Conformation of *tripod* Metal Templates in $CH_3C(CH_2PPh_2)_3ML_n$ (n = 2, 3): Neural Networks in Conformational Analysis☆

Stefan Beyreuther, Johannes Hunger, Gottfried Huttner*, Susanne Mann, and Laszlo Zsolnai

Anorganisch-Chemisches Institut der Universität Heidelberg,
INF 270, D-69120 Heidelberg, Germany
Telephone: (internat.) +49(0)6221-548446
Telefax: (internat.) +49(0)6221-545707
E-mail: stefan@indi.aci.uni-heidelberg.de

Received December 27, 1995

Key Words: Conformational analysis / Triphos ligands / Neural networks / Factor analysis

The conformational space spanned by *tripod* metal templates $CH_3C(CH_2Ph_2)_3M$ is analysed on the basis of the solid-state structures of 82 *tripod*Co templates in compounds *tripod*CoL_2 and *tripod*CoL_3. Systematic analysis, including the techniques of conformational space group scatter graphs, principal-component analysis, and partial least squares, reveals a series of basic regularities: The torsion of the phenyl groups is strongly linked to the torsional skew of the bicyclooctane-type framework of the chelate cage. For one sense of this skew there are two classes of low-energy conformation that differ by the helicity of the phenyl arrangement and by the degree of torsional skew in the chelate backbone. From the scatter graphs it is evident that a change in helicity may occur by one- or by two-ring flip mechanisms. The basic regularities found by the above methods are also evident from the analysis of the same data by a neural network approach. The fact that these regularities are found for *tripod*CoL_2 and *tripod*CoL_3, irrespective of the widely different coligands L and crystal environments, means that the conformation of the *tripod* metal templates is governed by the forces acting within them and not so much by the forces imposed on them by their individual chemical or crystal environment. It is shown that the classifications, although derived from a data basis only containing Co compounds, are characteristic for *tripod* metal templates irrespective of the specific metal involved.

In most cases specificity in coordination chemistry is dominated by a specific choice of the ligands surrounding the metal. Homogenous catalysis[1,2] and, even more so, the subset of enantioselective catalysis[3] are clear illustrations of this statement.

Part of this specificity will arise from steric interactions. In order to be able to predict specificity a thorough knowledge of these steric interactions is therefore needed. One basis for acquiring such a knowledge is currently available in many cases: A considerable amount of information exists on solid-state structures (determined by X-ray analysis), most of it deposited in an orderly and easily retrievable way in internationally available databases[4]. The wealth of information contained in these data files is, as yet, far from having been systematically exploited.

The current paper deals with the systematic analysis of all such structural data describing the solid-state conformation of *tripod*Co templates [*tripod* = $CH_3C(CH_2PPh_2)_3$] for compounds *tripod*CoL_2 and *tripod*CoL_3. The interest in this analysis stems from a research programme aiming at an understanding of the chemical reactivity of a subset of ligand metal templates in which the ligand is a neopentane-based tripodal entity. The programme involves the three following steps: 1) synthesis of a library of neopentane-based *tripod* ligands $RCH_2C(CH_2X)(CH_2Y)(CH_2Z)$ (X, Y, Z = donor groups), 2) understanding and modelling the shape of *tripod* metal templates *tripod*M, 3) correlating the results of a catalysis mediated by *tripod*M with the shape of *tripod*M. The synthesis of *tripod* ligands $RCH_2C(CH_2X)(CH_2Y)(CH_2Z)$ with three different donor groups X, Y, Z is well established including the enantioselective synthesis of ligands containing three different phosphorus donor groups[5–25]. Catalysis mediated by *tripod* metal templates[26–32] has been observed, and it is thus time to approach the understanding of the shape of *tripod* metal templates.

An experimental basis for the development of such an understanding may be found in the X-ray data pertaining to a specific *tripod* ligand $CH_3C(CH_2PPh_2)_3$. Since most of the data referring to this ligand are available for *tripod*Co templates the systematic analysis has been restricted to cover *tripod*CoL_2 and *tripod*CoL_3.

The traditional tools for this type of analysis, such as conformational space-group scatter graphs[33–37], principal-component analysis (PCA)[33–40], partial least squares (PLS)[38–40] and hierarchical clustering[39,40], have been applied to the data. They allow for the classification of conformations and the elucidation of pathways in conformational space. This type of information in itself forms a basis on which the construction of force field models may be built.

As a method that is not yet traditional in conformational analysis the methodology of neutral networks[41–44] has also been applied to the data. It is shown that this technique reproduces the results from the traditional methods but also

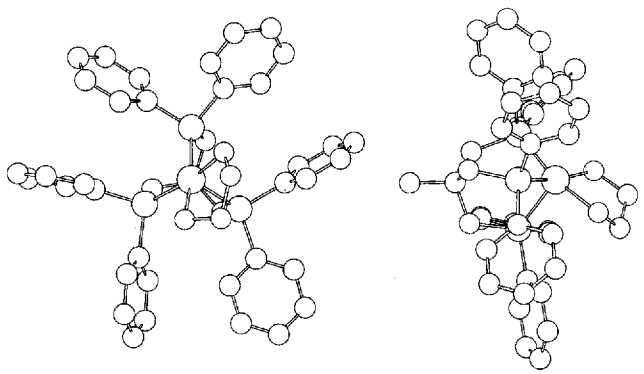
  0009–2940/96/0707–0745 $ 15.00+.25/0

has merits of its own that recommend its further application in conformational analysis.

Definition of Parameters

In order to systematically analyse the conformational space occupied by *tripod* metal templates, this space must be defined by an appropriate selection of conformational parameters. To this end the conformational manifold shown by *tripod* metal coordination compounds is defined by a small set of torsion, angle, and distance parameters. The relation between these parameters and the usual structure plots is visualised by comparison of Figures 1 and 2.

Figure 1. Ball and stick model of $CH_3C(CH_2PPh_2)_3(en)Co(II)$ (NABGUS, Table 1) as derived by X-ray analysis in two mutually orthogonal viewing directions

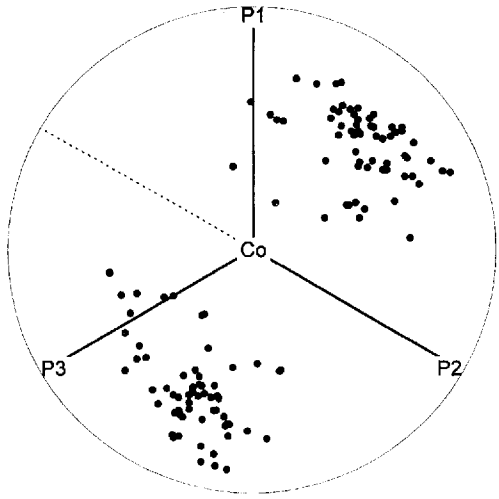


The rotational position of the phenyl groups is defined relative to the idealised C_3 axis of the chelate cage. This axis is vertical to the plane defined by the three phosphorus atoms and runs through the metal centre. The specific phenyl rotations for which this axis is parallel to the phenyl planes are assigned a φ value of 0° (Figure 2)[8,45]. The torsional skew of the chelate cage is measured by torsion angles τ ($C_{bridgehead}-C_{CH_2}-P-M$) (Figure 2).

