Coordination Chemistry Reviews
212 (2001) 133–163

COORDINATION
CHEMISTRY
REVIEWS

www.elsevier.com/locate/ccr

Molecular mechanics, data and conformational analysis of first-row transition metal complexes in the Cambridge Structural Database

Marc Zimmer *

Department of Chemistry, Connecticut College, 270 Mohegan Ave., New London, CT 06320, USA

Received 24 August 1999; accepted 18 January 2000

Contents

Keywords: Molecular mechanics; Cambridge Structural Database; Metal complexes

---

* Tel.: +1-860-4392476; fax: +1-860-4392477.
E-mail address: mzim@conncoll.edu (M. Zimmer).

PII: S0010-8545(00)00297-6

1. Introduction

1.1. Scope of the review

In this paper we will present an overview of how statistical methods, particularly cluster analysis, can be used in conjunction with inorganic molecular mechanics to structurally analyze first-row transition metal complexes that have a common backbone, these are also known as congeneric families. Using a crystal structure as the starting conformation for a molecular mechanical analysis, or comparing calculated structures with individual Cambridge Structure Database (CSD) structures is of limited utility. Significantly more information can be gathered by comparing all the conformations of a class of compounds in the CSD with all the low energy conformations available to the compounds in a crystal averaged environment. These conformations can be generated manually or using inorganic conformational search methods.

Numerous papers using correlation methods to analyze first-row transition metal complexes in the CSD have been published. However most of them do not use MM calculations to confirm or explain observed trends. Since all the reviews in this volume of Coordination Chemistry Reviews are devoted to inorganic molecular mechanics, Section 2 of this review only presents a single example of a CSD data analysis that was conducted without further analyzing the results by molecular mechanics methods. Sections 3 and 5 discuss how CSD correlation analyses have been combined with inorganic molecular mechanics and are much more comprehensive.

1.2. Inorganic molecular mechanics

Most commercially available MM programs are capable of routinely analyzing commonly occurring organic functionalities. Despite the fact that inorganic complexes were amongst the first molecules to be modeled by molecular mechanics [1], their analysis has been less common [2,3] than organic and protein calculations. Inorganic, especially transition metal, molecular mechanics are not that straightforward. The electronic effects of the partially filled d-orbitals require special care. The calculations need to take into account the multiple geometries available to the metal (square planar, tetrahedral, octahedral, etc.), parameters for the correct charge and spin state need to be applied and, if present, the *trans* and Jahn–Teller effects also need to be considered.

Just like organic molecular mechanics, inorganic molecular mechanics uses mathematical equations to mimic the strain energy present in the compound being analyzed. In order to use the equations to calculate the total strain in the molecule, one needs to know the force constant (K) for all the bonds and bond angles in the molecule, all the ideal bond lengths ($r°$) and bond angles ($\theta°$), the periodicity of the dihedral angles (n) and the barriers to their rotation (V). In addition, the van der Waals parameters (A_{ij}, B_{ij}) between the ith and jth atoms are required to simulate the nonbonded interactions, and the point charges (q_i and q_j) and the effective dielectric constant (ε) are needed to model the electrostatic potential.

In inorganic molecular mechanics these parameters are empirically derived; they have little or no relationship to physical observables such as the stretching force constants determined by infra-red spectroscopy. Generally, they have been derived by fitting a number of crystal structures and are found in the parameter set of the MM program being used. Since the parameters are normally derived by fitting a series of crystal structures, the calculated structure can be viewed as a complex in an averaged crystal lattice, not as solvated or in the gas phase [3b,e].This makes inorganic molecular mechanics an ideal method to analyze trends amongst the structures in the CSD.

In all cases the parameters have been derived for the specific force field employed, and are transferable from one molecule to another within the limits of the parameter set, but *not* between force fields. For example, it is not possible to take parameters derived for MM2 and simply use them in another force field.

Since the manner in which the geometry of the metal is calculated is of importance in conformational searching, the different methods used will be described in this section. There are three common ways of modeling valence angle deformations around transition metals within the molecular mechanics method [3]:

- In the electrostatic or ionic model, interactions between the metal and the ligands are taken to be purely ionic and there are no additional bond stretching or valence angle terms between them. This approach has been very successful with the more ionic complexes found with alkali and alkaline–earth metals [4,5] but has been less successful with transition metal compounds, where electronic effects are of importance. An advantage of the electrostatic model is that one set of parameters can model all coordination geometries. Its disadvantages include the necessity to obtain the ‘correct’ charges for the metal ion and its coordinating ligands, along with having to use either dummy atoms or specific electronic terms in the modeling of square planar compounds.
- In the valence force field model, the same potential functions used to model the organic ligand are used for metal–ligand interactions. Thus, metal–ligand stretching and ligand–metal–ligand bending parameters are required. Torsions around the metal are usually set to zero. This is the most common method of modeling small coordination compounds. It requires many parameters, including a separate set of parameters for each different geometry, and interconversion between geometries is impossible. The unique labeling problem must also be dealt with in this method [6].
- The points on a sphere (POS) approach is a mixed model in which a metal–ligand stretching function is used but the ligand–metal–ligand bond angle is replaced by 1,3-nonbonded interactions for the coordinating atoms [7–9]. This approach has the advantage of being able to model all coordination geometries with one set of parameters. Its only problem is that it is a molecular mechanical method and does not incorporate electronic effects, such as those of a ligand field. For this reason, the method always minimizes four coordinate complexes to a tetrahedral, not to a square planar geometry. This and other similar problems can be overcome by adding dummy atoms in the axial positions [10], by applying out-of-plane or plane twist functions [11–13], or by using a

molecular mechanics force field that has been enhanced by adding ligand field effects [14–16].

In order for a conformational search to find all geometric isomers of a transition metal complex, the valence angles around the metal have to be modeled with either the POS or ionic methods. This is because the valence force field model requires a separate set of parameters for each different geometry and set of geometric isomers; therefore, interconversion between geometric isomers is impossible with a valence force field description.

More detailed descriptions of inorganic molecular mechanics are available in the literature [2,3,6,17–19].

1.3. Conformational searching in molecular mechanics

Most modern MM programs have a graphics interface which allows the user to enter the structure by drawing it or by reading its Cartesian coordinates; the strain energy of the molecule in this given conformation can then be minimized. However, the resulting conformation is not necessarily the lowest energy (or global energy minimum) structure. In fact, it is quite possible that the minimized structure can be significantly higher in energy than the global minimum and therefore it may be of little practical importance.

The aim of a conformational search is to find as many minima as possible, including the global minimum, and to compute the Boltzmann population. In doing such a search a large number of high energy starting conformations are generated, minimized, compared with previously found conformers, and stored if they are unique. For a thorough search, the crude starting geometries must span the entire potential energy surface (PES); if only part of the surface is covered one cannot be sure that all important low energy minima will be found. Ideally a grid or deterministic conformational search, in which the starting conformations cover all of the conformational space, should be conducted in all molecular simulations. However, for large and/or very flexible molecules, the cpu time required to explore the entire energy surface is prohibitive. It is more common to use stochastic or Monte Carlo methods, which employ a random element to generate starting geometries or to sample structures in a molecular dynamics run. Numerous conformational searching methods have been presented in the literature. Some of them have been designed for use on small cyclic and acyclic peptides and organic molecules [20–22], while others were devised to search the conformational space available to small proteins [23], polymers [24], and carbohydrates [25]. Comparisons of different search methods have been published [26–28].

Conformational searches of transition metal compounds are complicated by the fact that inorganic complexes are often composed of multidentate ligands, resulting in many ring systems joined at the metal ion. Furthermore, metal ions can be found in a variety of coordination geometries (e.g. octahedral, trigonal prismatic, trigonal bipyramidal, square pyramidal, square planar, tetrahedral etc.), and transition metal compounds can often adopt geometrical isomers that are not available to organic compounds [29].

To date most inorganic MM calculations published have either ignored all isomers except for the one of interest, or have entered all the possible isomers graphically and minimized them individually [3,17]. The random kick and the Monte Carlo dihedral and torsional variation methods have recently been used to sample all the conformations and configurations of inorganic complexes [30]. In both cases the geometry around the metal ion was modeled using a POS approach.

1.3.1. The Monte Carlo dihedral and positional variation method

The Monte Carlo (MC) dihedral and positional variation method [31,32] uses internal coordinates, while the random kick method uses external coordinates. The advantage of using internal coordinates is that they cut down the molecular degrees of freedom. It is well known that valence angles and distances do not change much between the different conformations of a molecule, and that the main changes are in the torsion angles. For this reason it is the torsion angles that are varied to obtain the crude starting geometries in this MC search method. In each MC step a random number of flexible torsion angles are varied by a random amount, generating a new starting geometry that can be minimized. For cyclic systems, the ring needs to be cleaved at one point so that all other torsions can be varied. To prevent the two ends of the cleaved ring from being too close or too far from each other, a ring closure constraint is set. The situation is more complicated for inorganic systems, where coordination of multidendate ligands results in a number of adjoining ring systems, and ring closure constraints can only be set for bonds that are not shared by other cyclic systems. This problem can be overcome by 'breaking' the metal ligand bonds during each MC step and bonding them for the subsequent molecular mechanical minimization, see Fig. 13. The MC dihedral angle variation method can be supplemented by MC translational variations in which the position of individual atoms and/or molecular fragments are randomly moved and rotated. If the metal ligand bonds have been disconnected during the dihedral angle variation steps the metal ion can be moved in an additional MC translational step before minimizing the new starting geometry.

1.3.2. Cartesian stochastic (or kick) MC search method

In the Cartesian stochastic (or kick) MC search method [33,34], a random kick is applied to each atom in the molecule, generating a new high energy starting geometry for energy minimization. In a conformational search of cycloheptadecane, it was shown that the maximum kick size should be between 2.7 and 3.1 Å [28]. Larger kicks distorted the molecule so badly that the average time for energy optimization was dramatically increased. The probability of falling back to the starting geometry increased when the kick size was smaller than 1.5 Å. Advantages of this method are that it can be applied to inorganic systems just as easily as to organic systems, and that it requires only eight lines of Fortran to implement.

1.3.3. Molecular dynamics

Molecular dynamics searches have been shown to be less efficient at finding new minima than most conformational searching methods [28]. This is particularly evident when the minima are separated by large energy barriers, as can be expected for most of the inorganic geometric isomers that are formed by metal ligand bonds breaking and reforming. Therefore molecular dynamics sampling is not an efficient inorganic conformational/geometric isomer searching method. However, it can be used to sample the conformational space of coordinated ligand in fixed stereochemistries [35].

1.4. Data analysis in chemistry [36]

Social scientists often work with large multivariate data sets and have developed many useful statistical methods that simplify the interpretation of large multivariate datasets. These techniques have been adapted for chemical analysis and have mainly been applied to analytical problems. The increase in the number of structures in databases such as the CSD, and the large number of conformations generated in conformational searches and in molecular dynamics simulations, have also resulted in the adaptation of multivariate statistical techniques to conformational analyses.

