is similar to other dimer of Mo complexes(Chisholm, Cotton, Extine and Kelly, 1978). The major deviations can be found on bridging oxygen atoms, which have a smaller and a larger angle with apical atom N(2), N(2)-Mo-O(1), 99.9(2)°; N(2)-Mo-O(1)*, 117.6(2) °, compared with the standard ones. The angle between two oxygen atoms is also smaller, 68.7(1) °, than a normal value in a standard square pyramid.

Table 6.10 Selected bond lengths(Å) and angles(º) for Mo-coordination sphere. The same values in the other half molecule.

Mo(1)-N(1)	1.747(4)	Mo(1)-N(2)	1.760(4)
Mo(1)-O(1)	2.163(3)	Mo(1)-O(1)*	2.059(3)
N(1)-Mo(1)-N(2)	108.0(1)	N(1)-Mo(1)-O(1)*	96.2(2)
N(1)-Mo(1)-O(1)	152.0(2)	N(1)-Mo(1)-C(3)	91.9(2)
N(2)-Mo(1)-O(1)*	117.6(2)	N(2)-Mo(1)-O(1)	99.9(2)
N(2)-Mo(1)-C(3)	102.8(2)	O(1)-Mo(1)-C(3)	83.3(2)
O(1)*-Mo(1)-C(3)	133.5(2)	O(1)-Mo(1)-O(1)*	68.7(1)

* Symmetry transformations used to generate equivalent atoms 1-x, -y, -z.

References

Arndt, U. W. and Willis, B. T. M., *Single Crystal Diffractometry*, Cambridge University Press, Cambridge, UK, (1966).

Chisholm, M. H., Cotton, F. A. Cotton, Extine, M. W. and Kelly, R. L., *J. Am. Chem. Soc.*, **100**, 3354, (1978).

Cochran, W. and Woolfson, M. M., *Acta Cryst.*, **8**, 1, (1955)

Debaerdemaeker, T., Tate, C. and Woolfson, M. M., *Acta Cryst.*, A**41**, 353, (1988).

Dillon, K.B, Gibson, V. C., Howard, J. A. K., Redshaw, C., Sequeria, L. and Yao, J.W., *J. Organometallic Chem.*, **528**, 179, (1997).

Drew, M. G. B., *Prog. Inorg. Chem.*, **23**, 1, (1977).

Dyer, P. W., Gibson, V. C., Howard, J. A. K., Whittle, B. and Wilson, C., *Polyhedron*, **14**, 103, (1995).

Giacovazzo, S. C., Monaco, H. L., Viterbo, D., Scordari, F., Gilli, G., Zanotti, G and Catti, M., *Fundamentals of Crystallography*, OUP, Oxford, UK, (1992).

Glusker, J. P. with Lewis, M., Rossi, M., *Crystal Structure Analysis for Chemists and Biologists*, VCH publishers, Inc., New York, USA, (1994).

Glusker, J. P. and Trueblood, K. N., *Crystal Structure Analysis*, 2nd Edition, New York, USA, (1985).

Haker, D. and Kasper, J. S., *Acta Cryst.*, **1**, 70, (1948).

Karle, J. and Hauptman, H., *Acta Cryst.*, **9**, 635, (1956)

Karle, I. L. and Karle, J., *Acta Cryst.* **17**, 1356, (1964a).

Karle, I. L. and Karle, J., *Acta Cryst.* **17**, 835, (1964b).

McKie, D. and McKie, C., *Essentials of Crystallography,* Blackwell Scientific Publications, London, UK, (1986).

Overbeek, A. R. and Schenk, H., in *Computing in Crystallography*, Edited by H. Schenk, R. Olthof-Hazekamp, H. Van Koningsveld & G. C. Bassi, Delf Univ. Press, Holland, (1978).

Patterson, A. L., *Phys. Rev.*, **46**, 372, (1934).

SAINT Software Reference Manual, Siemens Energy & Automatic Inc. Analytical Instrumentation, Wisconsin, USA, (1995).

Sayre, D., *Acta Cryst.*, **5**, 60, (1952).

Sheldrick, G. M., *SHELXL-93*, Univ. of Göttingen, Germany, (1993).

Woolfson, M. M., *Acta Cryst.*, **11**, 393, (1958).

Woolfson, M. M., *Acta Cryst.*, A**33**, 219, (1977).

Woolfson, M. M., *An Introduction to X-ray Crystallography*, 2nd Edition, Cambridge University Press, Cambridge, UK, (1997).

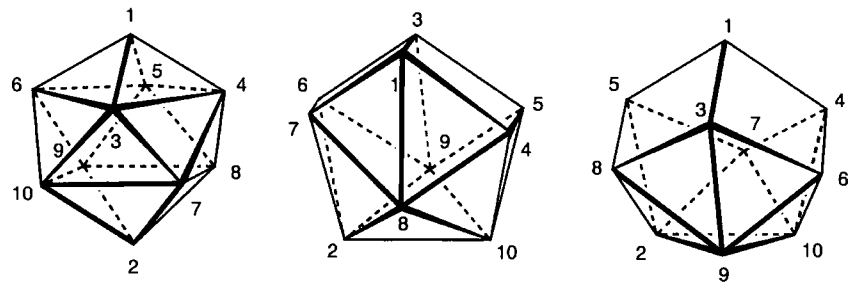
Yao, J-X., *Acta Cryst.*, A**37**, 642, (1981).

Chapter 7

Further Work

7.1 Geometry in Higher Coordination Number Spheres

The geometry analyses of the higher coordination number transition metal complexes has now been completed up to the 9-coordination level. Two other even higher coordination numbers are also observed in lanthanide or main group metal complexes, that is, 10- and 12-coordination. The basic geometries for these two coordination spheres are shown in Figure 7.1.



BSAP SPC TRH

(a)

Figure 7.1 (a) Geometries in 10-coordination

Three polyhedral forms, bicapped square antiprism (BSAP), sphenocorona (SPC) and trirhombohedron (TRH) exist in 10-coordination and four, icosahedron (ICOH), bicapped pentagonal prism (BCPP), anticuboctahedron (ACOT) and cuboctahedron (COCT) exist in 12-coordination.

With the increase of coordination number, the valence angles L-M-L have smaller differences. $R_{ang}(x)$ values may become unable distinguishable the different geometry models. This has been seen in 9-coordination, especially for those complexes with

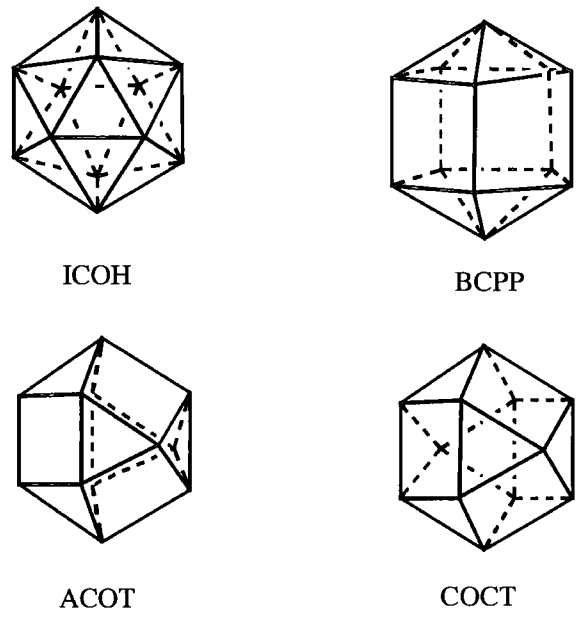


Figure 7.1 (b) Geometries in 12-coordination

multidentate ligands in which the valence angles, L-M-L are distorted due to the strains of the chelate rings.

It can be seen that information might be lost using the single $R_{ang}(x)$ value as a criterion to involve more angle parameters. Appropriate methods which are more sensitive to the smaller difference within these geometries in higher coordination spheres must be explored or more relevant parameters for these higher coordination must be applied. The interplanar angle δ between the dihedral planes in a polyhedron is another important geometrical parameter in characterizing different polyhedra. It seems to have a larger variable range than the conventional polyhedral angles. Using this value instead of the L-M-L angles may be more suitable for classifying the geometry of higher coordination. It may also be possible to extend this method to apply to geometry study of metal clusters or borane cages compounds. The importance in these applications also includes setting up standard geometries for different clusters or cages.

7.2 Study in Reaction Pathways

Another interest using systematic analysis is to examine or confirm some given reaction pathways using the available structural information. The deformation of a molecule from ideal states may reflect the distortion which these fragments would undergo along a given reaction coordinate.

The reaction mechanisms concerned with 7-coordination usually regards the 7-coordination complexes as intermediates. A 7-coordination intermediate may arise during reaction of 6-coordination complexes when an extra bond is formed or during reaction related to 8-coordination complexes when a bond is broken.

Considering a substituent reaction in a 6-coordination complex, one possibility is that the substituent ligand (Y) attacks to the 6-coordination complex XML_5 to form a 7-coordination intermediate, $[XML_5Y]$, and then X is eliminated and a new substituted 6-coordination complex ML_5Y is formed (Drew, 1977*). i.e.

$$XML_5 + Y \Leftrightarrow [XML_5Y] \Leftrightarrow X + ML_5Y$$

From the consideration of reaction pathways, the formation of idealized 7-coordination geometries can be seen as the addition of an extra ligand to the regular octahedron. Attack of ligand along an edge (E) of the square plane of octahedron will result in the other four ligands, on the vertices coplanar with this new ligand, being slightly rearranged so that five ligands sit on the vertices of a pentagon, which forms a 7-coordination PBP. Face attack (F) may generate either COC or CTP according to the movements of other vertices of the octahedron caused by the attack, shown as in Figure 7.2

Preliminary work to confirm this path has been undertaken by investigating the position of an extra non-bonded atom related to the central atom in 6-coordination complexes. This position is simulated as the optimal direction of the 7th atom attacking to the octahedron.

* This reference has been given in Chapter 3

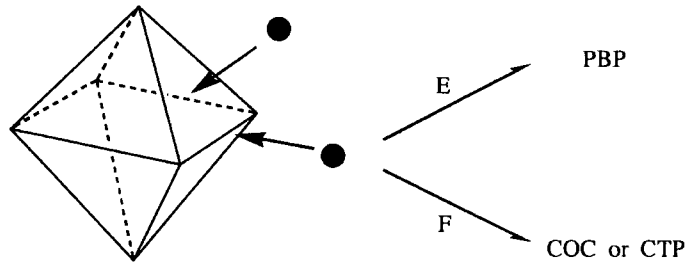


Figure 7.2 One-ligand attacks on octahedron to generate PBP, COC or CTP

This was carried out through a CSD search of 6-coordination complexes with a non-bonded atom and defining its relevant angles and non-bonded distance to the central metal atom. The non-bonded distances are subsequently normalized by subtracting the sum of the covalent radii of the relevant bonded atoms. If a non-bonded atom lies in the square base plane of octahedron and has angle of zero with the metal atom, that means that the 7th atom is approaching the metal atom from the edge (E). The results have shown that all non-bonded atoms locate between 0° and 60° from this plane (shown in Figure 7.3), which shows that the majority of the contacts approach the metal centre along a vector between the quadraplane and a face center of octahedra, and thus forming either "PBP" or "COC/CTP" 6+1 coordinated compounds.