The position of the donor atom of a specific coligand L is defined by the metal ligand distance, by a cone angle α as measured with respect to the idealised threefold axis and a rotational angle χ (Figure 2). In order to uniquely define these rotational angles a unique numbering scheme is essential. The following scheme has been adopted throughout: The vector M−P2 must lie in a plane that best approximates an idealised C_s symmetry of the ML_2 or ML_3 fragments.

A scatter graph (Figure 3) illustrating the idealised C_s symmetry of the five-coordinate compounds is obtained by projecting the positions of the donor atoms of the two ligands L1 and L2 onto the plane defined by P1, P2, and P3.

Figure 3. Projection of the positions of the donor atoms of the coligands L1 and L2 onto the plane defined by the three phosphorus atoms for the 62 five-coordinate compounds *tripod*CoL_2 contained in the database (Table 1). The numbering scheme is such that M−P2 lies in the idealised mirror plane of the ML_2 fragments (bisector of the mirrorplane indicated by dashed line)



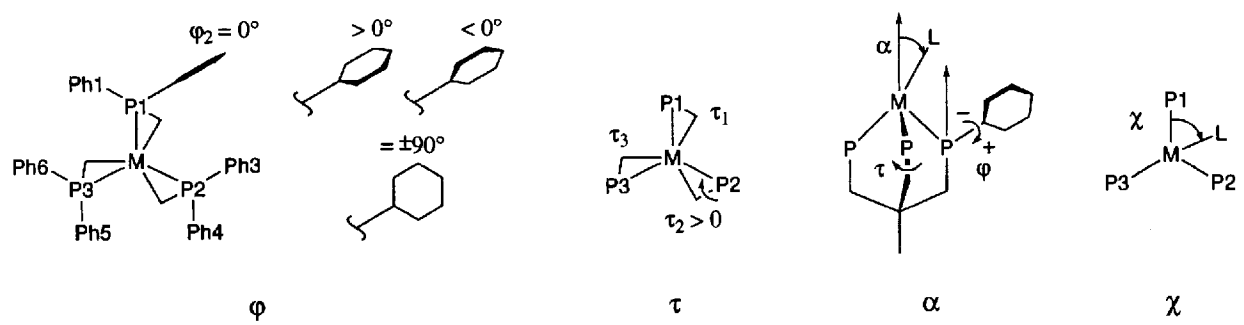
The Data Set

In this way the conformations of 82 compounds of the type *tripod*CoL_2 and *tripod*CoL_3, which form the basis of the following analysis, have been coded.

With respect to the helicity of the chelate cage (τ, Figure 2) there are two choices for standardising the data set corresponding to the two enantiomeric conformations that this cage may adopt. The choice corresponding to positive τ values as defined in Figure 2 has been set as the standard. Whenever necessary this standard has been met by inversion of the published crystallographic positional parameters.

The constitutional C_3 symmetry of the P_3 metal entity allows for three energetically equivalent arrangements of the coligands relative to the P_3M framework. Taking into account these three coligand positions only one half of conformational space is described: Ani conformation described

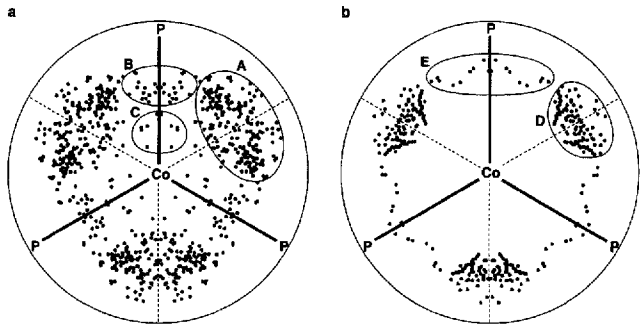
Figure 2. Definition of parameters



in this way has its mirror-symmetric equivalent of equal energy. Thus, the manifold describing all available isoenergetic conformations of the P_3ML1L2 or $P_3ML1L2L3$ entities in conformational space has C_{3v} symmetry.

The corresponding symmetric expansion of the original data leads to the scatter plot shown in Figure 4a (*tripod*CoL_2) and 4b (*tripod*CoL_3).

Figure 4. Scatter graphs. Projection of the positions of the donor atoms for (a) *tripod*ML_2 and (b) *tripod*ML_3 onto a plane defined by the three phosphorus atoms (symmetry-expanded dataset)

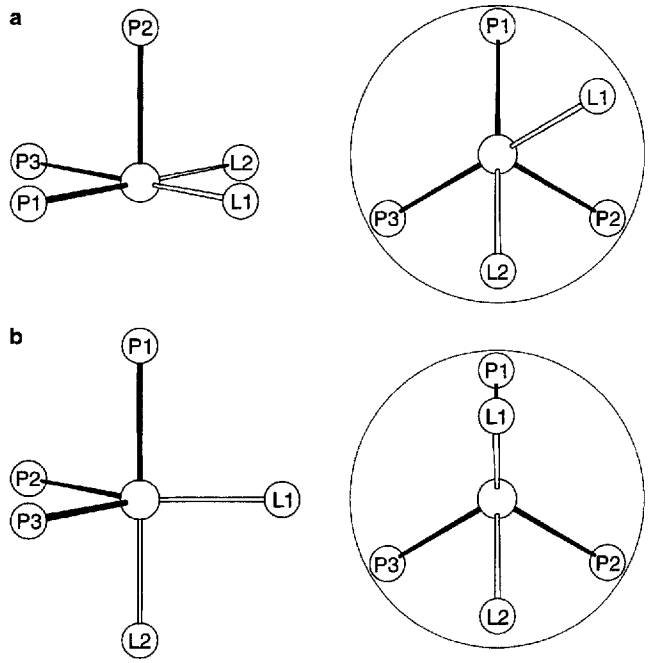


The Positions of the Coligands L

The data for the five-coordinate species *tripod*CoL1L2, already visualised in Figure 3, lead to the scatter plot in Figure 4a when expanded according to C_{3v} symmetry. It is evident that positions close to eclipsed arrangements of the M–P and M–L bonds (regions **B**, **C**) are far less populated compared with positions corresponding to staggered arrangements (region **A**, Figure 4a). At the same time, those positions corresponding to almost exactly staggered conformations are less densely populated (lines bisecting these segments) than region **A** on either side of the bisector. A further characteristic feature of the graph (Figure 4a) is the broad dispersion of data points. The data points in region **C** and its symmetric equivalents are close to the centre of the graph. All compounds represented by those data points contain ligands for which the donor atoms are directly connected (TRPCOS10, TPHCOA, KIRNII, SUHMIR, NABGOM; Table 1). It is a matter of taste to classify the compounds containing these ligands in an η^2-coordination mode as four-coordinate or five-coordinate. If the compounds are treated as four-coordinate then in an idealised tetrahedral arrangement the midpoint of the bond between the ligand atoms would lie on the trigonal axes and the endpoints would just be separated by one covalent bond. The projections (as shown in Figure 4a) must, therefore, have to be close to the centre (region **C**). In the view taken here the compounds are nevertheless considered as five-coordinate, with the positions of the donor atoms of the coligands necessarily close to the centre.