Auf der Hyde [37] and Allen [38] have written excellent tutorials on numerical methods of data analysis in chemistry. Both use analyses of crystallographic data to illustrate the techniques. More than 200 systematic structural correlations using the CSD have been published. Many of these analyses have investigated inorganic complexes, however only a limited number have combined the results with molecular mechanical calculations. They will be discussed in section Sections 3 and 5.

The two most commonly used methods are principal component analysis (PCA) and cluster analysis. In PCA the data is simplified by reducing the dimensionality of the data. In cluster analysis ‘natural’ groupings of similar points in a multidimensional space are found, thereby simplifying the data.

1.4.1. Cluster analysis

The aim of a cluster analysis is to place objects into groups, also called clusters, in such a way that all the objects within a cluster are very similar and that all the objects in different clusters are very dissimilar to each other.

In order to perform a cluster analysis one needs a similarity measure for comparing the data points. These criteria are very dependant on the type of data being analyzed and can be divided into distance, association, probabilistic and correlation measures [39]. Distance measures are the most commonly used similarity criteria in chemistry.

After deciding what measure to use to compare data points, a mathematical strategy for grouping the data points must be chosen. Non-hierarchical mathematical methods in which the ‘natural’ or best clustering is the only grouping produced are not very commonly used in chemistry, except in the Simple Optimized Non-Hierarchical Cluster Analysis (SONHICA) program [40] and the Jarvis–Patrick algorithm [41], which has been modified [42] and included in the GSTAT module [43,44] of the CSD. Hierarchical strategies which form clusters in a stepwise fashion, with later clusters being groupings of earlier ones, are much more common. Dendrograms are often used to show how the clusters develop as a function of the clustering level (i.e. step number). Different hierarchical methods use different measures of dissimilarity to determine the distance between clusters. In the single-link or nearest-neighbor method the distance between the nearest data points in neighboring clusters is used. In the complete-linkage strategy the distance between the furthest neighbors in adjoining clusters is measured, and in Ward's algorithm the centroid of each cluster is used. If a hierarchical method is used, a criteria for choosing the best grouping is also required. Examples of graphic and numeric methods for choosing the best clustering level are described in section Section 5.

Cluster analysis has effectively been used in the contraction of conformational space in the multiconformational analysis of solution NOE data [45,46], drug-discovery [47,48], database searching [49], analytical chemistry [50] and biochemistry [51]. Recently four programs have been released that cluster molecules on the basis of their conformations, three were written primarily for the analysis of structures generated in conformational searches and molecular dynamics simulations [40,52,53] and one was written primarily for conceptual database analysis [54].

The agglomerative, hierarchical, single-link clustering methods is described in a little more detail here as it is one of the more commonly used methods. The first step, in such a cluster analysis, is to find a similarity criterion and arrange all the molecules in a generic ordering according to their similarities. Therefore structure 2 is most similar to structures 1 and 3, while structure 3 is most similar to structures 2 and 4, and very dissimilar to molecules much higher in the genetic ordering. In clustering level 1 all the structures are generically ordered and placed in separate clusters. In each following clustering level the two molecules that are most similar to each other are joined to form a cluster. If the two structures are in separate clusters, the clusters are joined. At the end all the structures will be joined in one large cluster. For example, if N structures are being clustered, at clustering level 1 there will be N clusters each containing 1 structure, at level 2 there will be $N-1$ clusters, and at level N there will be one cluster containing all N structures. Fig. 10 shows a clustering mosaic which is a pictorial representation of an agglomerative, hierarchical, single-link clustering of a set of nickel(II) 14-membered macrocyclic rings.

One of the difficulties in an agglomerative, hierarchical, single-link method is to choose between all the clustering levels and find a level, or levels, at which the clusters are significantly different.

1.4.2. PCA

The goal of PCA is to generate linear combinations of principle components that describe the greatest variance over all the data. Ideally a small number (C) of principal components will describe a large proportion of the total variance. The principle components, $(\mathrm{PC})_j$, can be expressed as linear combinations of the original data [x_i ($i = 1, 2, 3, \ldots, p$) where p is the number of dimensions of the original data] as follows

$$(\mathrm{PC})_j = a_{1j}x_1 + a_{2j}x_2 + \cdots a_{pj}x_p$$

The coefficient a_{ij} indicates the contribution of each variable x_i to the PC_j. Coefficients of zero or close to zero indicate that their associated x_is do not significantly contribute to the PC, while coefficients that are significantly higher than the others are closely aligned to the PC_j. The coordinates of each fragment in the C-dimensional PC-space are known as PC scores. They are often used to generate scatterplots such as those shown in Fig. 1.

One of the difficulties in using PCA is that there is no clear relation between the PC and the structural parameters used in the analysis and it is not always possible to interpret the PCs in chemical terms. PCA can be very useful when there are hidden dependencies in the dataset. Although there is a difference the terms principal component and factor analysis are often used interchangeably [55].

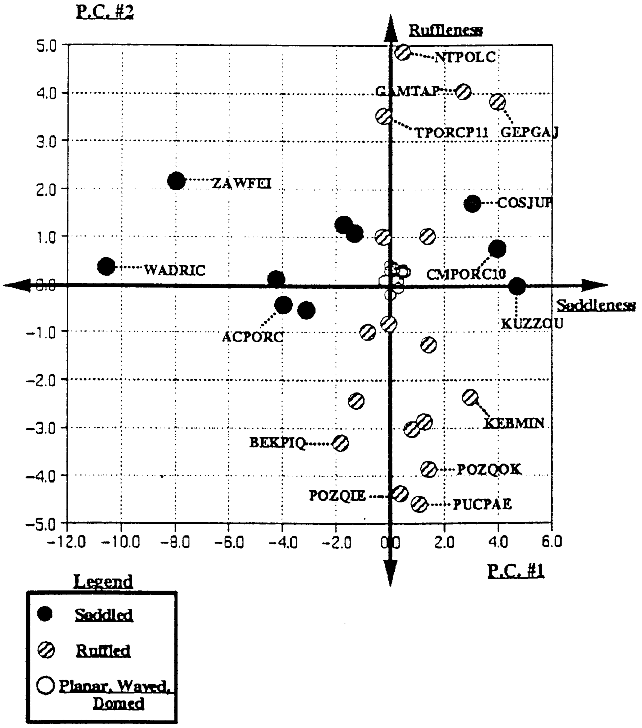


Fig. 1. Scattergram of the first two principle components obtained from a principle component analysis (PCA) of the distances of the C_β, C_{meso} and N atoms to the best plane and the Cα-N-N-Cα dihedral angles of all the cobalt porphyrins in the Cambridge Structure Database (CSD). The first two principal components account for 83% of the variance. The ruffled structures are located along the y axis with the more ruffled structures having higher y values. The saddled structures are along the x axis, while all the planar, domed and waved structures are located at the origin. Structures such as ZAWFEI and COSJUP have both ruffled and saddled character [121].

Table 1
Number of structures in the Cambridge Structural Database (CSD) v5.17

Metal ion (congeneric family)	No. structures
Sc	99
Ti	2245
V	1896
Cr	3415
Mn	3614
Fe	10 653
Co	7364
Ni	6258
Cu	10 472
Zn	2695
Transition metal (Porphyrin)	1129
Transition metal (14-Ane-N_4)	632
Transition metal (9-Ane-N_3)	594

2. Data analysis of first-row transition metal complexes in the CSD

The latest version of the CSD, version 5.17 has 197,481 X-ray and neutron diffraction structures of organocarbon compounds [56,57]. All the compounds in the CSD have less than 500 atoms, while those in the Protein Database (PDB) have more than 500 atoms [58]. Ninety thousand five hundred and eighty of the structures in the CSD contain one or more transition metal ion, see Table 1. There are many structures in the CSD that have a common backbone (e.g. porphyrins or 14-membered macrocycles) in this review we will refer to these groups as congeneric families.

Analyses of congeneric families are very useful as they can reveal the different conformations the backbone structure can adopt in the different environments found in the crystals. This can provide information about [59]:

- the conformations available to the backbone
- how the conformers can interconvert
- the environmental factors that are responsible for certain conformations.

Conformational analysis of the crystallographic data in the CSD can be done without a subsequent molecular mechanics analysis. In fact, this is the most common way of analyzing the structures in the CSD, and therefore examples of this type of analysis will be presented. However in order to extract the maximum amount of information from the conformational differences that occur within congeneric families, a molecular mechanical analysis is needed. This is particularly true for inorganic compounds because the force field parameters are derived from solid state structures and the conformations resulting from inorganic molecular mechanics calculations are ‘crystal averaged’ structures. One can therefore compare the calculated conformations with the conformations observed in the crystal structures that should be close to one or more minima on the PES. Furthermore if one assumes that the backbone will adopt a low energy conformation in most crystal structures, then the conformational space spanned by the crystal conformations can be equated and compared (qualitatively not quantitatively) with the PES of the backbone itself [60].

Although the majority of studies based on the analysis of the CSD are correlation studies that do not employ force field methods, the focus of this paper is on the use of correlation analysis used in conjunction with molecular mechanical techniques. Therefore I will only present one example of a CSD analysis not using molecular mechanics calculations. Auf der Hyde [61] published a review of all the papers that study reaction paths based on correlation analysis of solid state structures. It covers 40 papers published before 1991 and although it is slightly dated gives a good overview of the field.

The work of Huttner [62] will be presented as the sole example of a non-molecular mechanics correlation analysis. He has used cluster, neural network and PCA to study 82 cobalt complexes with the tripod ligand ($CH_3C(CH_2PPh_2)_3$) [62]. The ligand and all the parameters used in the correlation analyses are shown in Fig. 2.

A PCA of the six φ values and the three τ values for each compound revealed that two components described 78.7% of the total variance. The first component described a strong negative correlation between the φ and τ values, when φ increase τ decreases and vice versa. There seems to be no straightforward interpretation for the second component. A neural network analysis of the same nine parameters grouped the 82 complexes into three groups, Group I had large τ values and negative φ values, Group II had large positive φ and small τ values, and Group III had conformations between these two extremes. Hierarchical clustering analysis produces similar results, which can be seen by inspection of Fig. 3.

The figure compares the results from the three analyses and shows the dendrogram obtained in the cluster analysis. The clustering analysis also produces three clusters A, B, and those in between. All the classification procedures gave very similar results. Besides comparing the different statistical methods the authors also concluded that intra-tripod forces dominated the forces responsible for the conformations adopted by the tripod, and that the crystal forces and forces due to

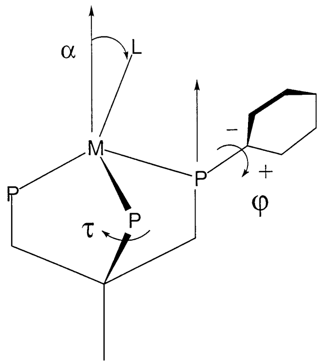


Fig. 2. Definition of the parameters used in the analysis of cobalt tripod complexes.