Meanwhile, let us assume at the position 3 in 7-coordination species (see Figure 3.1) is the final standing position of the attacking atom. Thus, the data for 7-coordination species are also used together with the data of 6-coordination non-bonded species of Co complexes. For comparison, the angle of ligand 3 away from the equatorial plane is calculated and plotted with 6-coordination non-bonded contact data. It shows that they all fall into the same angle range. But in the case of 7-coordination, since ligand 3 is bound to the metal atom, they have shorter normalised distances and then appear in the inner circle. This has proved that the 6+1 coordination metal data coincide with the data that could be similarly derived for the metals that exist as exactly 7-coordination. This means that it is possible to use these data to further explore the relationships from 6+1 coordination leading to 7-coordination by geometrical similarity.

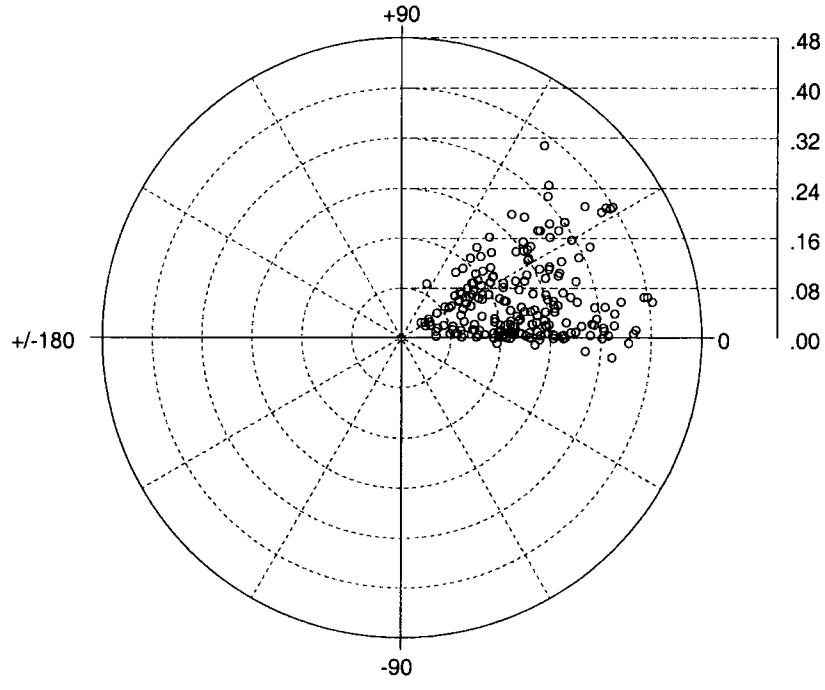


Figure 7.3 Polar scatterplot of angle *vs.* normalised distance between the metal atom and the non-bonded atoms in 6-coordination species, and L_3 in 7-coordination species.

The further study will be to find a meaningful reaction coordinates system to map the reaction pathways.

There also exist other possibilities of reaction pathways related to the 7-coordination complexes, such as, 7+1 to 8-coordination or 8-1 to 7-coordination (addition/elimination reaction) and so on. All these are to be considered in the same manner as those in 6+1 to 7-coordination.

Appendix I.

Atomic Coordinates

Atomic coordinates for the structures given in Chapter 6. U_{eq} is equivalent isotropic displacement parameters ($Å^2$), which is defined as: $U_{eq} = \frac{1}{3}\sum_i \sum_j U_{ij} a_i^* a_j^* \mathbf{a}_i \cdot \mathbf{a}_j$.

I. Tri-carbonyl-tris(methylcyanide)-chromium [$Cr(CO)_3(CH_3CN)_3 \cdot CH_3CN$]

Atom	*x*	*y*	*z*	U_{eq}
Cr(1)	.1441(1)	.3166(1)	.7423(2)	.021(1)
O(1)	-.0252(4)	.4086(8)	.8752(3)	.046(2)
O(2)	-.0289(5)	.4087(9)	.6088(3)	.048(2)
O(3)	.0405(1)	-.0116(2)	.7413(7)	.052(1)
N(1)	.2667(4)	.2580(9)	.8355(4)	.029(1)
N(2)	.2150(1)	.5485(2)	.7418(6)	.028(1)
N(3)	.2649(4)	.2550(9)	.6491(4)	.029(1)
C(1)	.3369(6)	.2228(10)	.8805(5)	.035(2)
C(2)	.2512(1)	.6757(2)	.7409(9)	.027(1)
C(3)	.3385(5)	.2239(10)	.6059(5)	.033(2)
C(4)	.0843(2)	.1144(2)	.7393(8)	.030(1)
C(5)	.0415(4)	.3757(9)	.8265(4)	.026(1)
C(6)	.0448(6)	.3722(10)	.6591(4)	.033(2)
C(11)	.4286(8)	.1843(10)	.9327(7)	.055(2)
C(12)	.2978(2)	.8380(2)	.7424(10)	.038(1)
C(13)	.4308(6)	.1800(10)	.5452(7)	.050(20)
C(01)	-.2520(9)	.2439(3)	.4913(9)	.048(1)
C(02)	-.2502(9)	.4197(3)	.4912(9)	.041(1)
N(01)	-.2641(11)	.1070(3)	.4938(11)	.085(2)

II. Trans-bis[1,2-bis(dimethylphosphino)cyclopentane]-dichloroiron

Atom	*x*	*y*	*z*	U_{eq}
Fe(1)	.0000	.0000	.5000	.018(1)
P(1)	.0655(1)	.1189(1)	.4078(1)	.025(1)
P(2)	.1490(1)	-.0542(1)	.3298(1)	.025(1)
Cl(1)	.2451(1)	.0053(1)	.6706(1)	.030(1)
C(11)	.1879(7)	.1910(2)	.5240(5)	.066(2)
C(12)	-.0966(6)	.1821(3)	.3143(7)	.089(2)
C(21)	.3383(4)	-.1115(2)	.3862(5)	.046(1)
C(22)	.0461(5)	-.1172(3)	.1826(4)	.059(1)
C(1)*	.2412(12)	.0963(4)	.2967(10)	.024(2)
C(5)	.1881(12)	.0235(4)	.1967(10)	.021(2)
C(1A)	.1497(13)	.1005(4)	.2352(10)	.025(2)
C(5A)	.2699(15)	.0312(4)	.2599(10)	.026(2)
C(2)	.2757(7)	.1584(2)	.1689(5)	.060(1)
C(3)	.3836(5)	.1079(2)	.0779(4)	.042(1)
C(4)	.3370(5)	.0211(2)	.1047(4)	.038(1)

* C(1) occup. 51%, C(1A) occup. 49%; C(5) occup. 51%, C(5A) occup. 49%.

III. Dichloro-(dimethoxyethane)-bis(t-butylimido)-molybdenum

Atom	*x*	*y*	*z*	U_{eq}
Mo(1)	-.0053(1)	.0890(1)	.3570(1)	.027(1)
Cl(1)	.0944(1)	-.0807(1)	.3766(1)	.040(1)
Cl(2)	-.0736(1)	.2395(1)	.3109(1)	.036(1)
O(1)	-.0451(3)	-.0103(2)	.2896(1)	.031(1)
O(2)	.1834(3)	.1012(2)	.3076(1)	.033(1)
N(1)	-.1683(3)	.0669(3)	.3759(1)	.032(1)
N(2)	.0793(3)	.1573(3)	.3986(1)	.033(1)
C(1)	-.1223(5)	-.1088(3)	.2922(2)	.038(1)
C(2)	.0729(5)	-.0250(4)	.2618(2)	.039(1)
C(3)	.1503(5)	.0779(4)	.2617(1)	.037(1)
C(4)	.2717(5)	.1921(4)	.3118(2)	.042(1)
C(01)	-.2967(4)	.0766(3)	.3996(1)	.032(1)
C(11)	-.2928(5)	.0042(4)	.4406(2)	.042(1)
C(12)	-.3163(5)	.1934(4)	.4129(2)	.042(1)
C(13)	-.4116(5)	.0408(5)	.3681(2)	.047(1)
C(02)	.1232(4)	.1971(3)	.4425(1)	.031(1)
C(21)	.2747(5)	.1720(5)	.4472(2)	.047(1)
C(22)	.0990(60	.3167(4)	.4434(2)	.046(1)
C(23)	.0423(5)	.1395(4)	.4786(2)	.043(1)

IV. Dichloro-bis(2,4,6-trfluoromethyl-phenylimido)-molybdenum

Atom	*x*	*y*	*z*	U_{eq}
Mo(1)	0.1508(1)	0.1313(1)	0.2479(1)	0.021(1)
Cl(1)	-0.1063(1)	0.1184(1)	0.2439(1)	0.037(1)
Cl(2)	0.2302(2)	-0.0807(1)	0.2763(1)	0.042(1)
N(1)	0.2734(4)	0.2895(4)	0.3278(2)	0.022(1)
N(2)	0.1737(4)	0.1618(4)	0.1464(2)	0.020(1)
C(11)	0.3837(5)	0.4329(5)	0.3612(2)	0.021(1)
C(12)	0.5457(5)	0.4550(5)	0.3628(2)	0.023(1)
C(13)	0.6540(5)	0.5999(5)	0.3934(3)	0.026(1)
C(14)	0.6033(5)	0.7239(5)	0.4243(2)	0.024(1)
C(15)	0.4453(5)	0.7029(5)	0.4249(3)	0.026(1)
C(16)	0.3350(5)	0.5589(5)	0.3943(2)	0.023(1)
C(1A)	0.6044(5)	0.3213(5)	0.3325(3)	0.031(1)
C(1B)	0.7227(6)	0.8796(5)	0.4554(3)	0.036(1)
C(1C)	0.1635(5)	0.5400(5)	0.3950(3)	0.027(1)
F(11)	0.5735(4)	0.2072(3)	0.3764(2)	0.045(1)
F(12)	0.5343(3)	0.2606(3)	0.2474(2)	0.038(1)
F(13)	0.7614(3)	0.3630(4)	0.3432(2)	0.054(1)
F(21)	0.8527(4)	0.8778(4)	0.5137(3)	0.086(1)
F(22)	0.7611(6)	0.9407(4)	0.3916(2)	0.094(1)
F(23)	0.6653(4)	0.9825(4)	0.4951(3)	0.075(1)
F(31)	0.1107(3)	0.4246(3)	0.4341(2)	0.039(1)
F(32)	0.0637(3)	0.5099(3)	0.3142(2)	0.038(1)
F(33)	0.1435(3)	0.6671(3)	0.4366(2)	0.049(1)
C(21)	0.2175(4)	0.2142(4)	0.0751(2)	0.020(1)
C(22)	0.2051(5)	0.3580(5)	0.0556(3)	0.023(1)
C(23)	0.2432(5)	0.4044(5)	-0.0179(3)	0.027(1)
C(24)	0.2934(5)	0.3083(5)	-0.0722(3)	0.028(1)
C(25)	0.3083(5)	0.1673(5)	-0.0535(3)	0.026(1)
C(26)	0.2698(4)	0.1184(4)	0.0194(3)	0.021(1)
C(2A)	0.1516(5)	0.4642(5)	0.1144(3)	0.028(1)
C(2B)	0.3282(5)	0.3581(5)	-0.1528(3)	0.039(1)
C(2C)	0.2800(6)	-0.0380(6)	0.0373(2)	0.027(1)
F(41)	0.2547(3)	0.5087(3)	0.1944(2)	0.034(1)
F(42)	0.0074(3)	0.3976(3)	0.1237(2)	0.035(1)
F(43)	0.1396(4)	0.5925(3)	0.0833(2)	0.041(1)
F(51)	0.3940(5)	0.5073(4)	-0.1415(2)	0.076(1)
F(52)	0.4300(5)	0.2911(6)	-0.1795(3)	0.090(1)
F(53)	0.1983(4)	0.3241(5)	-0.2191(2)	0.076(1)
F(61)	0.3874(3)	-0.0300(3)	0.1133(2)	0.034(1)
F(62)	0.3240(3)	-0.1162(3)	-0.0247(2)	0.040(1)
F(63)	0.1387(3)	-0.1257(3)	0.0403(2)	0.036(1)