The scatter of the more densely populated outer part of the graph is a representation of the notorious flexibility of fivefold coordination[46–48] generally described with trigonal-bipyramidal and square-pyramidal coordination as the cornerstones (Figure 5). While a trigonal-bipyramidal arrangement with one of the P–M–P angles close to 120° is impossible due to the rigidity of the neopentane framework of the *tripod* ligands considered here, arrangements approximating this idealisation are feasible (Figure 5b): two phosphorus atoms and one of the coligand donor atoms may lie in one plane with the metal atom (P2, P3, L1). The two remaining bonds, one to the apical phosphorus (P1) and the other to the donor atom of the remaining ligand (L2), are approximately vertical to this plane.

Figure 5. Idealised coordination polyhedra for *tripod*ML_2: (a) square-pyramidal, (b) trigonal-bipyramidal. The encircled figures correspond to projections onto the P_3 planes



A square-pyramidal conformation, with one phosphorus in the apical position and the two other phosphorus atoms and the donor atoms L1 and L2 forming the basal square plane, is obviously less restricted by this angular constraint (P–M–P = 90°, Figure 5a).

The two idealised forms find distinct representations in the scatter graph (Figure 4a). For the idealised trigonal-bipyramidal conformations one of the data points representing the position of the two coligands for a specific compound must project into region **B** (Figures 4a and 5b) and its symmetry equivalents close to the line representing an M–P bond (L1, Figure 5b), the other ligand is necessarily close to the bisector on the opposite side (Figure 5b), thus falling into the symmetry equivalent of region **A** (Figure 4a).

For the square-pyramidal idealisation both of the donor atoms of the two coligands, L1 and L2, should project onto lines bisecting the segments PMP of the graph (Figure 5a). Compounds whose geometry is close to a square-pyramidal arrangement will therefore be represented by data points in region **A** and its symmetry equivalents (Figure 4a).

The pattern observed for the five-coordinate compounds *tripod*ML_2 is complicated by the fact that a threefold (MP_3) and a twofold (ML_2) rotor are combined in one and the

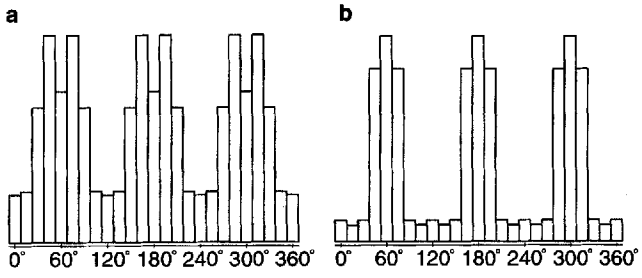
same compound. With six-coordinate *tripod*ML_3 the corresponding pattern should be simplified by the C_3 symmetry of both the MP_3 and the ML_3 rotors. The scatter graph (Figure 4b) based on a C_{3v} expansion of the original data shows a comparably simple structure with densely populated regions **D** corresponding to arrangements close to an idealised octahedral coordination[49]. The circular belt connecting these regions is only sparsely populated (Figure 4b). In fact, all of the 18 points contained in region **E** (Figure 4b) refer to just two compounds. Both are tripledecker-type species *tripod*Co(μ^3-X_3)ML_n (ASDPCO10, COPFEA; Table 1). That giving rise to 12 of the 18 points in region **E** contains a triarsenic entity ($X_3 = As_3$, ASDPCO10) embedded between two *tripod*Co fragments. The other, represented by the remaining six data points, contains a triphosphorus cycle ($X_3 = P_3$, COPFEA) coordinated on one side to a *tripod*Co species and on the other to a $CH_3C(CH_2PEt_2)_3Fe$ entity. Only the part containing Co is referred to in the data.

The strong deviation from octahedral coordination observed for these tripledecker compounds may be attributed in part to the fact that a strictly octahedral coordination would lead to an arrangement in which the MP_3 fragments of the two *tripod* metal templates would be strictly eclipsed, thereby leading to steric crowding. On the other hand, the electronic properties of the P_3 system taken as a cyclic π ligand would not differentiate very much between the different rotational orientations relative to the *tripod*M entities[50]. Nevertheless, compounds containing X_3 ligands in terminal μ^3-coordination fall into the octahedral regions **D**. At the border of region **D** a few compounds are represented that contain terminal μ^3-"X_3" groups ("X_3" = P_3, P_2S, P_2Se, As_2S) where, however, one (or two) atoms of the "X_3" species is (are) connected to bulky organic or inorganic groups (CEYMOU, DEWGAN, DUVXAT, COJHOY20, CEWGIU; Table 1).

It is informative to compare the scatter plots of five- and six-coordinate compounds (Figure 4): Going from sixfold octahedral coordination to fivefold square-pyramidal coordination in a least-motion sense corresponds to removing just one ligand from the octahedron. The two remaining ligands at the base of the square pyramid naturally find their projections in the regions **A** (Figures 4a, 5a). Region **D** (Figure 4b) is thus naturally mapped onto region **A** (Figure 4a). The fact that the density of points is close to its maximum at the lines bisecting the triangular sectors in Figure 4b for the octahedral compounds, while the bisecting lines mark positions of local minimum density, is already qualitatively apparent in Figure 4a.

This view is demonstrated in a more quantitative way by the histograms in Figure 6, in which the data points of Figure 4 are sampled with respect to the angular distribution. The histogram corresponding to sixfold coordination (Figure 4b) shows a narrow distribution with maxima at the bisecting angles (Figure 6b)[51]. The histogram referring to fivefold coordination (Figure 6a) reveals a much broader

Figure 6. Histograms of the angular distribution of coligand positions for (a) *tripod*ML_2 and (b) *tripod*ML_3. The horizontal scale refers to χ values as defined in Figure 2



distribution indicating a clear local minimum at the bisecting angles. These facts lend themselves to a straightforward interpretation: The additional free space created by removing one ligand from the octahedral arrangement allows for a relaxation such that the coligands make better use of the total coordination space available (Figure 3).

Conformational Analysis

The matter of real interest in this analysis is the search for regularities within the conformational space occupied by the *tripod* metal templates $CH_3C(CH_2PPh_2)_3M$ themselves. While it is to be expected that the various arrangements of the coligands will have an influence on the conformation of the *tripod* ligands it is also expected that this influence will not be so dominant as to override these regularities. This reasoning is intrinsically based on the amount of steric crowding within the *tripod* metal templates. The following analyses demonstrate that this reasoning is based on firm grounds.