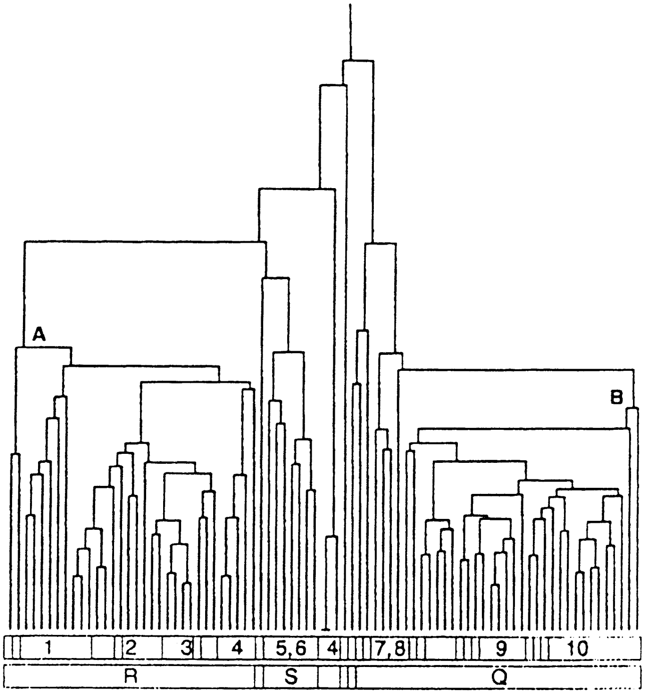


Fig. 3. Dendrogram obtained by the hierarchical cluster analysis of the six φ and three τ values for the 82 tripod cobalt complexes found in Cambridge Structural Database (CSD). The boxes beneath the dendrogram refer to the results from the neural network analysis (neurons 1–10) and principal component analysis (PCA) (regions R, S and Q).

differing coordination numbers as well as the differing nature of the coligands are not important to the conformations the tripod ligand adopts.

3. Data and molecular mechanical analysis of first-row transition metal complexes in the CSD

The structures in the CSD can be used to derive parameters for molecular mechanical analysis of first-row transition metal complexes [3,63–65], individual CSD structures can be analyzed by inorganic MM calculations [66,67] or they can be used to interpret trends observed in the analyses of the first-row transition metal complexes in the CSD. In this review we will only discuss the latter type of analyses, there are not many.

Raithby and co-workers have published three articles in which they have undertaken a systematic conformational analyses of metal and metal-free 12-, 14-, 15- and 16-membered oxa and thia macrocyclic ligands and their derivatives [68–70], see Fig. 4. The authors are of the view that “The problem of including metals in molecular mechanics calculations has not yet been fully resolved" [68]. As a result they only used molecular mechanics calculations to find the low energy conformations of the metal free ligands. In all cases the Jarvis–Patrick algorithm was used to cluster the macrocycles based on their intra-annular torsion angles, after they had been symmetry adapted [71]. The Jarvis–Patrick method is a non-hierarchical method that is part of the GSAT program [43,44].

The 12-, 14-, 15- and 16-membered oxa and thia macrocyclic ligands are fairly flexible ligands that can adopt many different conformations. The different conformations are named using Dale's nomenclature [72,73]: two adjacent *gauche* torsions of the same sign are counted as corners of the ligand and the length of the sides between the corners are sequentially listed between square brackets. Two adjacent *gauche* torsions with different signs are called pseudo-corners and are ignored. The starting point and direction are chosen so that one starts with the shortest length (i.e. smallest number in square brackets). In this article we have also italicized the conformations to differentiate them from references.

All the conformations of the 12-, 14-, 15- and 16-membered oxa and thia macrocyclic ligands were successfully clustered using symmetry-modified Jarvis–Patrick clustering. While PCA of the 12-, 14- and 16-membered macrocycles was able to represent the different conformations in torsional space, the conformations adopted by the 15-membered rings were not well by represented by PCA.

One hundred and thirty-two 12-Ane-O_4 fragments were found in v5.09 of the CSD, 117 coordinated a metal ion through all four oxygens and 13 were metal free macrocycles. The most common conformations found for the metal-free 12-membered oxa macrocycles are the *[3333]* and *[66]* conformations. A molecular mechanics analysis of all the possible conformations revealed that the *[3333]* and *[66]*

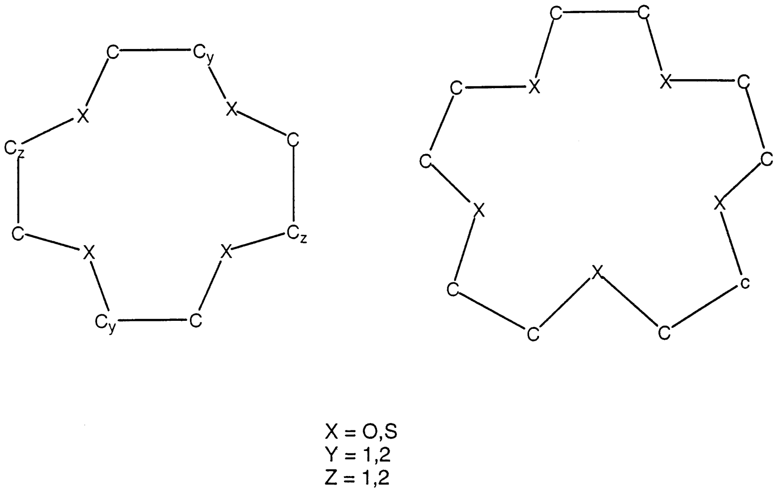


Fig. 4. A systematic conformational analyses of metal and metal-free 12-, 14-, 15- and 16-membered oxa and thia macrocyclic ligands and their derivatives was undertaken by Raithby and co-workers [68–70].

conformations are the lowest energy forms. When the macrocycles are bound to metal ions they are most commonly found to adopt a *[3333]* conformation if the metal has a square-planar, square pyramidal, seven coordinate 4:3 or square anti-prismatic geometry. Metals with a *cis* octahedral geometry coordinate the 12-membered oxa macrocycles in a *[48]* conformation. The conformations of the 12-membered rings when they were coordinated to a metal ion (i.e. mainly *[3333]* and *[66]*) did not differ much from the conformations in the low energy metal free conformations. Therefore, one can assume that the ligands were preorganized for metal binding [74]. The authors concluded that they were able to find the low energy conformations of metal bound 12-membered oxa macrocycles without having to resort to inorganic molecular mechanics calculations by noting the frequency of conformational occurrences [68].

One hundred and seventeen 15-Ane-O_5 fragments were found in v5.09 of the CSD, 99 of these bind a metal through all five oxygens and 15 are metal free macrocycles. Only ten 15-Ane-S_5 fragments were found. Due to the low number of thia macrocycles they were not analyzed in the same detail as the oxa macrocycles. The most commonly found 15-Ane-O_5 fragments were the ones that had the lowest molecular mechanics energies for the free ligand. The energy differences between the different conformations were found to be less than 10 kJ mol^{-1}. Small main group and transition metal ions were bound to macrocycles with anangular conformations, while larger cations (e.g. lanthanides) were coordinated to ligands with uniangular and *[348]* conformations [69].

Thia 14-membered macrocycles were found to be more common than the 14-Ane-O_4 fragments; 11 14-Ane-O_4, 32 14-Ane-S_4 and one 14-Ane-Se_4 macrocycle were reported [70]. Sulfur donors also predominate in the 16-membered fragments found in v5.09 of the CSD (two with oxygen, 34 with sulfur and two with Se donor atoms). As was the case for 15-membered rings, PCA was not able to reduce the dimensionality of the torsional space.

The metal free 14- and 16-membered oxa fragments in the CSD all have low molecular mechanics energy conformations. The conformations of the metal coordinated ligands are related to the amount of strain resulting from the conformation that is adopted by the ligand in order to direct the donor atoms towards the metal ion and maintain the correct metal–oxygen distance.

Just like the conformations of the metal free 14- and 16-membered oxa macrocycles, the conformations of their thia analogs are low energy conformations. Although certain trends have been observed, predicting the conformations of the metal bound thia ligands is not as straightforward as for the oxa-macrocycles.

All oxa-macrocycles studied adopt different low energy conformations from their carbocyclic analogs. This is due to the fact that the CCCC dihedral angles in carbocycles preferably adopt *gauche* conformations while the *gauche* CCOC dihedral angles in oxa-macrocycles are destabilized by the resultant short 1,4-hydrogen interactions.

Our work described in Section 5 reinforces the idea presented here [68–70] that low energy conformations of metal bound macrocycles can be found by finding the frequency of the conformational occurrences in the CSD. However, we have gone

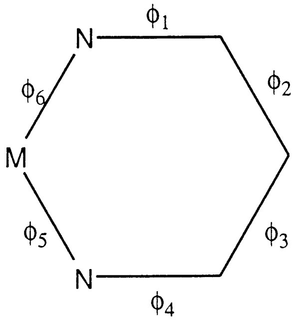


Fig. 5. The substructure used in the Cambridge Structural Database (CSD) search, M = nickel(II), copper(II) or cobalt(III), and all nitrogens as well as the carbons are sp^3 hybridized.

on to show that inorganic molecular mechanics calculations used in conjunction with conformational searches are able to find the same low energy conformations.

We have used cluster and molecular mechanical methods to analyze the conformations of cobalt(III) [75], copper(II) [76] and nickel(II) [77] propanediamine and substituted propanediamine ring systems. A total of 769 6-membered rings containing the substructure shown in Fig. 5 with nickel(II), cobalt(III) and copper(II) were obtained from the CSD search. Three hundred and eighty-seven of the rings were found with nickel(II), 174 for copper(II) and 208 for cobalt(III) [78].

Regardless of the metal being considered the same clusters were obtained at high clustering levels when using atomic overlap or torsional differences as similarity criteria; however, at lower clustering levels there were noticeable differences [78]. The reason for this is fairly obvious; when the dihedral angles are used as a measure of similarity, the lengths of the Ni–N, Cu–N and Co–N bonds are not important. However, if the difference in Cartesian coordinates between identical pairs of atoms is used, then differing bond lengths will result in larger differences and therefore in different clusters. The fact that the same clusters are obtained at the more important higher clustering levels indicates that the agglomerative, hierarchical, single-link clustering method is a fairly robust technique. The cluster analysis of nickel 14-membered tetraazamacrocycles, described in Section 5, will be used to illustrate the method.

As expected, the chair conformation is the most common conformation for all three metal ions. For cobalt(III) and nickel(II) there are more boat conformations than one would expect on the basis of the PES, which has the boat conformation as a transition state higher in energy than either the twist-boat or chair. In fact, minimizing a cobalt, nickel or copper propanediamine moiety that is in the boat conformation always results in a twist-boat energy minimum structure. Although some boat conformations are found for copper(II) there are significantly fewer than there are for cobalt(III) and nickel(II).