V. Bis(t-butylimido)-[1,2-bis(2,4,6-trifluoromethylphenyl)diphosphene]-trimethylphosphine- molybdenum

Atom	*x*	*y*	*z*	U_{eq}
Mo(1)	0.2089(1)	0.2303(1)	0.3196(1)	0.015(1)
P(1)	0.2935(1)	0.1903(1)	0.1755(1)	0.022(1)
P(2)	0.1528(1)	0.4190(1)	0.3665(1)	0.018(1)
P(3)	0.3029(1)	0.3932(1)	0.2838(1)	0.018(1)
N(1)	0.0641(2)	0.2487(2)	0.2865(2)	0.021(1)
N(2)	0.3057(2)	0.0774(2)	0.4114(2)	0.020(1)
C(1)	0.2638(3)	0.5444(3)	0.1657(2)	0.019(1)
C(2)	0.1505(3)	0.6483(3)	0.1170(2)	0.020(1)
C(21)	0.0301(3)	0.6415(3)	0.1340(2)	0.024(1)
C(3)	0.1396(3)	0.7676(3)	0.0460(2)	0.026(1)
C(4)	0.2398(3)	0.7880(3)	0.0152(2)	0.030(1)
C(41)	0.2276(4)	0.9169(3)	-0.0619(3)	0.045(1)
C(5)	0.3526(3)	0.6858(3)	0.0526(2)	0.027(1)
C(6)	0.3645(3)	0.5677(3)	0.1255(2)	0.022(1)
C(61)	0.4926(3)	0.4639(3)	0.1598(2)	0.029(1)
C(7)	0.2213(2)	0.3766(3)	0.4936(2)	0.018(1)
C(8)	0.1914(3)	0.2946(3)	0.5797(2)	0.021(1)
C(81)	0.0910(3)	0.2547(3)	0.5732(2)	0.024(1)
C(9)	0.2452(3)	0.2559(3)	0.6752(2)	0.024(1)
C(10)	0.3267(3)	0.3010(3)	0.6897(2)	0.024(1)
C(101)	0.3874(3)	0.2582(3)	0.7921(2)	0.032(1)
C(11)	0.3468(3)	0.3923(3)	0.6093(2)	0.024(1)
C(12)	0.2934(3)	0.4322(3)	0.5130(2)	0.021(1)
C(121)	0.3144(3)	0.5434(3)	0.4343(2)	0.026(1)
C(31)	0.2499(3)	0.3306(3)	0.0566(2)	0.033(1)
C(32)	0.4567(3)	0.1093(3)	0.1911(3)	0.034(1)
C(33)	0.2464(3)	0.0817(3)	0.1515(3)	0.032(1)
C(51)	-0.0570(3)	0.2629(3)	0.2641(2)	0.026(1)
C(52)	-0.1003(3)	0.3410(4)	0.1518(3)	0.041(1)
C(53)	-0.1367(3)	0.3365(4)	0.3213(3)	0.037(1)
C(54)	-0.0575(3)	0.1301(3)	0.2977(3)	0.040(1)
C(71)	0.3384(3)	-0.0581(3)	0.4799(2)	0.026(1)
C(72)	0.3788(3)	-0.0780(3)	0.5848(2)	0.037(1)
C(73)	0.4429(3)	-0.1446(3)	0.4451(3)	0.038(1)
C(74)	0.2302(3)	-0.0893(3)	0.4787(3)	0.034(1)
F(21)	0.0348(2)	0.5239(2)	0.1593(1)	0.029(1)
F(22)	-0.0208(2)	0.6835(2)	0.2022(1)	0.033(1)
F(23)	-0.0503(2)	0.7192(2)	0.0488(1)	0.037(1)
F(41)	0.1214(3)	1.0073(3)	-0.0817(4)	0.153(1)
F(42)	0.3061(4)	0.9537(3)	-0.0351(2)	0.091(1)
F(43)	0.2599(3)	0.9138(2)	-0.1476(2)	0.065(1)
F(61)	0.5029(2)	0.3524(2)	0.1593(2)	0.038(1)
F(62)	0.5685(2)	0.4984(2)	0.0982(2)	0.039(1)
F(63)	0.5396(2)	0.4355(2)	0.2511(2)	0.040(1)
F(81)	-0.0110(2)	0.3575(2)	0.5195(1)	0.034(1)
F(82)	0.0639(2)	0.1975(2)	0.6630(1)	0.040(1)
F(83)	0.1145(2)	0.1721(2)	0.5310(2)	0.031(1)
F(101)	0.5066(2)	0.2094(4)	0.7988(2)	0.108(1)
F(102)	0.3675(2)	0.1655(3)	0.8618(2)	0.096(1)
F(103)	0.3604(3)	0.3503(2)	0.8207(2)	0.083(1)
F(121)	0.3238(2)	0.6200(2)	0.4749(2)	0.044(1)
F(122)	0.2262(2)	0.6218(2)	0.3566(1)	0.032(1)
F(123)	0.4180(2)	0.5045(2)	0.3978(2)	0.038(1)

VI. Dichloro-oxo-[1,1-bis(3,5-di-t-butyl-2-hydroxyphenyl)ethane]-tungsten

Atom	*x*	*y*	*z*	U_{eq}
W(1)	.2244(1)	.1549(1)	.0507(1)	.018(1)
Cl(1)	.2935(2)	.0784(1)	.1863(1)	.036(1)
Cl(2)	.3860(1)	.1632(2)	.0592(1)	.040(1)
O(1)	.2000(3)	.2527(3)	-.0121(3)	.023(1)
O(2)	.1230(3)	.1786(3)	.0914(3)	.021(1)
O(3)	.1706(4)	.0873(3)	-.0344(3)	.028(1)
C(01)	.2471(5)	.3134(4)	.1747(4)	.017(1)
C(02)	.3345(5)	.3516(5)	.2563(5)	.023(2)
C(1)	.1954(4)	.3353(4)	-.0012(4)	.019(1)
C(2)	.1685(5)	.3817(4)	-.0820(4)	.021(2)
C(3)	.1605(5)	.4652(4)	-.0701(5)	.021(1)
C(4)	.1768(5)	.5002(4)	.0160(4)	.020(1)
C(5)	.2029(5)	.4499(4)	.0930(4)	.019(1)
C(6)	.2139(4)	.3667(4)	.0870(4)	.018(1)
C(21)	.1506(5)	.3455(5)	-.1772(4)	.026(2)
C(21A)	.2451(6)	.3028(5)	-.1733(6)	.032(2)
C(21B)	.1222(7)	.4099(5)	-.2533(5)	.031(2)
C(21C)	.0628(6)	.2856(5)	-.2057(5)	.031(2)
C(41)	.1661(5)	.5915(4)	.0262(5)	.025(2)
C(41A)	.2625(6)	.6245(5)	.1012(6)	.031(2)
C(41B)	.0779(6)	.6071(6)	.0541(6)	.035(2)
C(41C)	.1454(9)	.6362(5)	-.0641(6)	.041(2)
C(11)	.1020(5)	.2195(4)	.1576(4)	.019(1)
C(12)	.1611(5)	.2868(4)	.2006(4)	.017(1)
C(13)	.1403(5)	.3264(4)	.2684(4)	.019(2)
C(14)	.0637(5)	.3006(4)	.2950(4)	.019(1)
C(15)	.0065(5)	.2350(4)	.2484(5)	.021(2)
C(16)	.0226(5)	.1931(4)	.1785(4)	.017(1)
C(141)	.0417(5)	.3411(4)	.3734(4)	.023(1)
C(14A)	.0620(7)	.2796(6)	.4522(6)	.036(2)
C(14B)	.1058(7)	.4160(6)	.4128(7)	.043(2)
C(14C)	-.0673(6)	.3680(5)	.3354(6)	.029(2)
C(161)	-.0410(5)	.1190(4)	.1308(5)	.021(2)
C(16A)	-.1303(6)	.1082(5)	.1576(6)	.031(2)
C(16B)	.0233(6)	.0415(5)	.1607(6)	.030(2)
C(16C)	-.0827(6)	.1256(5)	.0242(5)	.028(2)

VII. Dichloro-[bis(diphenylphosphino)-N,N'-dimethylethylenediamine-P,P']-platinum

Atom	*x*	*y*	*z*	U_{eq}
Pt(1)	.5726(1)	.3786(1)	.2018(1)	.017(1)
Cl(1)	.5963(1)	.5095(1)	.2514(1)	.026(1)
Cl(2)	.3530(1)	.3904(1)	.2103(1)	.027(1)
P(1)	.7795(1)	.3842(1)	.1856(1)	.019(1)
P(2)	.5421(1)	.2592(1)	.1472(1)	.019(1)
N(1)	.8467(2)	.3040(1)	.1470(2)	.022(1)
C(01)	.9511(4)	.3126(2)	.0965(2)	.033(1)
N(2)	.6481(2)	.1919(1)	.1766(2)	.022(1)
C(02)	.6379(3)	.1108(2)	.1460(2)	.027(1)
C(1)	.8693(3)	.2344(2)	.1986(2)	.024(1)
C(2)	.7524(3)	.2080(2)	.2359(2)	.022(1)
C(11)	.8006(3)	.4607(2)	.1126(2)	.021(1)
C(12)	.8763(3)	.5273(2)	.1265(2)	.030(1)
C(13)	.8894(3)	.5811(2)	.0655(2)	.037(1)
C(14)	.8268(3)	.5695(2)	-.0077(2)	.034(1)
C(15)	.7490(3)	.5043(2)	-.0212(2)	.033(1)
C(16)	.7358(3)	.4504(2)	.0387(2)	.023(1)
C(21)	.8805(3)	.4076(2)	.2745(2)	.021(1)
C(22)	.8323(3)	.4043(2)	.3481(2)	.025(1)
C(23)	.9114(3)	.4171(2)	.4163(2)	.029(1)
C(24)	1.0376(3)	.4338(2)	.4114(2)	.031(1)
C(25)	1.0868(3)	.4358(2)	.3390(2)	.031(1)
C(26)	1.0085(3)	.4216(2)	.2704(2)	.028(1)
C(31)	.3965(3)	.2073(2)	.1627(2)	.024(1)
C(32)	.3891(3)	.1766(2)	.2384(2)	.032(1)
C(33)	.2896(4)	.1278(2)	.2543(3)	.040(1)
C(34)	.1960(4)	.1096(2)	.1956(3)	.043(1)
C(35)	.2021(3)	.1407(2)	.1214(3)	.040(1)
C(36)	.3021(3)	.1893(2)	.1043(2)	.031(1)
C(41)	.5415(3)	.2740(2)	.0419(2)	.023(1)
C(42)	.4524(3)	.3270(2)	.0062(2)	.030(1)
C(43)	.4594(4)	.3504(2)	-.0711(2)	.036(1)
C(44)	.5536(4)	.3215(2)	-.1137(2)	.035(1)
C(45)	.6407(4)	.2680(2)	-.0794(2)	.033(1)
C(46)	.6355(3)	.2449(2)	-.0017(2)	.026(1)
C(1A)	1.4620(5)	.4000(3)	.4095(2)	.051(1)
Cl(3)	1.3811(1)	.4208(1)	.4930(1)	.045(1)
Cl(4)	1.5412(1)	.3081(1)	.4176(1)	.051(1)