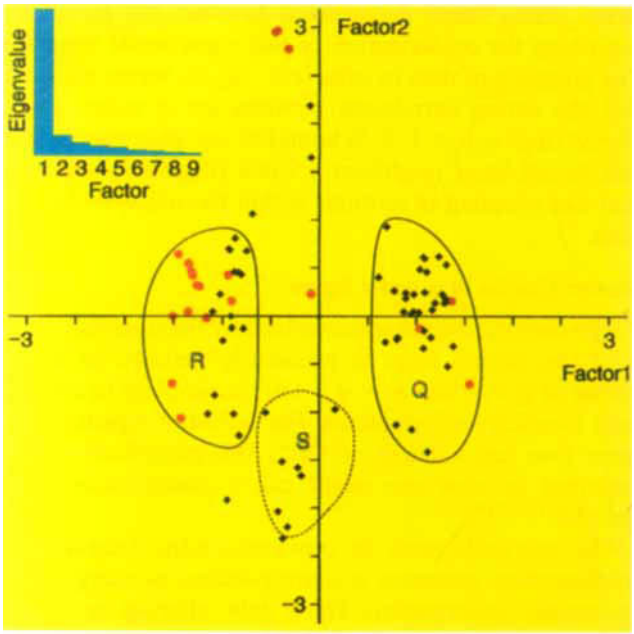
Principal-Component Analysis

Principal-component analysis (PCA) has become a standard tool for any attempt to extract regularities from a set of structures containing some constituents in common[33–40]. The essence of the method lies in the fact that it creates linear combinations (called principal components) of the parameters used to describe the data manifold in a ranked sequence in such a way that these principal components describe a maximum of the remaining variance at each step[38,39]. Applied to conformational analysis the method often allows reduction of the dimensionality of the conformational problem to just a few principal components. These components may, in fortunate cases, lend themselves to simple physical interpretations that provide a clue to the main conformational changes. In addition, classification of the data manifold into subsets may sometimes be indicated by PCA.

When applied to the φ and τ parameters used to describe the conformations of the 82 structurally characterised compounds containing *tripod*Co templates (Table 1) a few principal components clearly drop out from the 9-dimensional φ/τ space.

The first component already describes 69.3% of the total variance. Including the second component corresponds to a coverage of 78.7% of the total variance (Figure 7). When

Figure 7. Projection of the data describing 82 structurally characterised *tripod*Co templates on the two most important principal components obtained from a PCA using six φ values and three τ values for each compound (six-coordinate: dots; five-coordinate: squares). A histogram of the Eigenvalues is shown in the insert



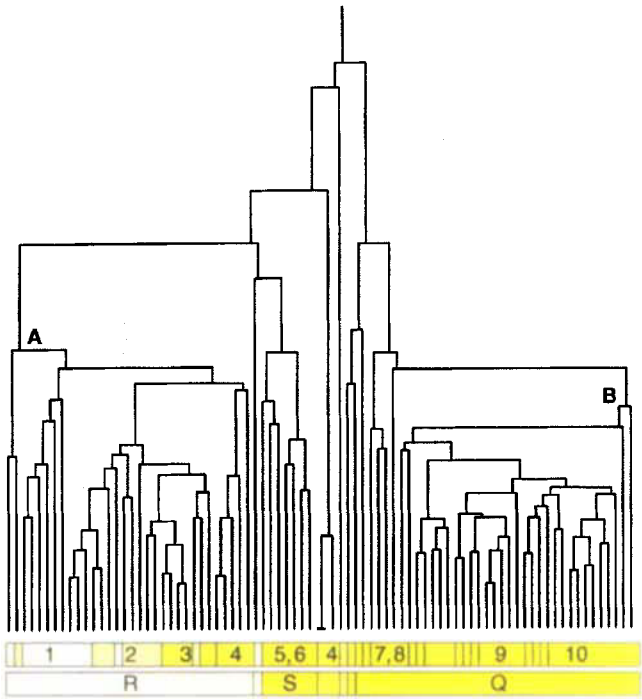
the data are projected onto these two most important principal components the corresponding graph (Figure 7) shows a clear separation of the data into subsets **Q**, **R** and **S**. Subsets **Q** and **R** are clearly separated by component 1 alone, subset **Q** having positive and subset **R** having negative coordinates (scores) along this component. Subset **S** is separated from the other two subsets by each of the two components; with respect to component 1 it projects close to a value of zero, thus being in the middle between **Q** and **R** (Figure 7).

Almost all of the data points projecting into region **Q** refer to five-coordinate compounds (squares). Subset **R** is subdivided into a left-hand part containing mainly data points referring to six-coordinate compounds (dots) and a right-hand part dominated by data points referring to five-coordinate species (squares, Figure 7). The overall conformational properties giving rise to the separation into classes **Q** and **R** are evident when traced back to the corresponding data for φ and τ. Compounds in subset **Q** have small positive values of τ and positive torsions φ (mean values: $\varphi_1...\varphi_6$: 17, 44, 21, 31, 26, 29; $\tau_1...\tau_3$: 15, 14, 13). Compounds projecting into region **R** have large positive τ and negative φ (mean values: $\varphi_1...\varphi_6$: −17, −17, −37, −21, −35, −12; $\tau_1...\tau_3$: 31, 31, 27). This observation agrees well with the loadings of component 1, which are all of approximately equal absolute size, being positive for all of the six φ parameters and negative for all of the three τ parameters (numbers refer to the parameters $\varphi_1...\varphi_6$, $\tau_1...\tau_3$ in that order; component 1: 0.875, 0.778, 0.876, 0.753, 0.848, 0.752, −0.886, −0.877, −0.830). Principal component 1 hence describes a strong correlation between φ and τ values in such a way that whenever φ increases τ must decrease and vice versa. Principal component 2, describing only an additional 9% of the total variance, in comparison to the 69% covered by component 1, has no straightforward mechanistic interpretation; it helps, however, to separate subset **S** from the rest of the data (Figure 7). The data points within this subset may tentatively be assigned to conformations representing a transition between subsets **Q** and **R**. This assignment is corroborated by other types of analyses (see below). To sum up, PCA indicates a strong negative correlation between the torsional motion of the phenyl groups and the torsional twist of the chelate cage and, at the same time, groups the data into various well defined subsets.

Cluster Analysis

A different approach to analyse the data manifold with respect to the occurrence of natural subsets is cluster analysis. This method starts by analysing some measure of distance between the individual data points. Data points closest to each other are combined into pairs, which are then represented by their average for the next iteration step. The process is repeated with these new data points until finally only one data point remains. Once this analysis is done a tree-like graph may be constructed that groups the data into clusters, such that points from the original data set that are similar to each other are grouped in the same cluster[39,40].

Figure 8. Dendrogram as obtained by cluster analysis for the conformations of 82 *tripod* metal templates each encoded by six φ and three τ values (Figure 2). The blocks given under the dendrogram refer to the neural network analysis (top box, neurons 1–10; Figure 11) and principal component analysis (regions R, S, Q; Figure 7)[52]



Applying cluster analysis to the 82 data points representing the conformations of the 82 structurally characterised *tripod*Co templates should lead to a dendrogram, which should group compounds with similar overall conformations close to each other[52]. The dendrogram actually

obtained (Figure 8) shows groups of closely related conformations at its left-hand and at its right-hand side. The relatively close relation between the conformations represented within each group is evident from the vertical distance to the branching points **A** and **B** from which these clusters are derived. The range in the middle of the dendrogram corresponds, as judged from the branching points of their groups, to much more diverse conformations. The overall grouping of the data as inferred from cluster analysis (Figure 8) qualitatively corresponds to the classification inferred from PCA: Data belonging to subsets **Q**, **R**, and **S** from the PCA classification are, with a few exceptions, grouped together in blocks by cluster analysis (Figures 7, 8). Subset **S** corresponding to a subgroup mainly discriminated by principal component 2 of PCA forms a group again well separated from the other two clusters.