In order to establish whether this could be due to the difference in metal ion sizes, we did some bite-size calculations on hypothetical cobalt(III), nickel(II) and copper(II) propanediamine complexes (i.e. Co(pn); Ni(pn) and Cu(pn)). The energies of the different conformations were obtained as a function of the metal–ligand distance using the method of Hancock [79]. Similar results were obtained when using Wiberg's method [80]. Since minimizing the boat conformation always results in a twist-boat conformation, we had to constrain the M(pn) to a boat by fixing the intraannular dihedral angles. The results from the bite-size calculations are shown in Fig. 6. The boat conformation has the smallest bite-size with an ideal metal–nitrogen distance of 1.84 Å, the chair has an ideal metal–nitrogen distance of 1.92 Å and the twist-boat prefers metal–nitrogen distances of 2.18 Å. The twist-boat conformation has the broadest parabola in Fig. 6 which indicates that it is the most flexible conformation [78].

The boat conformations are only found when the 6-membered rings are forced into the boat conformation by other parts of the compound; such compounds all have the have the substructures shown in Fig. 7. The fact that cobalt and nickel form 6-membered rings with boat conformations in the solid state more often than copper can be attributed to two factors. Firstly, the copper–nitrogen distance is slightly longer than for the other two metal ions and therefore the energy difference between the boat and chair conformations is larger. Secondly, the boat conforma-

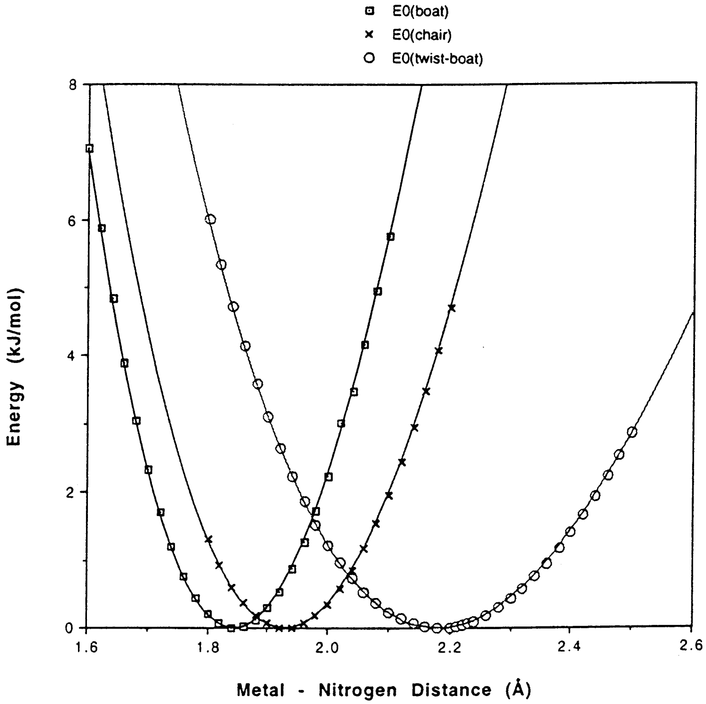


Fig. 6. A plot of 'relative energy' as a function of the metal–nitrogen distance in a metal propanediamine complex. The energies of the different conformations cannot be compared in these calculations as the boat conformation had to be constrained, while the other two conformations were not. For this reason we took the minimum energy of each conformation as having 0.00 kJ mol^{-1} relative energy.

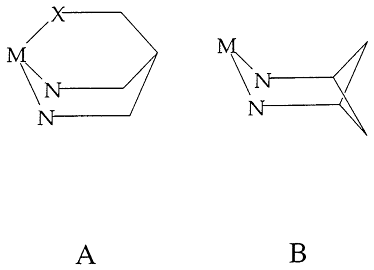


Fig. 7. Structural motifs responsible for all the boat and twist-boat conformations.

tions are only found in compounds having the substructure shown in Fig. 7, compounds having these substructures are most commonly prepared by template syntheses and copper does not lend itself to all template syntheses. Furthermore, copper(II) undergoes a Jahn–Teller distortion which lengthens the axial bonds beyond 2.10 Å, further decreasing the chances of forming boat conformations in compounds having the substructure shown in Fig. 7A.

Encapsulated transition metal compounds are good illustrations of how these factors influence the formation of boat conformations. A search of the CSD found 24 complexes of cobalt(III) with octahedral encapsulating ligands, three nickel(II) complexes and one copper(II) complex. The reason for this is that cobalt(III) is a ‘good’ templating center, while copper and nickel are not [81]. The copper and nickel complexes that have been crystallized were either formed in very low yield or were obtained from the reaction with the free ligand which is not easily synthesized [82]. Even though a copper hexaamine macrobicyclic complex has been crystallized [82] the structure is not rigid enough to force the 6-membered rings to adopt boat conformations, and although they have the substructure shown in Fig. 7A, they adopt strained twist-boats.

This example shows that the distribution of conformations belonging to a congeneric family in the CSD does not exclusively depend on the PES but is also influenced by other factors such as the chemical properties of the metal ions.

The structures in the CSD are not unique and the structures in the protein databank (pdb now located at http://www.rcsb.org/pdb/index.html) could also be used in correlation analyses. The only drawbacks with structures in the pdb is that the resolution is much lower that that for the structures in the CSD, and that the CSD has more user friendly search and analysis tools than the pdb. However, the pdb is rapidly improving, especially with the addition of the metalloprotein site database and browser (http://metallo.scripps.edu), which contains more than 3000 metalloproteins with a user friendly search interface [83]. Shelnutt [84,85] has analyzed 800 nonplanar structures of heme proteins in the pdb by normal-coordinate structural decomposition [86] and combined the results with molecular mechanics calculations [87].

4. Data and conformational analysis of first-row transition metal complexes in the CSD

There has only been one attempt to use the information in the CSD to determine the conformations available to inorganic compounds and although it does not involve molecular mechanics it has been included in this review.

The constitutional–configurational–conformational (CCC) program [88] uses the structural information in the CSD in a very interesting manner. It is an artificial intelligence program that utilizes a knowledge base of all the conformations adopted by common molecular fragments to do conformational searches. Molecular mechanics calculations are the method of choice for calculating the conformations available to organic fragments, while information from the CSD is the source of the inorganic knowledge base.

The best way to describe how the knowledge base and heuristics are used to find all the constitutional, configurational and conformational isomers available to transition metal coordination complexes is by way of an example.

$[Co(III)(trien)(gly)]^{2+}$ can adopt ten configurations. These arise from the different ways the two ligands can be arranged around the octahedral metal ion and from the two secondary amines which become asymmetric upon coordination, see Fig. 8.

A substructure search algorithm establishes which of the fragments in the knowledge base can be mapped onto the molecule. If the complex contains a metal ion it is assigned a geometry, e.g. cobalt(III) is octahedral. The ligands are inspected to find all possible ligating atoms and two heuristic rules are applied to propose possible constitutional isomers. The first rule states that if two donor atoms of a multidentate ligand bind to the metal ion to form a ring of seven or more atoms, in which there is a third donor atom that is not bound to the metal, this constitutional isomer is eliminated. For the trien ligand this means that all four nitrogen atoms will be coordinated to the cobalt ion. The second rule selects complexes that have the least number of ligands bound to the metal and with multidendate ligands making as many bonds as possible to the metal. This rule is derived from the chelate effect. Applying both rules together will produce a $[Co(III)(trien)(gly)]^{2+}$ complex rather than a $[Co(III)(gly)_2]^{+}$ complex or a $[Co(III)(trien)_2]$ complex in which two of the trien nitrogens are not coordinated.

After establishing the constitutional form of the complex, the configurations available to it need to be determined. In transition metal complexes, the configurations are a consequence of the way the ligand can wrap around the metal ion and of donor atoms which become asymmetric upon complexation. Two algorithms are therefore required. One determines which donor atoms become asymmetric when bonded to the metal and generates all the possible combinations of stereochemistries. The other arranges the donor atoms around the metal. There are 30 different ways in which six donor atoms can be arranged around a octahedral cobalt(III) ion; however, heuristic rules eliminate a number of them. For example, adjacent ligand atoms in a multidentate ligand cannot be *trans* to each other. An algorithm [89] is also used to eliminate all equivalent configurations.

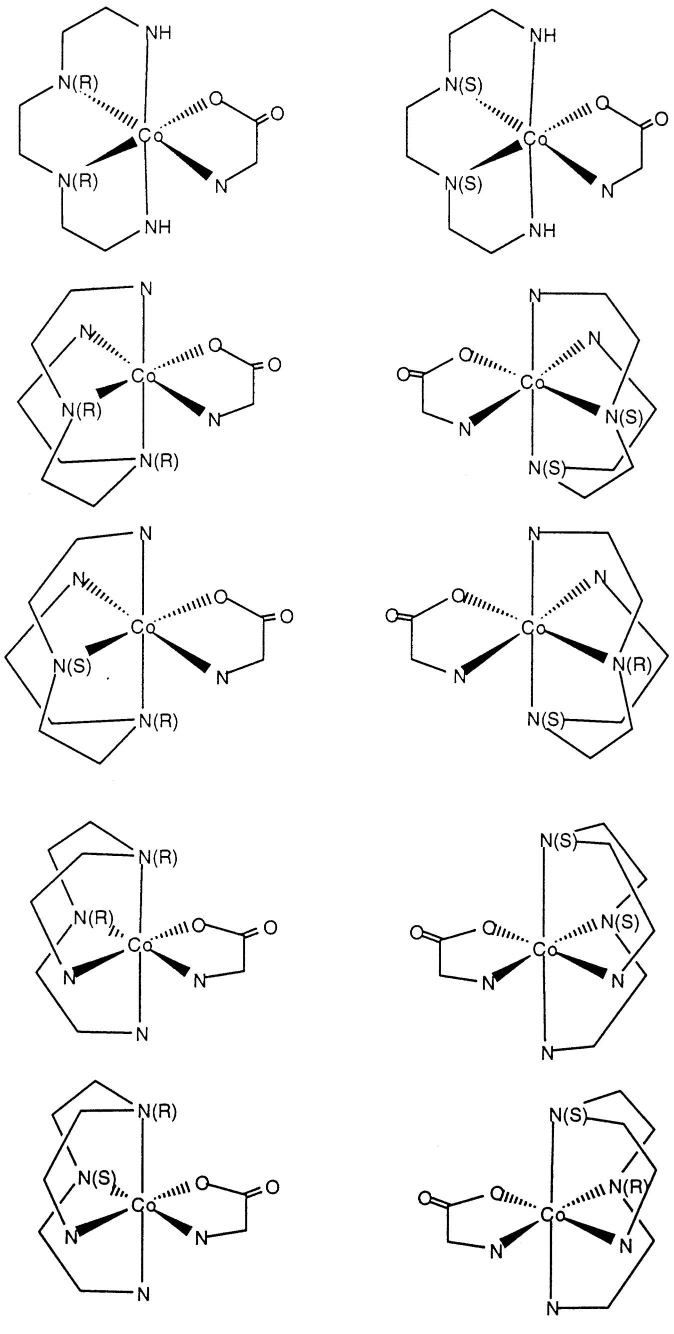


Fig. 8. The ten different configurations adopted by $[Co(III)(trien)(gly)]^{2+}$. These arise from the different ways the two ligands can be arranged around the octahedral metal ion and from the two secondary amines which become asymmetric upon coordination.