VIII. Bis(t-butylimido)-bis[2,4,6-tris(trifluoromethyl)phenyl]-molybdenum

Atom	*x*	*y*	*z*	U_{eq}
Mo(1)	.7450(1)	.6158(1)	.2787(1)	.023(1)
N(1)	.6742(2)	.6489(2)	.3271(2)	.029(1)
N(2)	.8157(2)	.6756(2)	.2832(3)	.029(1)
C(11)	.6859(3)	.6227(3)	.1649(3)	.025(1)
C(12)	.6877(3)	.6776(3)	.1095(3)	.030(1)
C(13)	.6502(3)	.6734(3)	.0401(3)	.032(1)
C(14)	.6092(3)	.6133(3)	.0228(3)	.034(1)
C(15)	.6052(3)	.5578(3)	.0751(3)	.034(1)
C(16)	.6428(3)	.5641(3)	.1451(3)	.027(1)
C(12A)	.7343(3)	.7428(3)	.1211(3)	.035(1)
C(14A)	.5692(4)	.6085(4)	-.0527(4)	.052(2)
C(16A)	.6320(3)	.5033(3)	.1974(3)	.035(1)
F(121)	.7275(2)	.7747(2)	.1888(2)	.039(1)
F(122)	.7196(2)	.7943(2)	.0690(2)	.046(1)
F(123)	.8036(2)	.7272(2)	.1160(2)	.040(1)
*F(141)	.5230(21)	.5621(22)	-.0574(14)	.170(14)
F(142)	.5506(12)	.6691(10)	-.0802(11)	.086(7)
F(14A)	.5481(11)	.5432(7)	-.0699(10)	.072(6)
F(14B)	.5159(25)	.6506(25)	-.0594(19)	.177(18)
F(143)	.6162(4)	.6052(4)	-.1085(3)	.114(2)
F(161)	.6839(2)	.5002(2)	.2540(1)	.040(1)
F(162)	.5713(2)	.5073(3)	.2326(2)	.065(1)
F(163)	.6327(3)	.4398(2)	.1640(2)	.067(1)
C(21)	.8064(3)	.5340(3)	.3484(3)	.031(1)
C(22)	.8571(3)	.4927(3)	.3099(3)	.038(1)
C(23)	.9022(4)	.4433(4)	.3459(4)	.052(2)
C(24)	.8995(3)	.4318(4)	.4227(4)	.047(2)
C(25)	.8504(4)	.4694(4)	.4631(4)	.045(2)
C(26)	.8058(3)	.5193(3)	.4268(3)	.033(1)
C(22A)	.8669(4)	.5024(4)	.2282(4)	.053(2)
C(24A)	.9500(5)	.3795(5)	.4614(5)	.074(2)
C(26A)	.7555(3)	.5580(4)	.4761(3)	.042(2)
F(221)	.8116(2)	.5383(2)	.1925(2)	.040(1)
F(222)	.9245(2)	.5426(4)	.2132(3)	.103(2)
F(223)	.8734(4)	.4428(3)	.1901(3)	.119(3)
F(241)	.9544(3)	.3178(3)	.4224(4)	.106(2)
F(242)	1.0142(3)	.4034(3)	.4680(5)	.144(3)
F(243)	.9281(4)	.3561(4)	.5266(3)	.120(2)
F(261)	.7659(2)	.6292(2)	.4751(2)	.048(1)
F(262)	.6866(2)	.5466(2)	.4551(2)	.051(1)
F(263)	.7626(2)	.5384(3)	.5493(2)	.067(1)
C(3)	.6162(3)	.6935(3)	.3555(3)	.036(1)
C(31)	.5508(4)	.6431(5)	.3606(6)	.064(2)
C(32)	.6359(5)	.7249(6)	.4317(5)	.078(3)
C(33)	.5963(4)	.7495(6)	.2946(6)	.061(3)
C(4)	.8635(3)	.7307(3)	.3180(3)	.039(1)
C(41)	.8982(4)	.7730(4)	.2557(4)	.045(2)
C(42)	.8199(5)	.7804(4)	.3673(4)	.052(2)
C(43)	.9195(4)	.6901(5)	.3664(5)	.058(2)

* F(141) and F(142) are disordered
F(141) occup. 57%; F(14A) occup. 43%;
F(142) occup. 57%; F(14B) occup. 43%.

IX. Dichloro-bis(2,4,6-trifluoromethylphenylimido)-bis-pyridine-molybdenum

Atomic coordinates x, y, z ($\times 10^4$) and equivalent isotropic displacement U_{eq} ($\times 10^3$)

Atom	x	y	z	U_{eq}
Mo(1)	9600(1)	8507(1)	8696(1)	25(1)
Cl(1)	8574(1)	9241(1)	9479(1)	34(1)
Cl(2)	11475(1)	8183(1)	8036(1)	33(1)
N(1)	9303(4)	7556(2)	8984(2)	27(1)
N(2)	8030(4)	8620(2)	8112(2)	24(1)
N(3)	11764(4)	8591(2)	9440(2)	30(1)
C(11)	8886(5)	6803(3)	9099(2)	27(1)
C(12)	9847(5)	6160(3)	9015(2)	29(1)
C(13)	9341(6)	5405(3)	9083(2)	36(1)
C(14)	7899(6)	5252(3)	9252(2)	33(1)
C(15)	6987(6)	5871(3)	9373(2)	34(1)
C(16)	7450(5)	6632(3)	9311(2)	29(1)
C(12A)	11421(6)	6271(3)	8830(3)	39(1)
F(121)	12277(3)	5623(2)	8937(2)	62(1)
F(122)	12186(3)	6851(2)	9128(2)	45(1)
F(123)	11381(4)	6409(2)	8217(2)	50(1)
C(14A)	7358(6)	4430(3)	9281(3)	43(1)
F(141)	8451(4)	3945(2)	9535(2)	72(1)
F(142)	6913(4)	4139(2)	8718(2)	62(1)
F(143)	6172(4)	4349(2)	9602(2)	70(1)
C(16A)	6442(5)	7284(3)	9470(2)	36(1)
F(161)	7151(3)	7738(2)	9925(1)	41(1)
F(162)	5167(3)	7023(2)	9683(2)	48(1)
F(163)	6001(3)	7754(2)	8984(1)	43(1)
C(21)	6728(5)	8632(3)	7693(2)	26(1)
C(22)	6378(5)	8004(3)	7276(2)	27(1)
C(23)	5067(5)	8010(3)	6854(2)	29(1)
C(24)	4034(5)	8619(3)	6847(2)	28(1)
C(25)	4343(6)	9241(3)	7248(2)	30(1)
C(26)	5675(5)	9259(3)	7662(2)	24(1)
C(22A)	7424(6)	7313(3)	7277(2)	34(1)
F(221)	8793(3)	7500(2)	7110(2)	50(1)
F(222)	7663(4)	6966(2)	7830(1)	51(1)
F(223)	6843(3)	6760(2)	6874(1)	41(1)
C(24A)	2621(6)	8606(3)	6388(2)	37(1)
F(241)	2931(4)	8791(2)	5813(1)	54(1)
F(242)	1961(4)	7904(2)	6345(2)	53(1)
F(243)	1569(3)	9117(2)	6528(1)	49(1)
C(26A)	5933(5)	9956(3)	8076(2)	34(1)
F(261)	5044(4)	10560(2)	7874(2)	59(1)
F(262)	5584(4)	9801(2)	8654(2)	59(1)
F(263)	7348(3)	10213(2)	8144(2)	48(1)
C(32)	11652(6)	8323(4)	10011(2)	48(2)
C(33)	12841(8)	8353(5)	10487(3)	63(2)
C(34)	14185(7)	8700(4)	10378(3)	54(2)
C(35)	14323(6)	8975(3)	9796(3)	41(1)
C(36)	13103(6)	8905(3)	9339(3)	34(1)
N(4)	10464(4)	9770(2)	8431(2)	29(1)
C(42)	10990(5)	10322(3)	8844(2)	31(1)
C(43)	11610(5)	11013(3)	8672(3)	37(1)
C(44)	11682(6)	11165(3)	8055(3)	39(1)
C(45)	11075(6)	10619(3)	7618(3)	36(1)
C(46)	10483(6)	9938(3)	7821(2)	33(1)

X. Bis(μ^2-methoxo)-tetra-bis(2,6-di-isopropylphenyl)imido-dimethyl-di-molybdenum

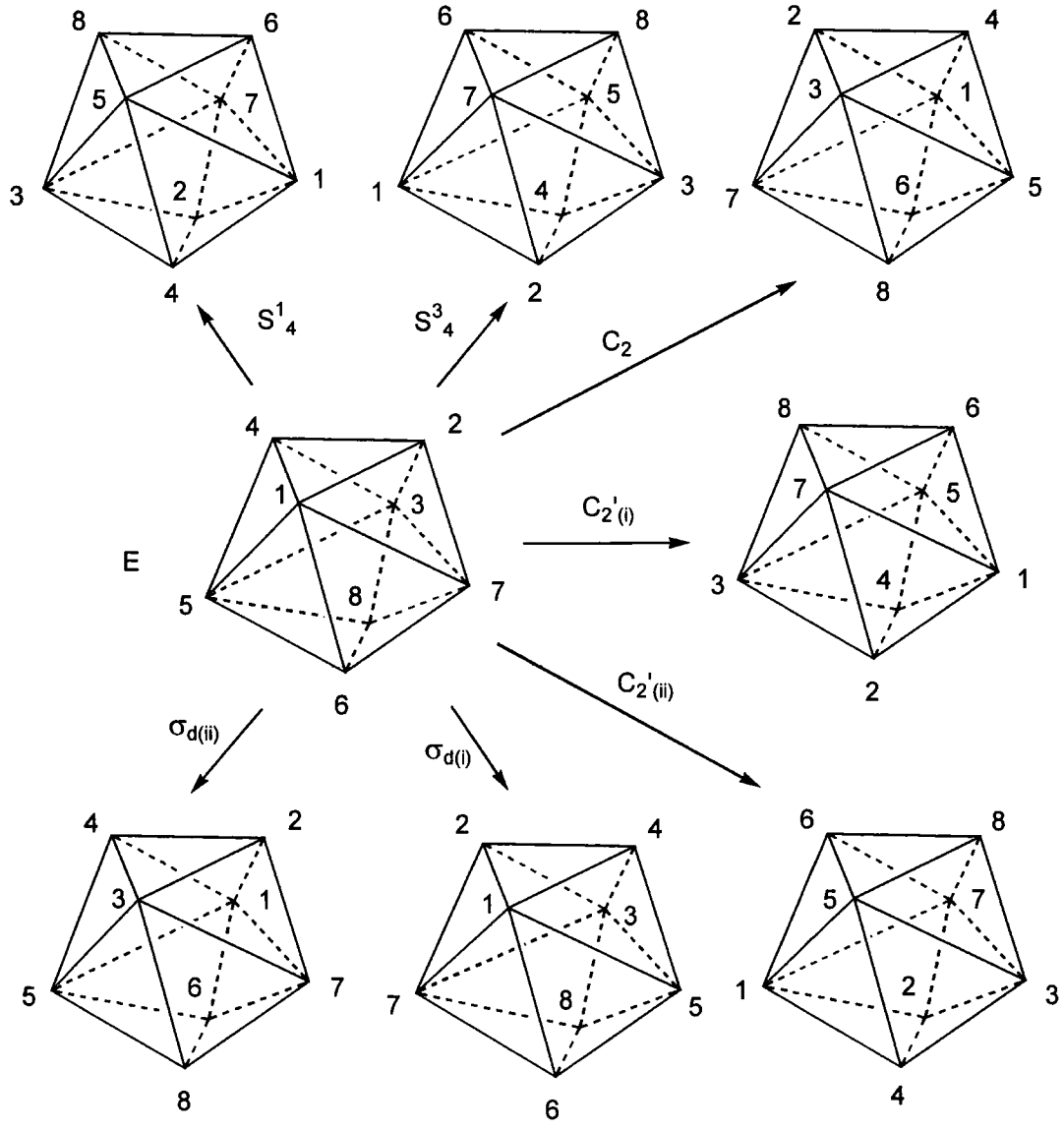
Atomic coordinates *x*, *y*, *z* ($\times 10^4$) and equivalent isotropic displacement U_{eq} ($\times 10^3$)