Partial Least Squares

The partial least squares (PLS) technique has been established as a method to quantify the degree of linear correlation between two data matrices[38–40]. In its current application to the above data set PLS has been applied to the data matrix of the six φ columns in such a way that this matrix was split into a matrix consisting of just one column φ_i and a matrix containing the other five columns φ_j ($j \neq i$). Applied in this way PLS corresponds to a kind of multiple linear regression analysis (MLRA) with the additional merit that the optimum number of components for the description of φ_i in terms of linear combinations of φ_j ($i \neq j$) is obtained. It is observed that in order to obtain a correlation coefficient of 62–77% two components are required for φ_i, $i = 1, 3, 5$, while three components are appropriate for φ_i, $i = 2, 4, 6$. This differentiation between the two sets of φ values φ_{2n-1} and φ_{2n} ($n = 1, 2, 3$) corresponds to the different types of positions that the corresponding phenyl rings occupy, corresponding to the C_3 symmetry of the geometric arrangement (Figures 1 and 2). Phenyl groups 2, **4**, and 6 are closer to the coordination space occupied by the coligands compared with the phenyl groups 1, 3, and 5, due to the positive τ angles [as defined for the standard setting (Figure 2)].

By geometrical reasoning it might be expected that the torsional position occupied by a specific phenyl ring φ_i is strongly influenced by the torsional positions of its two nearest neighbours, defined by φ_j ($j = |i + 1|_{\text{mod } 6}$) and φ_k ($j = |i - 1|_{\text{mod } 6}$). Linear regressions based on this model show a correlation coefficient of 59–72% for φ_i ($i = 1, 3, 5$); for φ_i ($i = 2, 4, 6$), a correlation between 60 and 65% is obtained.

If the individual φ columns are correlated with all the remaining eight columns containing five φ and three τ parameters the observed linear correlations are necessarily higher (ranging between 63 and 84%) as compared to those obtained with φ columns only (see above).

Searching for two-parameter models to describe φ_i in terms of one φ and one τ parameter it is found that correlations are better than 67% if φ_1 is expressed by the torsion φ_2 of the phenyl group bound to the same phosphorus atom and the torsion τ_3 of its left-hand neighbour (72%) and, correspondingly, φ_3 in terms of φ_4 and τ_1 (79%) and φ_5 in terms of φ_6 and τ_2 (67%).

Overall, the results of the PLS analyses again demonstrate strong linear correlations between the parameters describing the conformation of the *tripod* metal templates. The grouping of data in subsets φ_1, φ_3, φ_5 versus φ_2, φ_4, φ_6 and the strong correlation between the φ values of the phenyl rings φ_i ($i = 1, 3, 5$) bound to one phosphorus atom and its left-hand neighbour τ value (Figure 2) indicate a gear-like coupling of motions within the *tripod* metal template.

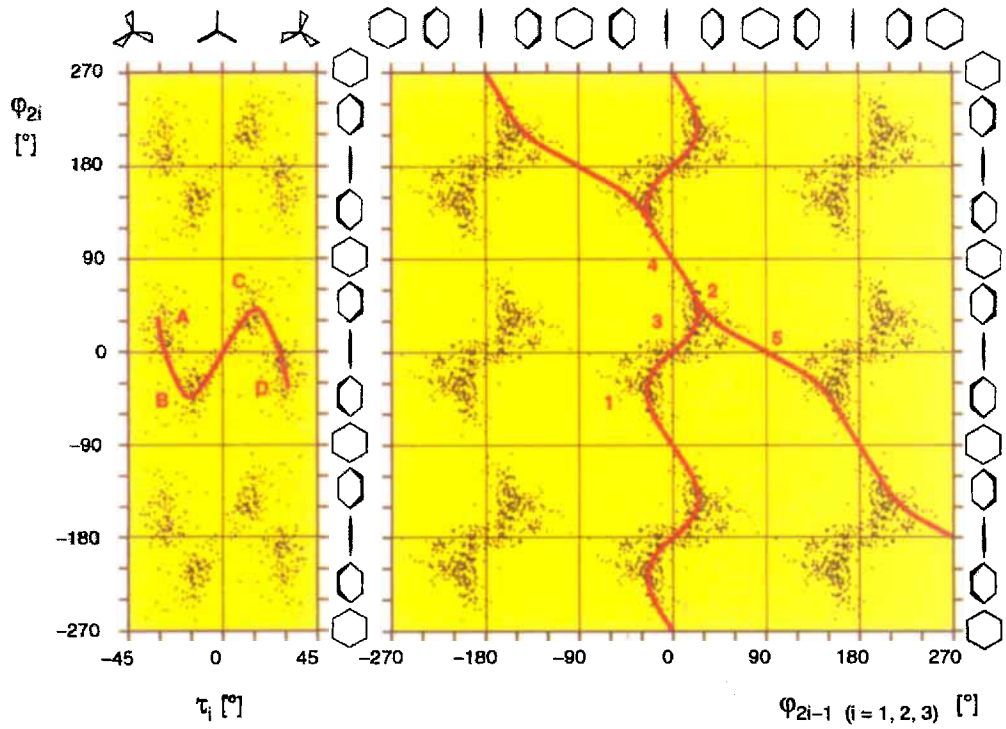
Scatter Graphs in φ and τ Space

The conformational space defined by the torsional angles φ of the phenyl rings is necessarily periodic in nature. Values of $\varphi_i = \varphi$ and $\varphi_i = \varphi \pm 180°$ correspond to energetically identical conformations. The period of repetition in φ space thus has a length of 180°. This periodicity may be described in what one might call a conformational unit cell plot[33–37].

The unit cells must be centrosymmetric because each conformation possesses a corresponding isoenergetic enantiomeric conformation. The φ data referring to the conformational parameters of the basis structures in their standard geometrical setting (Figure 2) may be expanded to include the data points for the enantiomeric conformations as well as the data points including any number of 180° φ rotations. A further threefold symmetry expansion based on the intrinsic C_3 symmetry of the problem (see above definition of parameters) leads to a set representing all of the symmetry-equivalent isoenergetic conformations[33–37]. The whole set of data may be described as composed of a sequence of multidimensional conformational unit cells. To visualise this space, two-dimensional projections are helpful.