Conformational searches of all the configurations are done using the COBRA program [90], which searches through its knowledge base to find fragments that can be combined to make-up the complex being studied. The $[Co(III)(trien)(gly)]^{2+}$ complex would be patched together by adding three fragments, gly, trien and cobalt(III). If some parts of the molecule are made up of fragments not found in the knowledge base the best possible fragments are used (e.g. replacing a sp^3 nitrogen with another sp^3 atom). The fragments are patched together to make up the molecule and all the combinations of the different conformations available to the fragments are generated, using the knowledge base obtained from the CSD.

The CSD has five crystal structures of $[Co(III)(trien)(gly)]^{2+}$, each with a different conformation [91,92]. All five conformations were found in a CCC search [88], with RMS fits between theoretical and experimental structures of less than 0.1 Å for all structures.

The method has three drawbacks:

- It assumes that the conformations of the fragments interact independently of each other.
- In order to be effective the knowledge base has to contain all the conformations a fragment can adopt.
- The technique can only be used if all the fragments (or related fragments) in the molecule are found in the knowledge base.

5. Molecular mechanics, data and conformational analysis of first-row transition metal complexes in the CSD

By the time this review comes out there will be more than 200 000 structures in the CSD and there are many congeneric families that are well represented in the database. If one examines enough structures of a congeneric family one should get a good representation of all the conformations available to the common backbone of the family, and one might even get a reflection of the PES. The same type of information can be obtained from a combination of molecular mechanics and conformational searching. Although the combination is a logical one it is not commonly used in inorganic chemistry, mainly because inorganic conformational search methods are not found in most molecular mechanics programs. Perhaps the introduction of inorganic molecular mechanics conformational search modules into MacroModel [93] and MOMEC [94,95] will result in more work in this area. MM conformational searches have been used to interpret trends in the CSD in many organic and biological studies, e.g. Lipkowitz's analysis of cyclodextrin structures [96,97].

We have used a combination of MM, conformational searching and database analysis to explain configurational/conformational trends observed in the solid state structures of first-row transition metal tetraazatetradecane macrocyclic complexes. The fact that macrocyclic complexes are thermodynamically more stable than their open-chain analogues, and that they are more selective metal ion binders, has led to a large amount of research [98–103]. One of the most commonly studied and used macrocycles is the 14-membered macrocycle 1,4,8,11-tetraazacyclotetradecane (cyclam, 14-ane-N_4).

Fourteen-membered tetraazamacrocycles, such as cyclam and its derivatives, are fairly flexible and can adopt any of the five configurations shown in Fig. 9, in both folded (*cis*) and planar (*trans*) geometries. The energies of the different configurations have been estimated on the basis of molecular models [104,105] and calculated using molecular mechanics [106–109], semiempirical methods [110] and local density functional theory [111].

Our cluster analysis of cobalt(III) [112], copper(II) [113] and nickel [114] complexes containing the 14-membered tetraaza substructure shown in Fig. 9 found that the macrocyclic ring structures were effectively separated according to their conformation by an agglomerative, hierarchical, single-link method. For cobalt(III), in which the metal ion is always octahedral similar results were obtained when all external dihedral angles, all intraannular (i.e. in the 5- and 6-membered rings) dihedrals, or a heavy atom atomic overlap was used as a clustering criteria. However, for the copper(II) and nickel complexes the most successful clustering

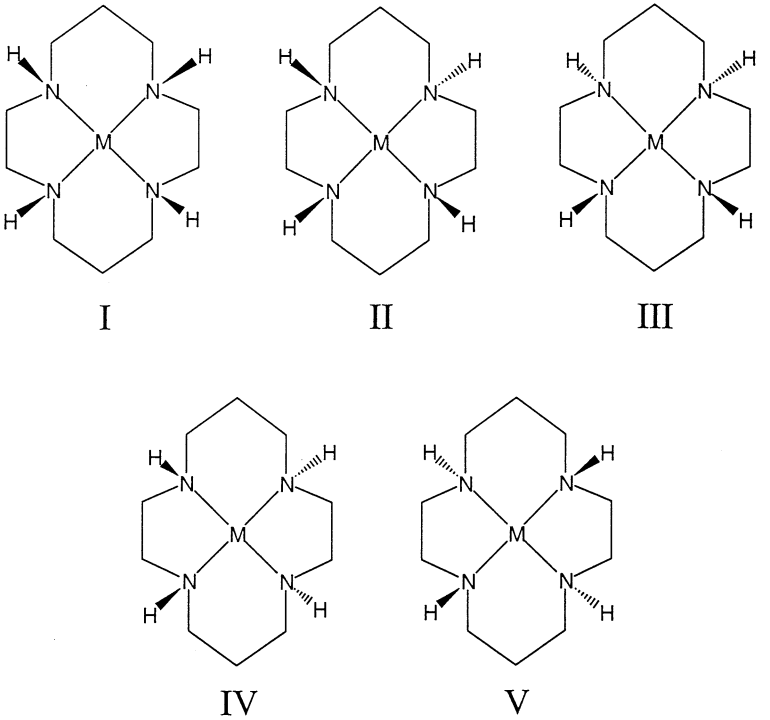


Fig. 9. Fourteen-membered tetraaza macrocycles can adopt the five different configurations shown above in both a planar (*trans*) and a folded (*cis*) geometry. The Cambridge Structure Database (CSD) search was conducted for all nickel, cobalt or copper complex having the macrocyclic backbone with any substituents on the any of the positions. However the macrocycles with double bonds in their backbone were rejected.

criteria was the rms displacement between identically numbered atomic pairs after rigid body superimposition. The cluster analyses separated all the different configurations shown in Fig. 9 from each other, and at lower clustering levels they also separated different conformations with the same configuration from each other. The structures were grouped according to their different configurations even though the clustering was based on the conformation of the macrocyclic backbone without taking into account the positions of the amine hydrogens.

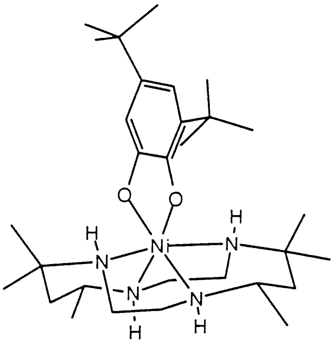
The cluster analysis of all the 6- and 14-membered macrocyclic complex followed similar procedures [75–78,112–114]. The following paragraphs describe the method using the analysis of nickel 14-membered tetraazamacrocycles [114] as an example. The same example will subsequently be used to describe how conformational searching and molecular mechanics calculations can be used in combination with the cluster analysis.

One hundred and thirty-nine nickel complexes containing the 1,4,8,11-tetraazacyclotetradecane backbone were found in the CSD. Examples of these structures are shown in Fig. 10. Five structures contained two or three tetraaza rings. These rings were separated so that they could be evaluated individually, resulting in 150 structures. One hundred and thirty-nine of the complexes were nickel(II), while only nine were nickel(III) and two were nickel(I). The properties of 1,4,8,11,-tetraazacyclotetradecane metal complexes can be altered by adding pendant arms to the macrocycle. These arms can bind to the metal raising the denticity of the macrocycle. Fifteen of the macrocycles bound the metal through one arm (e.g. fagtow [115] in Fig. 10) and 14 bound with two pendant arms (e.g. yiwbah [116] in Fig. 10). There were five macrobicyclic complexes. There were seventeen complexes with bidentate counterions, such as gibsal [117] which is shown in Fig. 10. They all had a *cis* V configuration. Eleven structures, such as zudjud [118] which is shown in Fig. 10, have one or more than one amide group.

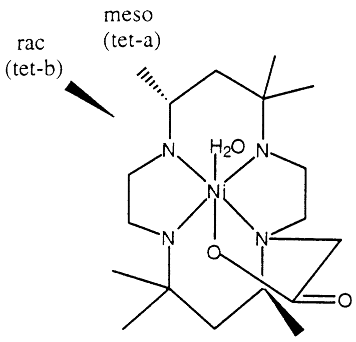
All hydrogen and substituent atoms were removed from the 1,4,8,11,-tetraazacyclotetradecane backbone so that each structure was composed of one nickel, four nitrogens and ten carbons. Xcluster [52] was used to perform three separate cluster analyses on the 150 structures. Each analysis was based on a different clustering criteria. The most successful attempt clustered the macrocycles according to the rms displacement between identically numbered atomic pairs after rigid body superimposition. The other two attempts, which based the clustering on the rms differences between the internal and external torsion angles, had trouble separating *cis* and *trans* structures. All discussions about cluster analysis from this point onward are based on the clustering with pairwise atomic overlap used as the clustering criteria.

The clustering mosaic and distance map, shown in Figs. 11 and 12, respectively, and the separation ratio indicated that the most significant clustering was at level 124. At a given clustering level, the separation ratio is the ratio of the largest conformational distance between any two structures in the same cluster at that level, to the shortest conformational distance between any two structures that are not in the same cluster at that clustering level.

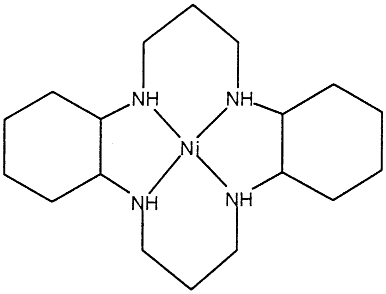
The distance map (Fig. 12), shows a matrix in which conformational similarity is indicated by the color of the cross peaks; the more similar two structures are, the



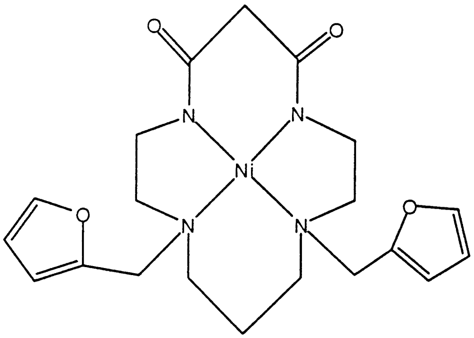
gibsal



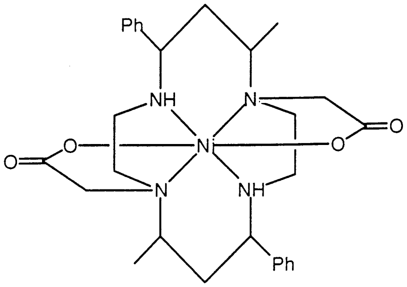
fagtow



kimwei, vovvil



zudjud



yiwbah

Fig. 10. Structures of gibsal [117], fagtow [115], kimwei [122], zudjud [118] and yiwbah [116].

darker the cross peak. Since all the diagonal elements indicate the similarity between identical conformations they should be black; however, they have been changed to white to increase the clarity of the map. Figs. 11 and 12 show that cluster 1 is the largest cluster. It contains 81 structures, all of which have *trans* III configurations. All but two of the members of cluster 1 have both of their 6-membered rings in the stable chair conformation and a λ and a δ 5-membered ring. The two exceptions to the general conformation are the last two macrocycles in the cluster. Limxou [119] and zudjud [118] both have one chair and one boat.