Atom	*x*	*y*	*z*	U_{eq}
Mo(1)	1278(1)	474(1)	346(1)	19(1)
C(3)	2018(4)	1641(6)	-364(3)	29(1)
O(1)	-187(2)	1039(3)	-257(2)	24(1)
N(1)	2413(3)	-512(4)	515(2)	23(1)
N(2)	1254(3)	1570(4)	1007(2)	21(1)
C(11)	3350(3)	-1250(5)	650(2)	21(1)
C(12)	3792(4)	-1854(5)	129(2)	26(1)
C(13)	4739(4)	-2547(6)	276(3)	34(1)
C(14)	5237(4)	-2663(5)	920(3)	30(1)
C(15)	4780(4)	-2096(5)	1421(3)	27(1)
C(16)	3845(4)	-1370(5)	1306(2)	23(1)
C(121)	3241(4)	-1785(6)	-575(2)	34(1)
C(12A)	3950(6)	-1120(8)	-1024(3)	51(2)
C(12B)	2915(6)	-3145(8)	-830(3)	52(2)
C(161)	3348(4)	-773(6)	1868(2)	30(1)
C(16A)	2849(8)	-1830(9)	2240(4)	94(3)
C(16B)	4115(6)	63(9)	2314(4)	90(3)
C(21)	1638(4)	2076(5)	1628(2)	21(1)
C(22)	2377(4)	3118(5)	1668(2)	24(1)
C(23)	2764(4)	3605(5)	2284(2)	27(1)
C(24)	2432(4)	3092(6)	2843(3)	33(1)
C(25)	1687(4)	2085(5)	2798(2)	27(1)
C(26)	1258(4)	1568(5)	2192(2)	21(1)
C(221)	2695(4)	3710(5)	1045(3)	33(1)
C(22A)	3774(6)	4416(8)	1141(4)	55(2)
C(22B)	1817(7)	4615(8)	734(4)	59(2)
C(261)	384(4)	536(5)	2128(2)	27(1)
C(26A)	239(6)	-114(7)	2786(3)	41(2)
C(26B)	-684(5)	1126(7)	1815(3)	37(1)
C(01)	-450(6)	2340(6)	-485(4)	49(2)

Appendix II.

Symmetry Isomers and Coordinates for 8-Coordination Sphere in Point Groups D_{2d} (DOD) and D_{4d} (SQAP)

Symmetry isomers in D_{2d}:



Symmetry coordinates in D_{2d}:

Derived from 28 valence angles, L-M-L, where the reducible representation is:

$$\Gamma_\theta = 7A_1 + A_2 + 3B_1 + 5B_2 + 6E,$$

only 13 of the following coordinates are independent.

A_1: $$S_1 = \frac{1}{\sqrt{8}}(\theta_{12} + \theta_{14} + \theta_{23} + \theta_{34} + \theta_{56} + \theta_{58} + \theta_{67} + \theta_{78})$$

$$S_2 = \frac{1}{2}(\theta_{15} + \theta_{17} + \theta_{35} + \theta_{37}); \qquad S_3 = \frac{1}{2}(\theta_{16} + \theta_{27} + \theta_{38} + \theta_{45});$$

$$S_4 = \frac{1}{2}(\theta_{18} + \theta_{25} + \theta_{36} + \theta_{47}); \qquad S_5 = \frac{1}{2}(\theta_{26} + \theta_{28} + \theta_{46} + \theta_{48});$$

$$S_6 = \frac{1}{\sqrt{2}}(\theta_{13} + \theta_{57}); \qquad S_7 = \frac{1}{\sqrt{2}}(\theta_{24} + \theta_{68}).$$

A_2: $$S_8 = \frac{1}{\sqrt{8}}(\theta_{12} - \theta_{14} - \theta_{23} + \theta_{34} + \theta_{56} - \theta_{58} - \theta_{67} + \theta_{78});$$

B_1: $$S_9 = \frac{1}{\sqrt{8}}(\theta_{12} - \theta_{14} - \theta_{23} + \theta_{34} - \theta_{56} + \theta_{58} + \theta_{67} - \theta_{78});$$

$$S_{10} = \frac{1}{2}(\theta_{15} - \theta_{17} - \theta_{35} + \theta_{37}); \qquad S_{11} = \frac{1}{2}(\theta_{26} - \theta_{28} - \theta_{46} + \theta_{48});$$

B_2: $$S_{12} = \frac{1}{\sqrt{8}}(\theta_{12} + \theta_{14} + \theta_{23} + \theta_{34} - \theta_{56} - \theta_{58} - \theta_{67} - \theta_{78});$$

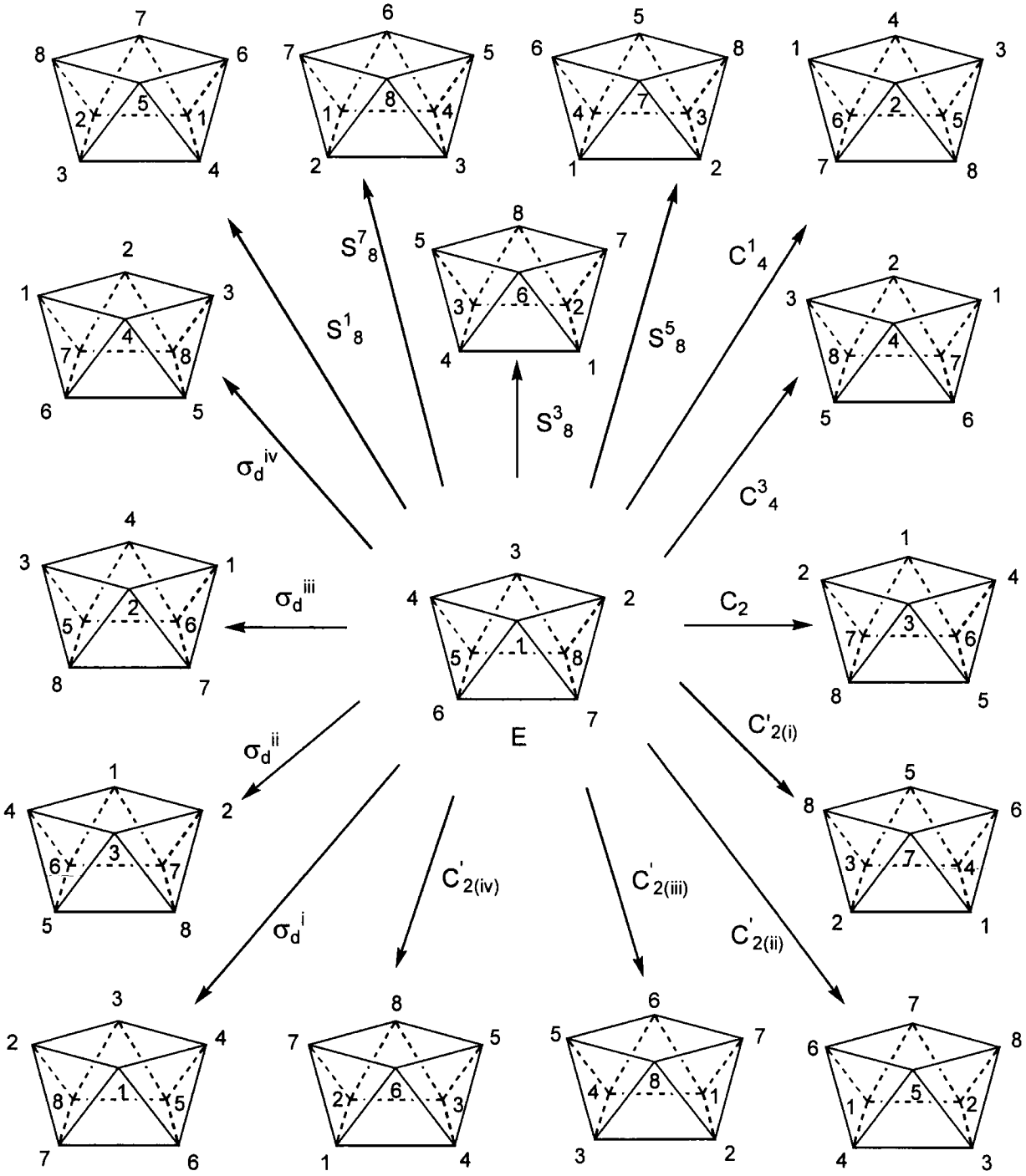
$$S_{13} = \frac{1}{2}(\theta_{16} - \theta_{27} + \theta_{38} - \theta_{45}); \qquad S_{14} = \frac{1}{2}(\theta_{18} - \theta_{25} + \theta_{36} - \theta_{47});$$

$$S_{15} = \frac{1}{\sqrt{2}}(\theta_{13} - \theta_{57}); \qquad S_{16} = \frac{1}{\sqrt{2}}(\theta_{24} - \theta_{68});$$

E:
$$\begin{cases} S_{17a} = \dfrac{1}{\sqrt{2}}(\theta_{12} - \theta_{34}) \\ S_{17b} = \dfrac{1}{\sqrt{2}}(\theta_{14} - \theta_{23}) \end{cases}; \quad \begin{cases} S_{18a} = \dfrac{1}{\sqrt{2}}(\theta_{15} - \theta_{37}) \\ S_{18b} = \dfrac{1}{\sqrt{2}}(\theta_{17} - \theta_{35}) \end{cases}; \quad \begin{cases} S_{19a} = \dfrac{1}{\sqrt{2}}(\theta_{16} - \theta_{38}) \\ S_{19b} = \dfrac{1}{\sqrt{2}}(\theta_{27} - \theta_{45}) \end{cases};$$

$$\begin{cases} S_{20a} = \dfrac{1}{\sqrt{2}}(\theta_{18} - \theta_{36}) \\ S_{20b} = \dfrac{1}{\sqrt{2}}(\theta_{25} - \theta_{47}) \end{cases}; \quad \begin{cases} S_{21a} = \dfrac{1}{\sqrt{2}}(\theta_{26} - \theta_{48}) \\ S_{21b} = \dfrac{1}{\sqrt{2}}(\theta_{28} - \theta_{46}) \end{cases}; \quad \begin{cases} S_{22a} = \dfrac{1}{\sqrt{2}}(\theta_{56} - \theta_{78}) \\ S_{22b} = \dfrac{1}{\sqrt{2}}(\theta_{58} - \theta_{67}) \end{cases}.$$

Symmetry isomers in D_{4d}:



Symmetry Coordinates in D_{4d}: Derived from 28 valence angles, L-M-L, where the reducible representation is:

$$\Gamma_\theta = 4A_1 + 2B_1 + 2B_2 + 3E_1 + 3E_2 + 3E_3,$$

only 13 of the following symmetry coordinates are independent)

A_1: $$S_1 = \frac{1}{\sqrt{8}}(\theta_{12} + \theta_{14} + \theta_{23} + \theta_{34} + \theta_{56} + \theta_{58} + \theta_{67} + \theta_{78});$$

$$S_2 = \frac{1}{\sqrt{8}}(\theta_{15} + \theta_{18} + \theta_{25} + \theta_{26} + \theta_{36} + \theta_{37} + \theta_{47} + \theta_{48});$$

$$S_3 = \frac{1}{\sqrt{8}}(\theta_{16} + \theta_{17} + \theta_{27} + \theta_{28} + \theta_{35} + \theta_{38} + \theta_{45} + \theta_{46});$$

$$S_4 = \frac{1}{2}(\theta_{13} + \theta_{24} + \theta_{23} + \theta_{57} + \theta_{68})$$

B_1: $$S_5 = \frac{1}{\sqrt{8}}(\theta_{15} - \theta_{18} - \theta_{25} + \theta_{26} - \theta_{36} + \theta_{37} - \theta_{47} + \theta_{48});$$

$$S_6 = \frac{1}{\sqrt{8}}(\theta_{16} - \theta_{17} + \theta_{27} - \theta_{28} - \theta_{35} + \theta_{38} + \theta_{45} - \theta_{46});$$

B_2: $$S_7 = \frac{1}{\sqrt{8}}(\theta_{12} + \theta_{14} + \theta_{23} + \theta_{34} - \theta_{56} - \theta_{58} - \theta_{67} - \theta_{78});$$

$$S_8 = \frac{1}{2}(\theta_{13} + \theta_{24} - \theta_{57} - \theta_{68});$$

E_1: $$\begin{cases} S_{9a} = \frac{1}{2}[\theta_{12} - \theta_{34} + \frac{\sqrt{2}}{2}(\theta_{56} + \theta_{58} - \theta_{67} - \theta_{78})] \\ S_{9b} = \frac{1}{2}[\theta_{14} - \theta_{23} - \frac{\sqrt{2}}{2}(\theta_{56} - \theta_{58} + \theta_{67} - \theta_{78})] \end{cases};$$

$$\begin{cases} S_{10a} = \frac{1}{2}[\theta_{15} - \theta_{37} + \frac{\sqrt{2}}{2}(\theta_{18} + \theta_{25} - \theta_{36} - \theta_{47})] \\ S_{10b} = \frac{1}{2}[\theta_{26} - \theta_{48} - \frac{\sqrt{2}}{2}(\theta_{18} - \theta_{25} - \theta_{36} + \theta_{47})] \end{cases};$$

$$\begin{cases} S_{11a} = \frac{1}{2}[\theta_{16} - \theta_{38} - \frac{\sqrt{2}}{2}(\theta_{17} - \theta_{28} - \theta_{35} + \theta_{46})] \\ S_{11b} = \frac{1}{2}[\theta_{27} - \theta_{45} - \frac{\sqrt{2}}{2}(\theta_{17} + \theta_{28} - \theta_{35} - \theta_{46})] \end{cases};$$

E_2:

$$\begin{cases} S_{12a} = \dfrac{1}{2}(\theta_{12} - \theta_{14} - \theta_{23} + \theta_{34}) \\ S_{12b} = \dfrac{1}{2}(\theta_{56} - \theta_{58} - \theta_{67} + \theta_{78}) \end{cases};$$

$$\begin{cases} S_{13a} = \dfrac{1}{2}(\theta_{15} - \theta_{26} + \theta_{37} - \theta_{48}) \\ S_{13b} = \dfrac{1}{2}(\theta_{18} - \theta_{25} + \theta_{36} - \theta_{47}) \end{cases};$$

$$\begin{cases} S_{14a} = \dfrac{1}{2}(\theta_{16} - \theta_{27} + \theta_{38} - \theta_{45}) \\ S_{14b} = \dfrac{1}{2}(\theta_{17} - \theta_{28} + \theta_{35} - \theta_{46}) \end{cases};$$

$$\begin{cases} S_{15a} = \dfrac{1}{\sqrt{2}}(\theta_{13} - \theta_{24}) \\ S_{15b} = \dfrac{1}{\sqrt{2}}(\theta_{57} - \theta_{68}) \end{cases};$$

E_3:

$$\begin{cases} S_{16a} = \dfrac{1}{2}[\theta_{12} - \theta_{34} - \dfrac{\sqrt{2}}{2}(\theta_{56} + \theta_{58} - \theta_{67} - \theta_{78})] \\ S_{16b} = \dfrac{1}{2}[\theta_{14} - \theta_{23} + \dfrac{\sqrt{2}}{2}(\theta_{56} - \theta_{58} + \theta_{67} - \theta_{78})] \end{cases};$$

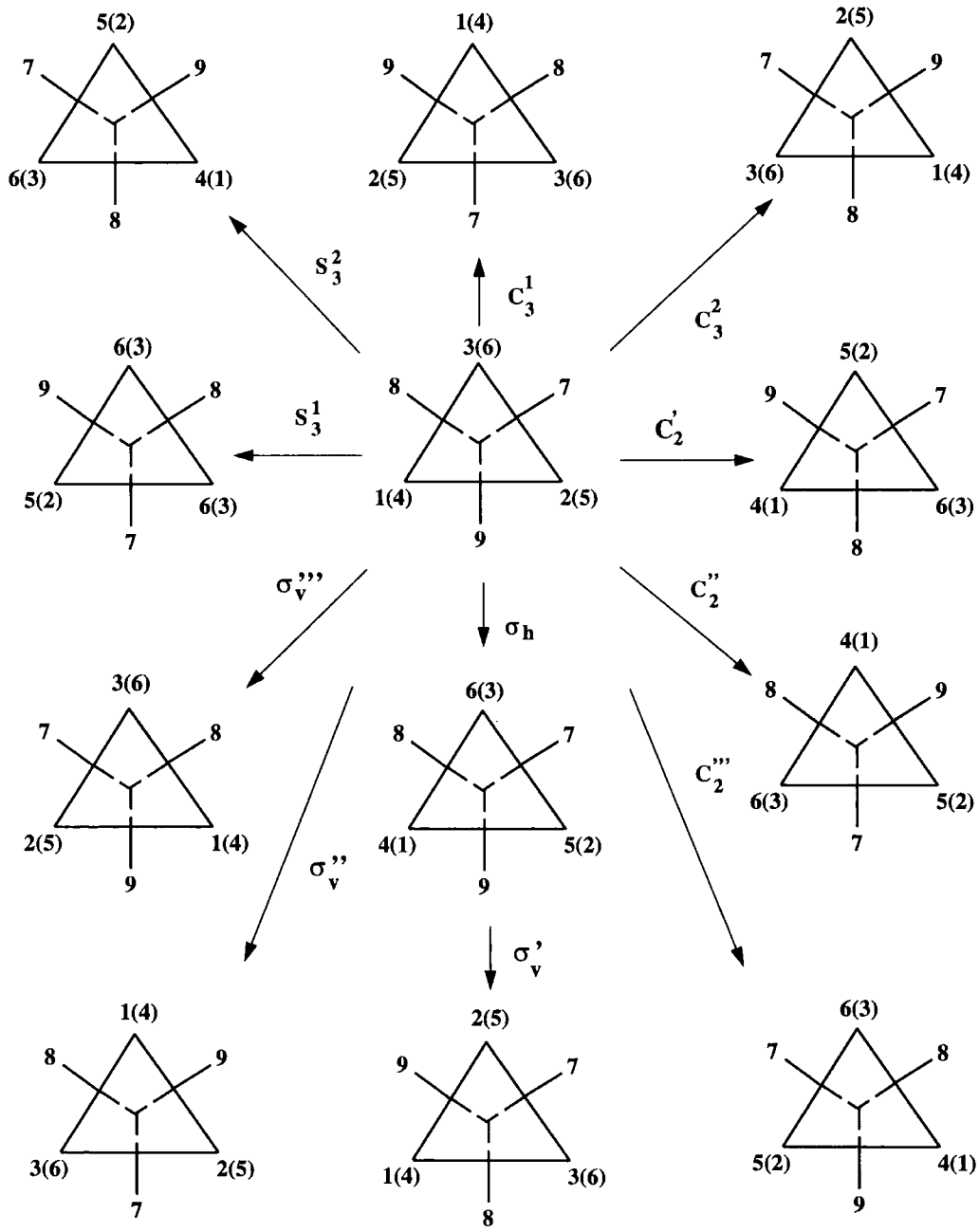
$$\begin{cases} S_{17a} = \dfrac{1}{2}[\theta_{15} - \theta_{37} - \dfrac{\sqrt{2}}{2}(\theta_{18} + \theta_{25} - \theta_{36} - \theta_{47})] \\ S_{17b} = \dfrac{1}{2}[\theta_{26} - \theta_{48} + \dfrac{\sqrt{2}}{2}(\theta_{18} - \theta_{25} - \theta_{36} + \theta_{47})] \end{cases};$$

$$\begin{cases} S_{18a} = \dfrac{1}{2}[\theta_{16} - \theta_{38} + \dfrac{\sqrt{2}}{2}(\theta_{17} - \theta_{28} - \theta_{35} + \theta_{46})] \\ S_{18b} = \dfrac{1}{2}[\theta_{27} - \theta_{45} - \dfrac{\sqrt{2}}{2}(\theta_{17} + \theta_{28} - \theta_{35} - \theta_{46})] \end{cases}.$$