To illustrate the relation between the torsions of phenyl groups bound to one and the same phosphorus atom plots referring to the pairs of φ_1 and φ_2 may be overlaid with those referring to pairs φ_3/φ_4 and φ_5/φ_6 due to the constitutional C_3 symmetry (Figure 9). A high density of points is observed for those regions in conformational space where both of the rings are tilted in the same direction either to the left (**1**) or to the right (**2**). The periodical scatter graph obtained this way also contains information about the different ways in which a pair of phenyl rings of a PPh_2 entity passes from one torsional conformation to its opposite. From the high density of points in region **3** it may be inferred that a correlated motion, in which both rings flip over the 0° position, is the dominant pathway (two-ring flip)[35,53–55]. Region **4** is less densely populated. The dispersion of points in this region, however, indicates that **4** as well corresponds to a possible path. The path involving region **4** corresponds to a conformational change for which φ_{2i} ($i = 1, 2, 3$) has to rotate over the 90° orientation, while φ_{2i-1} (i = 1, 2, 3) simply makes a flip over the 0° position (one-ring flip). Region **5**, again containing only a few points, suggests a pathway analogous but obverse that just

Figure 9. Conformational unit cell plots. Overlays of scatter graphs for pairs τ_1/φ_2, τ_2/φ_4, τ_3/φ_6 (left) and φ_1/φ_2, φ_3/φ_4, φ_5/φ_6 (right) each representing one leg of the *tripod* ligand

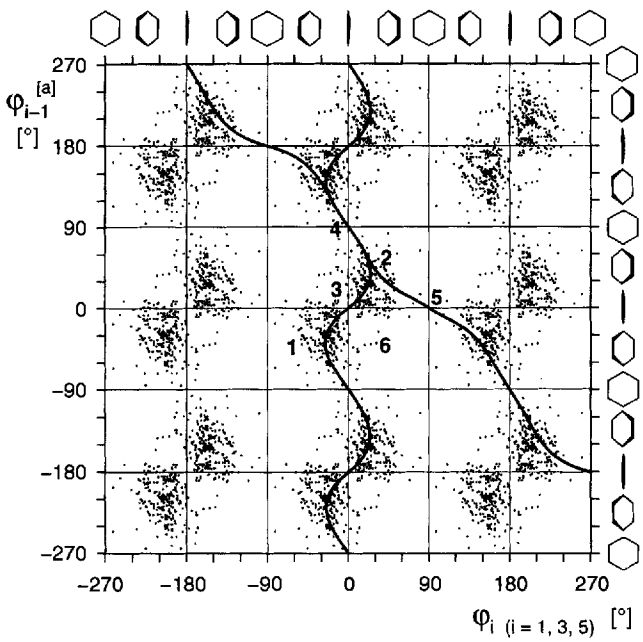


discussed: now φ_{2i-1} must rotate over the 90° position while φ_{2i} (i = 1, 2, 3) flips over the 0° orientation. The two pathways, that along **4** and the one along **5**, do not appear to be equally populated, with that along **4** containing a higher density of data points. There is no symmetry requirement that they be so: φ_{2i} is closer to the space occupied by the coligands compared with φ_{2i-1}. The inequivalence of φ_{2i} and φ_{2i-1} is also evident from the elongation of each of the nine symmetry equivalent patterns in Figure 9 along φ_{2i} (vertical) and also from the lack of mirror symmetry of these patterns: were the two sets φ_{2i} and φ_{2i-1} equivalent, the pattern should show mirror symmetry with the mirror planes running along the diagonals of the cell.

From the two curves indicated in the figures, that oscillating around the vertical position corresponds to an alternating sequence of the two-ring and one-ring flips. The one elongated along the diagonal maps a consecutive series of one-ring flips (Figures 9 and 10).

An overlay of the projections of the conformational unit cell onto the φ_{2i}/τ_i planes (Figure 9) shows that there are two preferred torsional positions on each side of the τ axis, which are, of course, related in pairs due to the inversion symmetry of the problem. Focusing on regions **C** and **D** it is evident that small values of τ are accompanied by positive phenyl torsions, whereas large values of τ are correlated with negative phenyl torsions. During the conformational change leading from the utmost positive torsion to the scaffolding (region **D**) to its opposite, the utmost negative torsion (region **A**), the phenyl rings are thus bound to undergo a pendular motion with three changes of helicity imposed on them.

Figure 10. Conformational unit cell plot. Overlay of scatter graphs for pairs φ_1/φ_6, φ_3/φ_2, φ_5/φ_4, i.e., pairs of neighbouring phenyl groups belonging to different PPh_2 entities. [a] $i - 1 \equiv |i - 1|_{mod\ 6}$, e.g. $i = 1$: $\varphi_{i-1} = \varphi_6$



If the φ values of a phenyl ring on one phosphorus atom are plotted against the φ values of its closest neighbour bound to another phosphorus atom (in the same type of

overlay projection as discussed above, Figure 9) the scatter graph Figure 10 is obtained that – with the densely populated regions **1** and **2** – shows that these pairs of phenyl rings again strongly tend to adopt the same sense of torsional twist.

With respect to idelised mirror planes along the diagonals of the unit cell the pattern is definitely more symmetrical than the pattern discussed above (Figure 10). This means that two proximate phenyl rings bound to alternative phosphorus atoms tend to have similar torsion angles, more so than phenyl rings bound to the same phosphorus atom. Regarding the pathways along which one torsional skew is transformed to its opposite the situation is similar to that found for geminal phenyl pairs (Figure 9) with respect to one-ring flip transitions mapped by regions **4** and **5**. The two-ring flip (region **3**) which is so dominant for geminal pairs (Figure 9) is not as strongly preferred for the adjacent pairs of Figure 10 (compare the density of regions **3**, **4**, and **5** in Figures 9 and 10). The coupling between the torsional motions is not as strong for the adjacent pairs as it is for the geminal ones. This is also evident from the broader scatter of points into the quadrants where the torsional arrangements are of different sense (region **6**).

Neural Networks

All of the above methods are in essence based on deciphering neighbourhood relations within conformational space. As a way to systematically analyse such neighbourhood relations and to recognise the patterns originating from such relations the methodology of neural networks possesses unrivalled power[41–44]. The logic underlying the neural network approach known as fuzzy logic[56,57] is most appropriate for conformational analysis: it naturally allows for a mapping such that a conformation belongs to different classes at the same time, whereby the degree of association with one class or the other is naturally introduced from the very beginning. At the same time, an optimal classification in this sense is also the result of this type of analysis. With algebraic methods based on Boolean logic this type of result is only a posteriori implemented by statistical reasoning while with the strategy of neural networks it is at the basis of reasoning. Neural networks thus account more appropriately for the real situations in nature that are frequently not so clear cut to allow for a simple yes/no alternative. There is yet another advantage: Neural networks do not consider the type of correlations dealing with linear correlations as well as with non-linear ones.

There are many different strategies based on this approach, all of them aimed at the recognition of specific patterns. One specifically helpful type of algorithm has been developed by Kohonen and is called the self-organising mapping network[58–60]. This algorithm results in projections of the data space that strictly sustain the main topological feature of this space even when the dimension of the projection space is strongly reduced as compared to the dimension of the data space.