Besides the *trans* III macrocycles located in cluster 1, there are also four macrocycles with a *trans* III configuration in cluster 23. All four have both of their 6-membered rings in the boat conformation. They are located in a separate cluster from the other *trans* III macrocycles because the clustering is based on the macrocyclic conformation, not configuration.

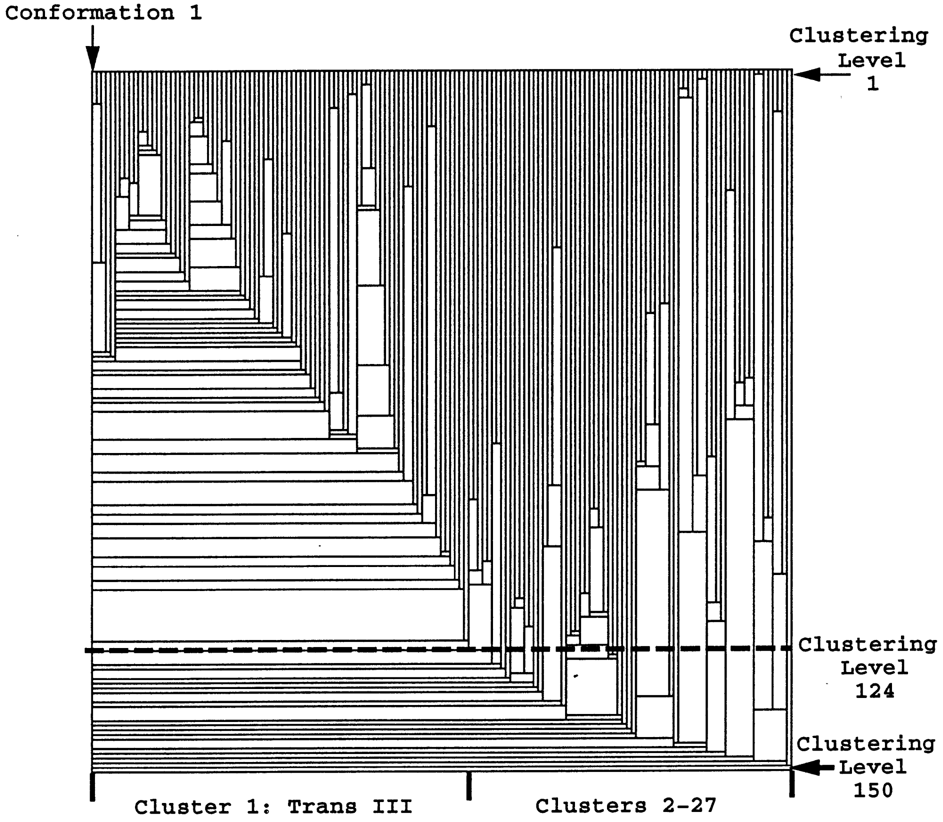


Fig. 11. At the top of the clustering mosaic, each narrow band corresponds to one of the 150 structures, each of which is in its own cluster. The first structure (on the far left) was chosen randomly and all the others are ordered according to their conformational proximity to the first structure. In each following level the two closest clusters are joined, so that in clustering level *N* there are 150-*N* clusters. The clusters joined at low clustering levels, and shown near the top of the figure, contain conformationally similar macrocycles, whereas those joined at higher levels are not very similar. At the bottom of the mosaic, there is one long band, corresponding to the final clustering level, in which all the structures are joined into one cluster. Clustering level 124 is indicated by a dashed line, and each block along the line is one of the level's 27 clusters.

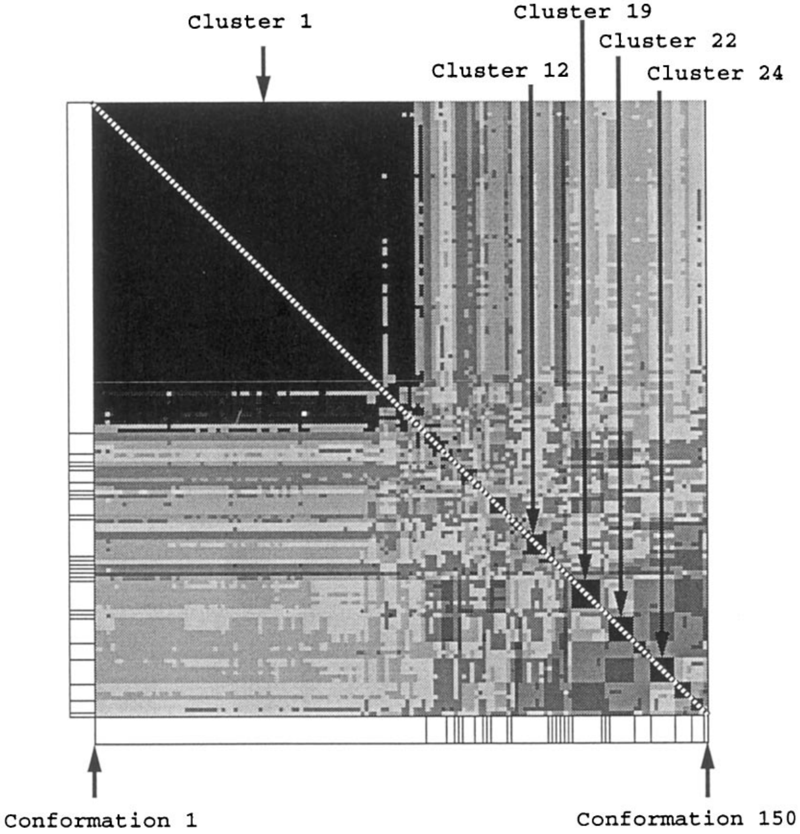


Fig. 12. The distance map showing clusters 1, 12, 19, 22 and 24. The cross-coordinates of conformationally similar macrocycles are darkly colored, while dissimilar structures are lightly colored.

All the macrocycles in clusters 2–17 are in the *trans* I and *trans* V configuration. Cluster 18 contains tetraazacyclotetradecan-2,5,7,10,13-pentanone (vixlix [120]). It cannot be assigned a configuration because its amide nitrogens are deprotonated. Every macrocycle in clusters 19–27 is in the *cis* V configuration, except those in cluster 23 which have *trans* III configurations. No *trans* II or IV or *cis* I, II, III, or IV macrocycles were found.

The agglomerative hierarchical single link clustering method used is very effective at separating the different configurations of cobalt, copper and nickel 14-membered macrocycles. Because pairwise atomic superimposition was used as a clustering criteria, the analysis was able to separate macrocycles with identical configurations, but with different 6-membered ring conformations.

The results of a cluster analysis, such as the one described above for nickel 14-membered macrocyclic complexes, can be used to provide some information about the PES of the complexes being analyzed, especially if sufficient complexes are used. For example the number of *trans* I macrocycles relative to the *trans* III macrocycles was higher for square planar complexes than for octahedral complexes, see Table 2. This confirms the molecular mechanical calculations which show that the *trans* I configuration becomes more stable relative to the *trans* III configuration in going from octahedral nickel(II) cyclam complexes to their square-planar analogues [107]. The majority (77.8%) of octahedral nickel(II) complexes with a 1,4,8,11-substituted 1,4,8,11-tetraazacyclotetradecane backbone adopted a *trans* III configuration, while only 22.3% were *trans* I. For the square planar, square pyramidal, and trigonal bipyramidal molecules, the situation was reversed. Only 26.7% adopted the *trans* III configuration, while 73.3% had the preferred *trans* I configuration. These preferences support the molecular mechanical calculations that the *trans* I configuration is the most stable configuration of square-planar 1,4,8,11-tetramethyl-1,4,8,11-tetraazacyclotetradecane nickel(II), while *trans* III is favored by the octahedral complex [10] (Table 3).

In some cases the addition and/or location of substituents on the macrocyclic backbone can change the identity of the lowest energy configuration, this was most dramatically shown for 5,5,7,12,12,14 hexamethyl-1,4,8,11-tetraazacyclotetradecane nickel(II) complexes. All the octahedral nickel (II) tet-b complexes (see Fig. 9) found in our CSD search were in the *cis* V configuration. On the other hand all but three octahedral nickel(II) tet-a complexes were found in the *trans* III configuration. In order to compare the PES described by the structures in the CSD with the results from molecular mechanics calculations it is important to compare all possible configurations and conformations. The number of conformations available to inorganic complexes is often quite large and the most common practice in inorganic

Table 2
Configurational distribution of nickel(II) 1,4,8,11-tetraazatetradecane macrocycles (amide containing macrocycles are not included)[a]

Config.	Overall (%)	Oct (%)	Spl (%)	Spy (%)	Tbp (%)	Ni(II) L.S.	Ni(II) H.S.
# Structures	138	75	58	4	1	54	74
Trans I	20.3	9.33	29.3	75.0	100	29.6	14.9
Trans III	57.2	52.0	67.2	25.0	0.00	66.7	45.9
Trans V	1.45	0.00	3.45	0.00	0.00	3.70	0.0
Cis V	21.0	38.7	0.00	0.00	0.00	0.00	39.2

[a] Oct, octahedral; Spl, square-planar; Spy, square-pyramidal; Tbp, trigonal bipyramidal.

Table 3
Configurational and conformational distribution of 1,4,8,11-substituted 1,4,8,11-tetradecane macrocycles (%)[a]

	Octahedral	Spl, Spy, and Tbp
Trans III	77.8	26.7
Trans I	22.3	73.3

[a] Spl, square-planar; Spy, square-pyramidal; Tbp, trigonal bipyramidal.

molecular mechanics is to ignore all isomers except the one of interest or have enter all the possible isomers graphically and minimize them individually [3,17]. Since this is a frequently occurring problem in inorganic molecular mechanics we used three test complexes, $[Co(dien)_2]^{3+}$, $[Co(dien)(dpt)]^{3+}$, and $[Co(hexamethyl\text{-}cyclam)(Cl_2)]^{3+}$, to evaluate the ability of the random kick (Cartesian stochastic MC search) method and the MC dihedral and positional method to find all conformations and geometric isomers of typical coordination compounds [30]. Our studies showed that both the MC dihedral and translational variation search and the random kick MC search are significant improvements on the current method by which all possible isomers are entered graphically and minimized individually. These conformational search methods allow the inorganic molecular mechanics practitioner to use a one step procedure to find all low energy conformations as well as all structural isomers. The techniques are much quicker at finding low energy conformations and structural isomers than the commonly used manual method. Furthermore, both the MC dihedral and translational variation method and the random kick approach are capable of finding conformations that are far from ideal and which would not be found if all possible ideal isomers were entered and minimized individually.