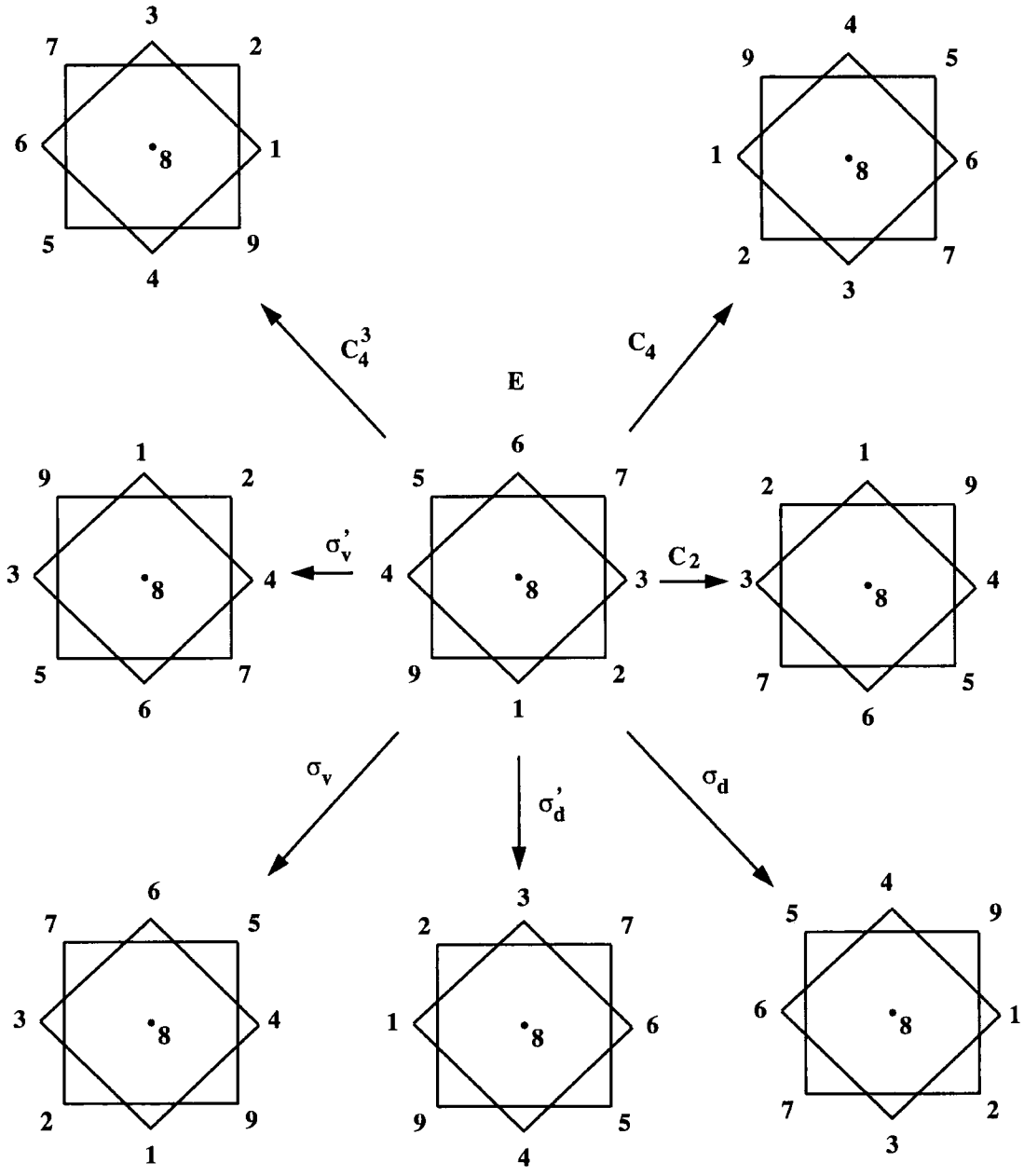
Appendix III

Symmetry Isomers for 9-Coordination Sphere in Point Groups D_{3h} (TTP) and C_{4v} (CSA).

12 isomers in D_{3h} symmetry



8 isomers in C_{4v}



Appendix IV

Parameters Used to Link to NAG Fortran Library for PCA (G03AAF) and FA (G03CAF, G03CCF) Calculations*

G03AAF performs a principal component analysis (PCA) on data matrix; both the principal component loadings and the principal component scores are returned. As G03AAF uses a singular value decomposition of the data matrix, it will be less affected by ill-conditioned problems than traditional methods using the eigenvalue decomposition of the variance-covariance matrix.

Specification

```
      SUBROUTINE G03AAF(MATRIX, STD, WEIGHT, N, M, X, LDX, ISX, S,
     1                  WT, NVAR, E LDE, P, LDP, V, LDV, WK, IFAL)
      INTEGER          N, M, LDX, ISX(M), NVAR, LDE, LDP, LDV, IFAL
      REAL             X(LDX,M), S(M), WT(*), E(LDE,6), P(LDP,NVAR),
     1                 V(LDV,NVAR), WK(NVAR*NVAR+5*(NVAR-))
      CHARACTER*1      MATRIX, STD, WEIGHT
```

Parameters

1: MATRIX –CHARACTER*1

On entry: indicates for which type of matrix the principal component analysis is to be carried out.

If MATRIX = 'C' or 'c', then it is for the correlation matrix.

If MATRIX = 'S' or 's', then it is for a standardized matrix, with standardization given by S.

If MATRIX = 'U' or 'u', then it is for the sums of squares and cross-products matrix.

If MATRIX = 'V' or 'v', then it is for the variance-covariance matrix.

2: STD – CHARACTER*1

On entry: indicates if the principal component scores are to be standardized to have variance equal to 1.0.

If STD = 'S' or 's', then the principal component scores are standardized.

* NAG Fortran Library Introductory Guide 18, The Numerical Algorithms Ltd., Oxford, UK, (1997)

If STD = 'U' or 'u', then the principal component scores are unstandardized.

3: WEIGHT – CHARACTER*1.

On entry: indicates if weights are to be used.

If WEIGHT = 'U' or 'u' (Unweighted), then no weights are used.

If WEIGHT = 'W' or 'w'(Weighted), then weights are used and must be supplied in WT.

4. N – INTEGER.

On entry: the number of observations, n. $N \geq 2$.

5: M – INTEGER.

On entry: the number of variables in the data matrix, m. $M \geq 1$.

6: X(LDX,M) – real array.

On entry: X(*i*,*j*) must contain the *i*th observation for the *j*th variable, for $i = 1,2,\ldots,n$; $j = 1,2,\ldots,m$.

7: LDX – INTEGER.

On entry: the first dimension of the array X as declared in the (sub) program from which G03AAF is called. $LDX \geq N$.

8: ISX(M) – INTEGER array.

On entry: ISX(*j*) indicates whether or not the *j*th variable is to be included in the analysis. If ISX(*j*) > 0, then the variable contained in the *j*th column of X is included in the principal component analysis, for $j = 1,2,\ldots,m$.

9: S(M) – real array.

On entry: the standardization to be used, if any.

If MATRIX = 'S' or 's', then the first m elements of S must contain the standardization coefficients, the diagonal elements of σ.

Constraint: if ISX(*j*) > 0, then S(*j*) > 0.0, for $j = 1,2,\ldots,m$.

On exit: if MATRIX = 'S' or 's', then S is unchanged on exit.

If MATRIX = 'C' or 'c', then S contains the variances of the selected variables. S(*j*) contains the variance of the variable in the *j*th column of X if ISX(*j*) ≥ 0.

If MATRIX = 'U', 'u', 'V' or 'v'. then S is not referenced.

10: WT(*) – real array.

On entry: if WEIGHT = 'W' or 'w', then the first n elements of WT must contain the weights to be used in the principal component analysis.

If WT(*i*) = 0.0, then the *i*th observations is not included in the analysis. The effective number of observations is the sum of the weights.

If WEIGHT = 'U' or 'u', then WT is not referenced and the effective number of observations is *n*.

WT(*i*) ≥ 0.0, for *i* = 1,2,...,n and the sum of weights ≥ NVAR + 1.

11: NVAR – INTEGER.

On entry: the number of variables in the principal component analysis, *p*.

1 ≤ NVAR ≤ min(N-1,M).

12: E(LDE,6) – real array.

On exit: the statistics of the principal component analysis.

E(*i*,1), the eigenvalues associated with the *i*th principal component, λ_i^2, for *i* = 1,2,...,*l*.

E(*i*,2), the proportion of variation explained by the *i*th principal component, for *i* = 1,2,...,*l*.

E(*i*,3), the cumulative proportion of variation explained by the first *i*th principal components, for *i* =1,2,...,*l*.

E(*i*,4), the χ^2 statistics, for *i* = 1,2,...,*l*.

E(*i*,5), the degrees of freedom for the χ^2 statistics, for *i* = 1,2,...,*l*.

If MATRIX ≠ 'C' or 'c', the E(*i*,6), contains significance level for the χ^2 statistic, for *i* = 1,2,...,NVAR.

If MATRIX = 'C' or 'c', E(*i*,6) is returned as zero.

13: LDE – INTEGER.

On entry: the first dimension of the array E as declared in the (sub)program from which

G03AAF is called. LDE ≥ NVAR.

14: P(LDP,NVAR) – real array

On exit: the NVAR columns of P contain the principal component loadings, a_i. The *j*th column of P contains the NVAR coefficient for the *j*th principal component.

15: LDP – INTEGER.

On entry: the first dimension of the P as declared in the (sub)program from which G03AAF is called. LDP ≥ NVAR.

16: V(LDV,NVAR) – real array.

On exit: the first NVAR columns of V contain the principal component scores. The *j*th column of V contains the N scores for the *j*th principal component.

If WEIGHT = 'W' or 'w', then any rows for which WT(i) is zero will be set to zero.

17: LDV – INTEGER.

On entry: the first dimension of the array V as declared in the (sub)program from which G03AAF is called. LDV $\geq$ N.

18: WK(NVAR*NVAR+5*(NVAR-1) – real array.

19: IFAIL – INTEGER.

On entry: IFAIL must be set to 0, -1, or 1.

On exit: IFAIL = 0 unless the routine detects an error.

G03CAF computes the maximum likelihood estimates of the parameters of a factor analysis model. Either the data matrix or a correlation/covariance matrix may be input. Factor loadings, communalities and residual correlations are returned. The factor loadings may be orthogonally rotated by using G03BAF and factor score coefficients can be computed using G03CCF.

Specification

```
      SUBROUTINE G03CAF(MATRIX, WEIGHT, N, M, X, LDX, NVAR, ISX, NFAC,
     1                  WT, E, STAT, COM, PSI, RES, FL, IDFL, IOP,
                       IWK,
     2                  WK, LWK, IFAIL)
       INTEGER          N, M, LDX, NVAR, ISX(M), NFAC, LDFL, IOP(5),
     1                  IWK(4*NVAR+2), LWK, IFAIL
       REAL             X(LDX,M), WT(*), E(NVAR), STAT(4), COM(NVAR),
     1                  PSI(NVAR), RES(NVAR*(NVAR-1)/2), FL(LDFL,NFAC),
     2                  WK(LWK)
       CHARACTER*1      MATRIX, WEIGHT
```

Parameters

1: MATRIX – CHARACTER*1

On entry: selects the type of matrix on which factor analysis is to be performed.

If MATRIX = 'D' (Data input), then the data matrix will be input in X and factor analysis will be computer for the correlation matrix.

If MATRIX = 'S', then the data matrix will be input in X and factor analysis will be computed for the covariance matrix.

If MATRIX = 'C', then the correlation/variance-covariance matrix will be input in X and factor analysis computed for this matrix.

2: WEIGHT – CHARACTER*1

On entry: if MATRIX = 'D' or 'S', WEIGHT indicates if weights are to be used.

If WEIGHT = 'U', then no weights are used.

If WEIGHT = 'W', then weights are used and must be supplied in WT.

If WEIGHT = 'C', then WEIGHT is not referenced.

3: N – INTEGER.

On entry: if MATRIX = 'D' or 'S' the number of observations in the data array X.

If MATRIX = 'C' the (effective) number of observations used in computing the (possibly weighted) correlation/variance-covariance matrix input in X. N > NVAR.

4: M – INTEGER.

On entry: the number of variables in the data/correlation/variance-covariance matrix. M ≥ NVAR.

5: X(LDX,M) – real array.

On entry: the input matrix.

If MATRIX = 'D' or 'S', then X must contain the data matrix, i.e. X(i,j) must contain the ith observation for the jth variable, for $i = 1,2,\ldots,n$; $j = 1,2,\ldots,M$.

If MATRIX = 'C', then X must contain the correlation or variance-covariance matrix. Only the upper triangular part is required.

6: LDX – INTEGER.

On entry: the first dimension of the array X as declared in the (sub)program from which G03CAF is called. If MATRIX = 'D' or 'S', then LDX ≥ N, if MATRIX = 'C', then LDX ≥ M.

7: NVAR – INTEGER.

On entry: the number of variables in the factor analysis, p. NVAR ≥ 2.

8: ISX(M) – INTEGER array.

On entry: ISX(j) indicates whether or not the jth variable is variable in the factor analysis. If ISX(j) ≥ 1, then the variable represented by the jth column of X is included in the analysis; otherwise it is excluded, for $j = 1,2,\ldots,M$. ISX(j) > 0 for NVAR values of j.