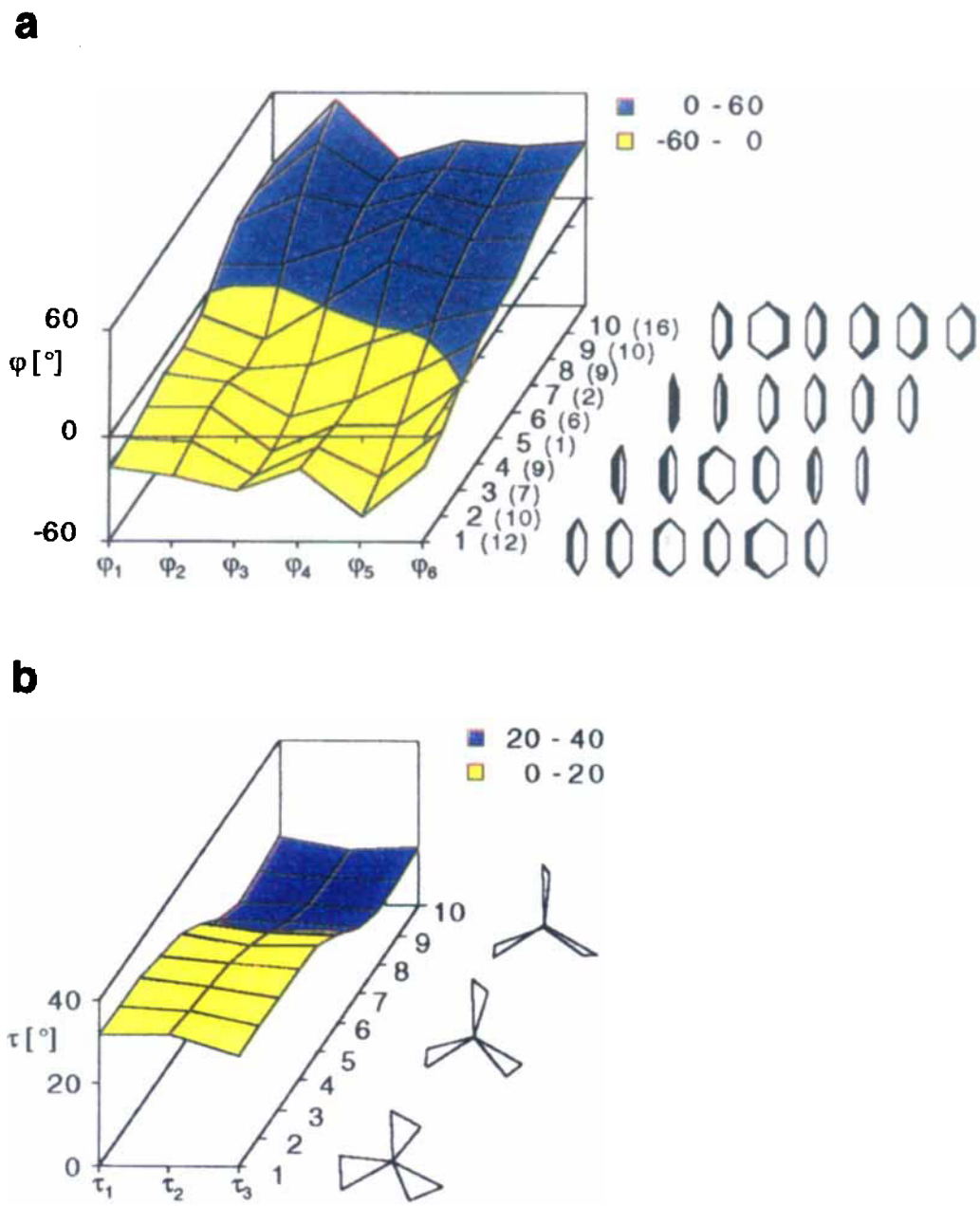
The analysis has been carried out based on the non-symmetry-expanded original data set, as represented by a set of nine-dimensional data vectors composed of the individual six φ and three τ values. This manifold in conformational space is mapped onto a linear array of ten neurons[58] with a network just consisting of the nine input units, holding the values of the nine conformational parameters for each template, and the ten neuron units. The result of applying this strategy to sort the data may be described in a very simple way: At the end of the procedure the network is equivalent to an algorithmic machine, which does nothing different but sorting a bunch of molecular models (each model represented by a data vector) into individual boxes (neurons) in such a way that the models contained in a given box are most similar to each other. At the same time each box is labelled such that the label best descibes the models contained in that box. The boxes find themselves sorted in such a way that the degree of similarity between the labels continuously decreases with the distance between the boxes increasing.

The is achieved by an iterative process for which, given an appropriate similarity criterion (Euclidean distance) and the number of neurons, the network needs nothing else than the data. Filling, sortin, and labelling of the boxes is a self-organised process.

In this way neural networks are able to solve a problem well known in mathematics as the problem of the travelling salesman; in fact, they are the most efficient way to solve this problem[44]. The travelling salesman problem can be described as follows: A salesman must visit a number of towns more or less dispersed in some region. The salesman wishes to organise his travelling in such a way that starting at a given point he pays a visit to each town once and only once with the additional condition that the distance he travels be as short as possible. Transfigured to a journey through conformational space the optimum pathway found by the neural network corresponds to a pathway transforming one set of conformations to the other such that the sets at the extremes are connected by the shortest route possible if all the occupied regions in conformational space are visited just once. In effect, this corresponds to a least-motion pathway through conformational space. Neural networks have been applied in conformational analysis with the aim of predicting the folding of proteins[61–69]. To the best of our knowledge they have not yet been applied to conformational analysis in the sense described here.

The result of the process as applied to the conformational data in φ/τ space is shown in Figure 11. The diagram may be considered as a coherent representation of the labels of the ten "neuron boxes": the characteristic φ and τ values represented by individual curves for each neuron represent the characteristic features of the contents – i.e., the labels – of each neuron box. The sequence of these curves is taken as representing a surface in the diagrams (Figures 11a, 11b). While the analysis is based on nine-dimensional vectors in full φ/τ space the result is represented in two separated diagrams, one pertaining to phenyl torsions φ (Figure 11a), the other one pertaining to τ (Figure 11b), taking into account the basically different meaning of φ and τ. To increase the readability of the diagrams, conformations

Figure 11. Diagram representing the sequence of the labels of a ten-neuron chain, trained with the 82 nine-dimensional data vectors corresponding to the 82 compounds of the data basis. (a) Diagram representing the phenyl torsions $\varphi_1 - \varphi_6$ (symbols to the right: number of neuron (number of conformations mapped onto neuron); visualisation of φ for neurons 1, 4, 7, 10). (b) Diagram representing the cage torsions $\tau_1 - \tau_3$ (symbols to the right: number of neuron; visualisation of an averaged τ for neurons 1, 4, 7, 10)



characteristic of a few individual neurons ("labels of the boxes") are implemented (Figure 11). In Figure 11a the φ values characteristic for neurons 1, 4, 7, and 10 are visualised by a series of phenyl rings tilted according to the values of φ_1 to φ_6 in that order. In Figure 11b the meaning of skew angles τ is represented by diagrams characterising the "labels" of neurons 1, 6, and 10.