The hexamethyl cyclam ligands, tet-a and tet-b may coordinate in a planar (*trans*-configuration of the two monodentate donors) or folded (*cis*-configuration of the two monodentate donors) form. If the two ligands tet-a and tet-b are considered as one species, then the coordinated ligand has six chiral centers, i.e. the four coordinated amines (N1, N4, N8, N11) and the two methyl-substituted carbon atoms (7C and 14C) (Fig. 9). This gives rise to 64 stereoisomers for the *trans*-configuration of the complex and to 2· 64, stereoisomers of the *cis* configuration because there is no pseudo-symmetry plane through the MN_4 fragment that relates the two monodentate donors. Not all of these configurations are non-degenerate, i.e. some are symmetry related [note that this analysis does not include different conformations due to the 5-membered chelate rings (λ or δ) and the 6-membered chelate rings (chair, twist boat)]. The N-based configurations RSRS, RSRR, RSSR, RRSS and RRRR have been named *trans* I, *trans* II, *trans* III, *trans* IV and *trans* V and a similar nomenclature may be adapted for the *cis*-configurations, see Fig. 9. Due to the corresponding pseudo-symmetries of the square planar M (tet-n) fragment (only considering the *N*-configurations; C_2, C_i, C_1, C_2) the number of non-degenerate configurations for the *trans*-geometry is 16; for the *cis*-geometry there are 32 independent configurations (note again that the number of isomers is much larger due to the chelate ring conformations). Therefore, if the chirality of carbons-7 and -14 is not restrained, then a conformational search starting from $[Co(tet\text{-}a)Cl_2]^+$ or from $[Co(tet\text{-}b)Cl_2]^+$ should find 16 *trans* isomers and 32 *cis* isomers.

A MC dihedral and translational variation search found all the possible isomers of nickel, cobalt and copper tet-a and tet-b, however a large number of them were highly strained. In the search all external torsion angles (i.e. dihedrals not involving the metal ion) were varied between 0 and 180°. Metal and monodentate ligands were translated by between 0 and 0.5 Å, ring closure atoms were defined with

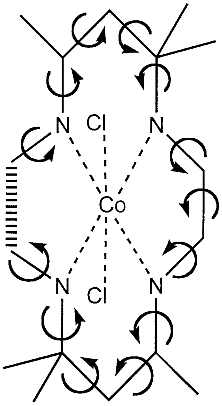


Fig. 13. The Monte Carlo (MC) dihedral and translational variation search of hexamethylcyclam rotated all the dihedral angles with curly arrows, translated the cobalt ion, chloride ions and the macrocycle, and rotated the macrocycle. The bonds shown with thin dotted lines are the metal ligand bonds that were temporarily broken and the thick dotted line is the closure bond.

Table 4
Relative energies of all the configurations within 25 kJ mol^{-1} of the lowest energy conformation of nickel(II) tet-a and tet-b compounds[a]

	tet-a (spl)	tet-a (oct)	tet-b (spl)	tet-b (oct)
trans I	–	–	5.80	–
trans II	11.87	–	0.00	–
trans III	0.00	0.00	15.89	–
trans V	20.59	–	0.00	–
cis V	[b]	14.94	[b]	0.00

[a] tet-a: C-*meso*-5,5,7,12,12,14 hexamethyl-1,4,8,11-tetraazacyclotetradecane; tet-b: C-*rac*-5,5,7,12,12,14 hexamethyl-1,4,8,11-tetraazacyclotetradecane.

[b] Square-planar compounds cannot adopt a *cis* configuration.

default constraints, and all bonds to the metal were broken to generate the translated multidentate ligand, metal, and monodentate ligand fragments. All multidentate ligand fragments were rotated by between 0 and 180° in each MC step (see Fig. 13).

Table 4 gives the relative energies of all the configurations within 25 kJ mol^{-1} of the lowest energy conformation for the square planar and octahedral complexes of nickel tet-a and tet-b. The most stable configuration of square-planar nickel(II) tet-a is the *trans* III configuration and all but one of the square-planar tet-a structures in the CSD are *trans* III structures. The most stable configuration of octahedral tet-a nickel(II) complexes is *trans* III, the only other configuration within 25 kJ mol^{-1} of the lowest energy configuration is a *cis* V form. The CSD has nine *trans* III structures and three *cis* V structures. *Cis* V is the only configuration found, within 25 kJ mol^{-1} of the lowest conformation of octahedral nickel(II) tet-b. The next lowest configuration, *trans* III, is more than 40 kJ mol^{-1} higher in energy. Not surprisingly the only solid state nickel(II) tet-b structures in the CSD have a *cis* V configuration.

6. Conclusion

The CSD contains a wealth of information. It is easy to search and the VISTA module that accompanies it has various statistical functions that make correlation analyses of CSD structures fairly straightforward. Yet the number of CSD analyses remains limited. Possible reasons are the annual cost of the CSD, which has just been decreased, and the memories of the old CSD commands from more than a decade ago when the CSD did not have its user friendly graphic interface.

If the conformations of a sufficiently number congeneric family were to be examined they would approximate the shape of the PES. Therefore the results of a CSD analyses are complementary to the results of inorganic molecular mechanics conformational/configurational searches. Studies combining CSD correlation analyses and inorganic molecular mechanics conformational searches are very rare due to the difficulties in conducting inorganic molecular mechanics conformational searches. Hopefully this will be changing with the introduction of conformational searching methods that can handle transition metal ions in both MacroModel and MOMEC.

References

[1] (a) J.-P. Mathieu, Ann. Phys. 19 (1944) 335. (b) J.C. Bailar, E.J. Corey, J. Am. Chem. Soc. 81 (1959) 2620.
[2] C.R. Landis, D.M. Root, T. Cleveland, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Rev. Comp. Chem, vol. 6, VCH, Weinheim, Germany, 1995.
[3] (a) R.D. Hancock, Prog. Inorg. Chem. 37 (1989) 187. (b) P. Comba, Coord. Chem. Rev. 123 (1993) 1. (c) B.P. Hay, Coord. Chem. Rev. 126 (1993) 177. (d) T.R. Cundari, J. Chem. Soc. Dalton Trans. 1998 (1998) 2771. (e) P. Comba, T.W. Hambley, Molecular Modeling of Inorganic Compounds, VCH, Weinheim, Germany, 1995.
[4] (a) Y. Sun, P.A. Kollman, J. Am. Chem. Soc. 117 (1995) 3599. (b) J. Aqvist, J. Phys. Chem. 94 (1990) 8021.
[5] T.J. Marrone, K.M. Merz, Jr, J. Am. Chem. Soc. 114 (1992) 7542.
[6] P. Comba, M. Zimmer, J. Chem. Ed. 73 (1996) 108.
[7] T.W. Hambley, C.J. Hawkins, J.A. Palmer, M.R. Snow, Aust. J. Chem. 34 (1981) 45.
[8] D.L. Kepert, Inorganic Sterochemistry, Springer, Berlin, 1982.
[9] T.R. Cundari, J. Chem. Soc. Dalton Trans. (1998) 2271.
[10] T.W. Hambley, J. Chem. Soc. Dalton Trans. (1986) 565.
[11] T.W. Hambley, Comm. Inorg. Chem. 14 (1991) 1.
[12] P. Comba, W. Goll, B. Nuber, K. Várnágy, Eur. J. Inorg. Chem. (1998) 2041.
[13] P. Comba, Molecular Modeling and Dynamics of Bioinorganic Compounds, Kluwer, Dorderecht, 1997, p. 21.
[14] V.J. Burton, R.J. Deeth, C.M. Kemp, P.J. Gilbert, J. Am. Chem. Soc. 117 (1995) 8407.
[15] P. Comba, M. Zimmer, Inorg. Chem. 33 (1994) 5368.

[16] P. Comba, T.W. Hambley, M. Ströhle, Helv. Chim. Acta 78 (1995) 2042.
[17] M. Zimmer, Chem. Rev. 95 (1995) 2631.
[18] A.K. Rappe, C.J. Casewit, Molecular Mechanics Across Chemistry, University Science Press, Sausalito, USA, 1997.
[19] Papers appearing in this issue of Coord. Chem. Rev.
[20] I. Kolossvary, W.C. Guida, J. Am. Chem. Soc. 118 (1996) 5011 and references therein.
[21] H. Senderowitz, F. Guarieri, W.C. Still, J. Am. Chem. Soc. 117 (1995) 8211.
[22] J.T. Ngo, M. Karplus, J. Am. Chem. Soc. 119 (1997) 5657.
[23] M.J. Dudek, H.A. Scheraga, J. Comp. Chem. 11 (1990) 121.
[24] E. Leontidis, J.J. de Pablo, M. Laso, U.W. Suter, Adv. Polymer Sci. 116 (1994) 283.
[25] C.W. von der Lieth, T. Kozar, W.E. Hull, Theochem. J. Mol. Struct. 395 (1997) 225.
[26] A.M. Treasurywala, E.P. Jaeger, M.L. Peterson, J. Comput. Chem. 17 (1996) 1171.
[27] J.C. Meza, R.S. Judson, T.R. Faulkner, A.M. Treasurywala, J. Comp. Chem. 17 (1996) 1142.
[28] M. Saunders, K.N. Houk, Y.D. Wu, W.C. Still, M. Lipton, G. Chang, W.C. Guida, J. Am. Chem. Soc. 112 (1990) 1419.
[29] R.M. Hartshorn, D.A. House, J. Chem. Soc. Dalton Trans. (1998) 2577.
[30] J. Bartol, P. Comba, M. Melter, M. Zimmer, J. Comp. Chem. 20 (1999) 1549.
[31] G. Chang, W.C. Guida, W.C. Still, J. Am. Chem. Soc. 111 (1989) 4378.
[32] G. Chang, W.C. Guida, W.C. Still, J. Am. Chem. Soc. 112 (1990) 1419.
[33] M. Saunders, J. Am. Chem. Soc. 109 (1987) 3150.
[34] M. Saunders, J. Comput. Chem. 10 (1989) 203.
[35] K.M. Williams, L. Cerasino, F.P. Initini, G. Natile, L.G. Marzilli, Inorg. Chem. 37 (1998) 5260.
[36] H.-B. Buergi, J.D. Dunitz, Structure Correlation, vol. 1, 2, VCH Press, Weinheim, Germany, 1994.
[37] T.P.E. Auf der Heyde, J. Chem. Ed. 67 (1990) 451.
[38] R. Taylor, F.H. Allen, In: H.-B. Buergi, J.D. Dunitz (Eds.), Structure Correlation, Ch. 4, VCH, Weinheim, Germany, 1994.
[39] P.H.A. Sneath, R.R. Sokal, Numerical Taxonomy, W.H. Freeman and Co, San Francisco, 1973.
[40] G. Bravi, E. Gancia, A. Zaliani, M. Pegna, J. Comp. Chem. 97 (1997) 1295.
[41] R.A. Jarvis, E.A. Patrik, IEEE Trans. Comput. 22 (1973) 1025.
[42] F.H. Allen, M.J. Doyle, R. Taylor, Acta Crystallogr. B47 (1991) 50.
[43] P. Murray-Rust, J. Rafferty, J. Mol. Graph. 3 (1985) 50.
[44] P. Murray-Rust, J. Rafferty, J. Mol. Graph. 3 (1985) 60.
[45] (a) C.R. Landis, V.S. Allured, J. Am. Chem. Soc. 113 (1991) 9493. (b) J.S. Giovannetti, C.M. Kelly, C.R. Landis, J. Am. Chem. Soc. 115 (1993) 4040.
[46] C.P. Casey, S.L. Hallenbeck, J.M. Wright, C.R. Landis, J. Am. Chem. Soc. 119 (1997) 9680.
[47] J.M. Barnard, G.M.J. Downs, Chem. Inf. Comput. Sci. 32 (1992) 644.
[48] C.H. Reynolds, R. Drucker, L.B. Pfahler, J. Chem. Comput. Sci. 38 (1998) 305.
[49] T. Okada, T. Wipke, Tetrahedron Comp. Meth. 2 (1989) 249.
[50] B.R. Kowalski, S. Wold, in: P.R. Krishnaiah, L.N. Kanal (Eds.), Handbook of Statistics, vol. 2, North-Holland, Amsterdam, 1982.
[51] M.E. Karpen, D.J. Tobias, C.L. Brooks, III, Biochemistry 32 (1993) 412.
[52] P.S. Shenkin, D.Q. McDonald, J. Comp. Chem. 15 (1994) 899.
[53] A.E. Torda, W.F. Van Gunsteren, J. Comp. Chem. 12 (1994) 1331.
[54] D. Conklin, S. Fortier, J.I. Glasgow, F.H. Allen, Acta Crystallogr. B52 (1996) 535.
[55] C. Chatfield, A.J. Collins, Introduction to Multivariate Analysis, Chapman and Hall, London, 1980.
[56] F.H. Allen, J.E. Davies, J.J. Galloy, O. Johnson, O. Kennard, C.F. Macrae, E.M. Mitchell, G.F. Mitchell, J.M. Smith, D.G. Watson, J. Chem. Inf. Comp. Sci. 31 (1991) 187.
[57] F.H. Allen, O. Kennard, Chem. Des. Automation News 1 (1993) 31.
[58] F.C. Bernstein, T.F. Koetzle, J.G.B. Williams, E.F. Meyer, M.R. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, M. Tasumi, J. Mol. Biol. 122 (1977) 535.
[59] A.G. Orpen, J. Chem. Soc. Dalton Trans. (1993) 191.
[60] J.D. Dunitz, H.-B. Buergi, Acta Crystallogr. B44 (1988) 445.