9: NFAC – INTEGER.

On entry: the number of factors, k. $1 \leq \text{NFAC} \leq \text{NVAR}$.

10: WT(*) – real array.

On entry: if WEIGHT = 'W' and MATRIX = 'D' or 'S', WT must contain the weights to be used in the factor analysis. The effective number of observations in the analysis will then be the sum of weights. If WT(i) = 0.0, then the ith observation is not included in the analysis. If WEIGHT = 'U' or MATRIX = 'C'. WT is not referenced and effective number of observations is n.

Constraint: if WEIGHT = 'W', then WT(i) $\geq$ 0.0, for i = 1,2,...,n, and the sum of weights > NVAR.

11: E((NVAR) – real array.

On exit: the eigenvalues θ_i, for i = 1,2,...,p.

12: STAT – real array.

On exit: the test statistics.

STAT(1) contains the value F(Ψ).

STAT(2) contains the test statistic, χ^2.

STAT(3) contains the degrees of freedom associated with the test statistic.

STAT(4) contains the significance level.

13: COM(NVAR) – real array.

On exit: the communalities.

14: PSI(NVAR) – real array.

On exit: the estimates of Ψ_i, for i = 1,2,...,p.

15: RES(NVAR*(NVAR-1)/2) – real array.

On exit: the residual correlations. The residual correlation for the ith and jth variables is stored in RES((j-1)(j-2)/2+i), $i < j$.

16: FL(LDFL,NFAC) – real array.

On exit: the factor loadings. FL(i,j) contains λij, for i = 1,2,...,p; j = 1,2,...,k.

17: LDFL – INTEGER.

On entry: the first dimension of the array FL as declared in the (sub)program from which G03CAF is called. LDFL $\geq$ NVAR.

18: IOP(5) – INTEGER array.

On entry: options for the optimization. There are four options to be set:

iprint – controls iteration monitoring;

if *iprint* ≤ 0, then there is no printing of information else if *iprint* > 0, then information is printed at every *iprint* iterations. The information printed consists of the value of $F(\Psi)$ at that iteration, the number of evaluations of $F(\Psi)$, the current estimates of the communalities and an indication of whether or not they are at the boundary.

19: IWK(4*NVAR+2) – INTEGER array.

20: WK(LWK) – real array.

21: LWK – INTEGER.

On entry: the length of the workspace.

Constraints: if MATRIX = 'D' or 'S', then

$$LWK \geq \max((5\times NVAR\times NVAR+33\times NVAR-4)/2,$$
$$N\times NVAR + 7\times NVAR + NVAR\times(NVAR-1)/2).$$

If MATRIX = 'C', then

$$LWK \geq (5\times NVAR\times NVAR+33\times NVAR-4)/2.$$

22: IFAIL – INTEGER.

On entry: IFAIL must be set to 0, -1 or 1.

On exit: IFAIL = 0 unless the routine detects an error or gives a warning.

Appendix V.

Publications, Attended Conferences, Seminars and Courses

Publications:

(1) Howard, J. K. A., Copley, R. C. B., **Yao, J. W.** and Allen, F. H., "Systematic analysis of metal coordination sphere geometry from crystallographic data: a general method for detection geometrical preferences, deformations and interconversion pathways", *Chem. Commn.*, 2175, (1998).

(2). Bryce, M. R., Chalton, M. A., **Yao, J.W.** and Howard, J. A. K., "Synthesis and nitrosation reactions of π-extended 1,3-dithiol-2-ylidene system", *Tetrahedron*, 54, 3919-3928, (1998).

(3). **Yao, J. W.,** Copley, R.C.B, Allen, F.H. and Howard, J.A.K. "Automated geometrical analysis of metal 9-coordination sphere from crystallographic data", *27th international school of crystallography*, P64, May, 1998, Erice, Italy.

(4). **Yao, J. W.,** Copley, R.C.B, Allen, F.H. and Howard, J.A.K. "An identification and classification of geometry of eight-coordination transition metal complexes using crystallographic data", *BCA Annual Meeting CP15*, April, 1997, University of Leeds, UK.

(5). Bricklebank, N., Howard, J.A.K., Rawson, J.M. and **Yao, J.W.**, "The first example of side-on bonding in a tertiary phosphine-dihaloga donor-acceptor complex", submitted to *Angew. Chem. Int. Ed. Engl.* (1998).

(6). Dillon, K.B., Gibson, V.C., Howard, J.A.K., Redshaw, C., Sequeira, L. and **Yao, J. W.**, "Group 6 transition metal complexes containing the σ-bonded 2,4,6-tris(trifluoromethyl) phenyl("fluoromes") ligand", *J. Organomet. Chem.* **528**, 179-183, (1997).

(7). **Yao, J. W.**, Howard, J. A. K., Allen, F. H. and Pitchford, N. A., "A database study on geometry of transition metal alkyne and alkene complexes", *BCA Annual Meeting CP20,* April, 1996, Univ. of Cambridge, UK.

(8). Dillon, K. B., Gibson, V. C., Howard, J. A. K., Sequeira, L and **Yao, J.W.,** "Bis(imido) molybdenum(IV) complexes containing η^2-diphosphene ligands", *Polyhedron*, **15**(23), 4173-4177, (1996).

(9) Harris, R. H., Howard, J. A. K., Samadi-Mayodi, A. and **Yao, J. W.**, " A new silicate clathrate hydrate: An X-ray diffraction and nuclear magnetic resonance study of a system with octameric silicate anions and trivalent cations", *J. of Solid State Chem.*, **120**, 231-237, (1995).

The author has attended Durham X-ray crystallography group seminars weekly to report the research results and discuss with other members during 1995-1998.

Attended research conferences and presented results during the period of tuition of this thesis.

(1) BCA 5th Intensive Course in X-ray Structure Analysis, 2-8 April, 1995, Aston Birmingham. (before the tuition term).
(2) UK Charge Density Meeting, 15 - 17, December, 1995, Durham.
(3) BCA Chemical Crystallography Group Autumn Meeting 1995, *Structural Refinement,* 17 November, 1995, Manchester.
(4) BCA Annual Meeting, 1-4, April, 1996. University of Cambridge, Cambridge.
Poster Presented: A Database Study of the Geometry of Transition Metals Alkyne and Alkene Complexes.
(5) BCA Chemical Crystallography Group Autumn Meeting 1996, *Dynamic Crystallography*, 14 November, CCLRC Daresbury Laboratory.
(6) BCA Annual Meeting, 14-17, April, 1997, University of Leeds, Leeds.
Poster Presented: An Identification and Classification of the Geometry of Eight-Coordination Transition Metal Complexes Using Crystallographic Data.
(7) Databases for Macromolecular Crystallographers, 9-10, January, 1998, University of Reading, Reading.

(8) 27th International School of Crystallography, *Implications of Molecular and Materials Structure for New Technologies*, 28 May – 7 June, 1998, Erice, Sicily, Italy.

Poster Presented: Systematic Analysis of Nine-Coordination Sphere Geometry Using Crystallographic Data.

(9) ACA Annual Meeting, 18-23 July, Arlington, Virginia, USA.

Lecture Presented: Systematic Analysis of Metal 7-Coordination Sphere Geometry from Crystallographic Data.

Attended Seminars and lectures from invited speakers to University of Durham, Department of Chemistry.

1995/1996

(1) Prof. Dennis Tuck, *University of Windsor, Ontario, Canada,* "New Indium Coordination Chemistry", 4 October, 1995.

(2) Prof. P. Lugar, *Frei Univ. Berlin, FRG*, "Low temperature Crystallography", 11, October, 1995.

(3) Prof. R. Schmultzer, *Univ. Braunschweig, FRG*, "Calixarene-Phosphorus Chemistry: A new Dimension in Phosphorus Chemistry", 13, October, 1995.

(4) Dr. Andrea Sella, *UCL, London*, "Chemistry of Lanthanides with Polypyrazoylborate Ligands", 15 November, 1995.

(5) Dr. Bill Henderson, *Waikato Univ., NZ,* "Electrospray Mass Spectrometry - a new sporting technique", 10 January, 1996.

(6) Prof. J.W. Emsley, *Southampton Univ.*, "Liquid Crystal: More than Meets the Eye", 17, January, 1996.

(7) Dr. Alan Armstrong, *Nottingham Univ.*, "Alkene Oxidation and Natural Product System", 24 January, 1996.

(8) Prof. E.W. Randall, *Qeueen Mary & Westfield College*, "New Perspectives in NMR Imaging", 28 February, 1996.

(9) Prof. V. Balzani, *(RSC Endowed Lecture), Univ. of Bologna*, "Supramolecular Photochemistry", 12 March, 1996.

1996/1997

(1) Prof. G. Bowmaker, *Univ. Aukland, NZ,* "Coordination and Materials Chemistry of the Group 11 and Group 12 Metals: Some Recent Vibrational and Solid State NMR Studies", 1 October, 1996.

(2) Dr. P. Muntford, *Nottingham Univ.*, " Recent Developments in Group IV Imido Chemistry", 30 October, 1996.

(3) Dr. G. Resnati, *Milan*, Perfluorinated Oxaziridines: Mild Yet Powerful Oxidising Agents", 13 November, 1996.

(4) Prof. J. Earnshaw, *Dept. of Physics, Belfast*, "Surface Light Scattering: Ripples and Relaxation", 20 November, 1996.

(5) Dr. R. Templer, *Imperial College, London*, " Molecular Tubes and Sponges", 27 November, 1996.

(6) Prof. K. Muller-Dethlefs, *York Univ.*, " Chemical Applications of Very high Resolution ZEKE Photoelectron Spectroscopy", 1 December, 1996.

(7) Dr. C. Richards, *Cardiff Univ.*, "Sterochemical Games with Metallocences", 11 December, 1996.

(8) Dr. Geert-Jan Boons, *Univ. of Birmingham*, "New Developments in Carbohydrate Chemistry", 12 February, 1997.

1997/1998

(1) Prof. W. Atkins, *Dept. of Physics, Univ. of Bristol*, "Advances in the Control of Architecture for Polyamides: from Nylon to Genetically Engineered Skills to Monodisperse Oligoamides", 8 October, 1997.

(2) Prof. R. J. Puddephatt, *Univ. of Western Ontario, Canada,* "Organoplatinum Chemistry and catalysis", 22 October, 1997.

(3) Dr. L. Spiccia, *Monash Univ., Melbourne, Australia*, "Polynuclear Metal Complexes", 20 November, 1997.

(4) Prof. A. P. Davis, *Dept. of Chemistry, Trinity College Dublin*, " Steroid-Based Frameworks for Supramolecular Chemistry", 3 December, 1997.

(5) Sir G. Higginson, Former Professor of Engineering in Durham and Retired Vice-Chancellor of Southampton Univ., "1981 and all that", 10 December, 1997.

(6) Prof. David Andrews, Univ. of East Anglia, "Energy Transfer and Optical Harmonics in Molecular Systems", 14 January, 1998.

(7) Prof. P. Fowler, Dept. of Chemistry, Strathclyde Univ., "Classical and Non-Classical Fullerences", 4 February, 1998.

(8) Dr. C. Jones, Swansea Univ., "Low Coordination Arsenic and Antimony Chemistry", 25 February, 1998.

Completed Courses

(1) Practical Nuclear Magnetic Resonance (PG3).

(2) Synthesis Methodology in Organometallic and Coordination chemistry (PG5).

(3) Advanced Mass Spectrometry (PG6).