[61] T. Auf der Heyde, In: H.-B. Buergi, J.D. Dunitz (Eds.), Structure Correlation, Ch. 4, VCH, Weinheim, Germany, 1994.
[62] S. Beyreuther, J. Hunger, G. Huttner, S. Mann, L. Zsolnai, Chem. Ber. 129 (1996) 645.
[63] K.R. Adam, M. Antolovich, L.G. Brigden, L.F. Lindoy, J. Am. Chem. Soc. 113 (1991) 3346.
[64] R.D. Hancock, S.M. Dobson, A. Evers, P.W. Wade, M.P. Ngwenya, J.C.A. Boeyens, K.P. Wainwright, J. Am. Chem. Soc. 110 (1989) 2788.
[65] J.J. Gajewski, K.E. Gilbert, T.W. Kreek, J. Comput. Chem. 19 (1998) 1167.
[66] E. Yuriev, J.D. Orbell, Inorg. Chem. 37 (1998) 6269.
[67] R. Cini, S.J. Moore, L.G. Marzilli, Inorg. Chem. 37 (1998) 6890.
[68] P.R. Raithby, G.P. Shields, F.H. Allen, Acta Crsytallogr. B53 (1997) 241.
[69] P.R. Raithby, G.P. Shields, F.H. Allen, Acta Crsytallogr. B53 (1997) 476.
[70] F.H. Allen, P.R. Raithby, G.P. Shields, Struct. Chem. 8 (1997) 385.
[71] F.H. Allen, M.J. Doyle, R. Taylor, Acta Crystallogr. B47 (1991) 41.
[72] J. Dale, J. Chem. Soc. (1963) 93.
[73] J. Dale, Acta Chem. Scand. 27 (1973) 1115.
[74] B.P. Hay, D. Zhang, J.R. Rustad, Inorg. Chem. 35 (1996) 2650.
[75] M.F. DaCruz, M. Zimmer, Inorg. Chem. 35 (1996) 2872.
[76] M.F. DaCruz, M. Zimmer, Inorg. Chim. Acta 261 (1997) 181.
[77] K.M. Norenberg, C.M. Shoemaker, M. Zimmer, J. Chem. Soc. Dalton Trans. 1997 (1997) 1521.
[78] M.F. DaCruz, M. Zimmer, Inorg. Chem. 37 (1998) 366.
[79] (a) R.D. Hancock, Prog. Inorg. Chem. 37 (1989) 187. (b) V.J. Thoem, C.C. Fox, J.C.A. Boeyens, R.D. Hancock, J. Am. Chem. Soc. 106 (1984) 5947.
[80] (a) K.B. Wiberg, R.H. Boyd, J. Am. Chem. Soc. 94 (1972) 8426. (b) M.G.B. Drew, P.C. Yates, J. Chem. Soc. Dalton Trans. (1986) 2506.
[81] A.M. Sargeson, Pure Appl. Chem. 58 (1986) 1511.
[82] P.V. Bernhardt, R. Bramley, L.M. Engelhardt, J.M. Harrowfield, D.C.R. Hockless, R. Korybut-Daszkiewicz, E.R. Krausz, T. Morgan, A.M. Sargeson, B.W. Skelton, A.H. White, Inorg. Chem. 34 (1995) 3589.
[83] J.M. Castagnetto, S.W. Hennessy, M. Pique, E.D. Getzoff, J.A. Tainer, J. Inorg. Biochem. 74 (1999) 90.
[84] J.A. Shelnutt, X.-Z. Song, J.-G. Ma, S.-L. Jia, W. Jentzen, C.J. Medforth, Chem. Soc. Rev. 27 (1998) 31.
[85] J.A. Shelnutt, in: K.M. Kadish, K.M. Smith, R. Guilard (Eds.), The Porphyrin Handbook, vol. 7, Academic Press, New York, 2000.
[86] W. Jentzen, X.-Z. Song, J.A. Shelnutt, J. Phys. Chem. B 101 (1997) 1684.
[87] J.-G. Ma, J. Zhang, R. Franco, S.-L. Jia, I. Moura, J.J.G. Moura, P.M.H. Kroneck, J.A. Shelnutt, Biochemistry 37 (1998) 12431.
[88] A.R. Leach, J. Comput. Aided Mol. Des. 7 (1993) 225.
[89] A.R. Leach, D.P. Dolata, K. Prout, J. Chem. Inf. Comput. Sci. 30 (1990) 316.
[90] A.R. Leach, K. Prout, J. Comput. Chem. 11 (1990) 1193.
[91] D.A. Buckingham, M. Dwyer, G.J. Gainsford, V.J. Ho, L.G. Marzille, W.T. Robinson, A.M. Sargeson, K.R. Turnbull, Inorg. Chem. 14 (1975) 1739.
[92] B.F. Anderson, J.D. Bell, D.A. Buckingham, P.J. Creswell, G.J. Gainsford, L.G. Marzilli, G.B. Robertson, A.M. Sargeson, Inorg. Chem. 16 (1977) 3233.
[93] F. Mohamadi, N.G.J. Richards, W.C. Guida, C. Liskamp, C. Caufield, G. Chang, T. Hendrickson, W.C. Still, J. Comp. Chem. 11 (1990) 440.
[94] P. Comba, T.W. Hambley, G. Lauer, N. Okon, Lauer & Okon Chemische Verfahrens- & Softwareentwicklung, (1997). For information contact: Lauer&Okon, Heidelberg, Germany, e-mail: CVS-HD@T-Online.de
[95] J.E. Bol, C. Buning, P. Comba, J. Reedijk, M. Ströhle, J. Comput. Chem. 19 (1998) 5.
[96] K.B. Lipkowitz, K. Green, J.-A. Yang, Chirality 4 (1992) 205.
[97] K.B. Lipkowitz, Chem. Rev. 98 (1998) 1829.
[98] L.F. Lindoy, The Chemistry of Macrocyclic Ligand Complexes, Cambridge University Press, Cambridge, 1989.

[99] A.E. Martell, R.D. Hancock, in: J.P. Fackler (Ed.), Metal Complexes in Aequeous Solutions, Mod. Inorg. Chem, Plenum, New York, 1996.

[100] D.K. Cabbiness, D.W. Margerum, J. Am. Chem. Soc. 91 (1969) 6540.

[101] R.M. Izatt, K. Pawlak, J.S. Bradshaw, R.L. Bruening, Chem. Rev. 91 (1991) 1721.

[102] M. Ciampolini, N. Nardi, B. Vantancoli, M. Micheloni, Coord. Chem. Rev. 120 (1992) 223.

[103] L. Lindoy, Pure Appl. Chem. 69 (1997) 2179.

[104] B. Bosnich, C.K. Poon, M.L. Tobe, Inorg. Chem. 4 (1965) 1102.

[105] L.G. Warner, D.H. Busch, J. Am. Chem. Soc. 91 (1969) 4092.

[106] V.J. Thom, C.C. Fox, J.C.A. Boeyens, R.D. Hancock, J. Am. Chem. Soc. 106 (1984) 5947.

[107] P.J. Connolly, E.J. Billo, Inorg. Chem. 26 (1987) 3224.

[108] K.R. Adam, M. Antolovich, I.M. Atkinson, L.G. Brigden, L.F. Lindoy, J. Mol. Struct. 323 (1994) 223.

[109] T.W. Hambley, J. Chem. Soc. Dalton Trans. (1986) 565.

[110] K.R. Adam, I.M. Atkinson, L.F. Lindoy, J. Mol. Struct. 384 (1996) 183.

[111] K.R. Adam, I.M. Atkinson, L.F. Lindoy, Inorg. Chem. 36 (1997) 480.

[112] C. Cooper, M. Zimmer, Struct. Chem. 10 (1999) 17.

[113] M. Bakaj, M. Zimmer, J. Mol. Struct. 508 (1999) 59.

[114] M.A. Donnelly, M. Zimmer, Inorg. Chem. 38 (1999) 1650.

[115] L. Yujuan, N. Shisheng, X. Jide, J. Huaxue, J. Struct. Chem. 5 (1986) 5.

[116] L. Yu-Juan, C. Li-Qin, N. Shi-Sheng, X. Ji-De, J. Huaxue, J. Struct. Chem. 14 (1995) 44.

[117] C. Benelli, A. Dei, D. Gatteschi, L. Pardi, Inorg. Chem. 27 (1988) 2831.

[118] X.H. Bu, X.C. Cao, Z.H. Zhang, Z.A. Zhu, Y.T. Chen, M. Shionoya, E. Kimura, Polyhedron 15 (1996) 1203.

[119] J.K. Moran, M.M. Olmstead, C.F. Meares, Acta Crystallogr. Sect. C (Cr. Str. Comm.) 51 (1995) 621.

[120] T.J. Collins, T.R. Nichols, E.S. Uffelman, J. Am. Chem. Soc. 113 (1991) 4708.

[121] D.L. Cullen, J.D. Urgitis, M. Zimmer, Unpublished results.

[122] K. Kobiro, A. Nakayama, T. Hiro, M. Suwa, K. Kakiuchi, Y. Tobe, Y. Odaira, Chem. Lett. (1990) 1979.