The sorting apparent from Figure 11 takes into account the whole range of φ and τ values contained in the data set. Along the chain of neurons moving from neuron 1 to neuron 10 the overall feature represented is obviously an increase along the φ axis from all negative φ values labelling neuron 1 to all positive φ values labelling neuron 10.

A general feature of the distribution of φ values characteristic for the individual neurons is an oscillation along the $\varphi_1 - \varphi_6$ axis. The φ parameters thus form two groups: φ_1, φ_3, φ_5 and φ_2, φ_4, φ_6. These corresponding sets of phenyl groups are different according to the symmetry of the problem (see above).

The spread of φ values for an individual conformation, compared with the whole range of φ space occupied, is nevertheless relatively small, which means that the phenyl torsions must be strongly coupled.

The continuous increase of φ along the chain of neurons

is mirrored by a continuous decrease of τ values along this chain, as shown in Figure 11b. The oscillation of τ values for the conformations represented by the individual neurons are also much smaller than the whole range of τ values representative of the complete data set, which reflects the immediately intelligible strong coupling of the three cage torsions $\tau_1-\tau_3$ (Figure 2). This means that conformations characterised by large positive phenyl torsions φ are at the same time characterised by small τ values and vice versa (see the sections on PCA and PLS).

Grouping of the conformations into three major classes is thus indicated by the graph: Group I is characterised by negative φ and large τ values, as represented by the neurons 1–3, and group II is characterised by large positive φ and small τ values (neurons 8–10). These two groups contain the major part of conformations (78%, compare the number of compounds per neuron in Figure 11a). The remaining neurons (4–7) then represent group III with conformations intermediate between these two extremes.

It is obvious from Figure 11a that the cuts of the surface describing the two main classes of conformation both show a specific modulation characterised by an extreme value of φ_5 (neurons 1, 2) and φ_2 (neurons 9, 10). The two surface cuts are almost mirror images of each another with the numbering of φ reversed. This specific pattern is due to a specific behaviour of the five-coordinate compounds (the same type of pattern is obtained when the five-coordinate compounds alone are used as the basis of the neural network analysis). It reflects the similarity of φ_2 and φ_5 with respect to the positions of the coligands L. Although the helicity of the scaffolding (τ) is not inverted between the two classes the helicity of the phenyl groups is inverted. The interaction between L1 (Figure 5) and φ_2 (Figure 2) is an approximate mirror image of the interaction between L2 (Figure 5) and φ_5 (Figure 2) biased by the torsion τ of the cage.

The major part of the six-coordinate compounds is mapped into group I (neurons 1–3), which means that sixfold and fivefold coordination lead to characteristic differences in the φ and τ values characterising the *tripod* metal templates.

The six-coordinate compounds contained in the data set that are not mapped into group I refer to data that have already been pinpointed as exceptions from the scatter graphs in Figure 4b (ASDPCO10, COPFEA mapped into group II and DEWGAN, CEYMOU, and COJHOY20 mapped onto neuron 4 pertaining to group III; Table 1).

Group III (neurons 4–7), containing conformations intermediate between those represented by groups I and II, is characterised by φ values that are much closer to 0° compared with the φ values of the two other groups. In these cases the $\varphi_1-\varphi_6$ values actually have positive and negative values (Figure 11a). Within group III the transition between the conformations is thus well represented.

In the described sense the sequence of neurons therefore represents a least-motion pathway by virtue of the property inherent to the network algorithm used. This kind of information is specific to the analysis by neural networks.

At the same time, due to specific sorting, a neural network architecture as applied presents a way of classifying a set of data with no classification scheme available at the beginning (see above). Presupposing a rudimentary classification scheme, a closely similar architecture of neural networks may be used to find the classifying power of the appropriate classes (and thus the degree of a specific correlation) for data not originally contained in the basis dataset[42,70]. In this case the network tries to establish box labels (reference vectors), which best accomplish a distinction between the given classes within the data after having sorted them.

This procedure has been applied to the classification of all the five- and six-coordinate *tripod*ML_n (n = 2, 3) structures available in the CSD irrespective of the nature of the coligands L and the type of the metal M[71]. As a training basis only those data referring to the *tripod*Co templates have been used, with the phenyl and backbone torsions as the sole parameters characterising the individual compounds. Nevertheless, the resulting classifying algorithmic machine was able to assign the appropriate classifiers to 85% of all the five-coordinate compounds and to 77% of all the six-coordinate ones.

The fact that it is possible to classify *tripod* metal compounds with respect to their coordination numbers on the basis of a few selected torsional parameters alone should mean that these torsional parameters selectively respond to the steric load imposed by the coligands.

Conclusion

Comparison of Methods

A basic problem underlying any kind of conformational analysis consists of the high dimensionality of the space spanned by the parameters describing the conformations. Any method for systematic conformational analysis must, therefore, reduce the dimensionality of the problem and must find a subspace that adequately describes the variance of the data manifold. To do this a measure of similarity must be defined. The measure underlying all the methods applied is the Euclidean distance in conformational space: the closer the points are the more similar are the conformations. The ways of reducing dimensionality, however, are different for the different methods.

Principal-component analysis reveals a specific set of linear combinations of the original parameters as those factors of primordial importance. In the case studied PCA shows that the component of primordial importance is linked to the average φ and average τ values. This means that scatter graphs based on φ and τ do in fact reveal the most important characteristic of the conformational manifold. The scatter graphs are easily interpreted as two-dimensional plots. The grouping of conformations in two classes of conformations already evident from PCA may also be extracted from the scatter graphs. The very advantage of the scatter graphs is that they suggest conformational transformation pathways that are not so obviously evident from PCA alone.

Cluster analysis is another efficient way of grouping the data into representative sets. In the present case the results of this type of classification procedure are similar to the classification extracted from PCA.

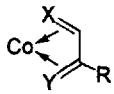
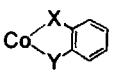
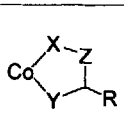
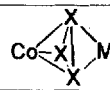
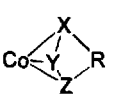
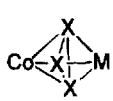
PLS analysis, in the case analysed, helps to rationalise the gear-like motion of the *tripod* metal templates as strongly coupled rotors. Both methods PCA and PLS are limited in the sense that they only recognise linear correlations within the data set. They will only give appropriate results if the data analysed are linearly separable[71].

Neural networks are free from both these restrictions[42,44,59]. As shown in the analysis they reveal all the essential regularities already found by the above methods with the additional advantage of immediately suggesting a least-motion pathway for conformational changes. The method thus well deserves broader application in conformational analysis.

Chemical Implications

The analysis shows that the conformation of *tripod* metal templates in the compounds studied is dominated by the inner forces. The outer forces from the many different crystal environments are obviously too weak to destroy the conformational coherence within the *tripod* metal templates. Even the very much stronger inner forces imposed on the *tripod* metal entities by the coligands L_n in *tripod*ML_n (n = 2, 3) are not sufficient to destroy this coherence. With the broad variation of the kind and type of coligands present

Table 1. Symbolic representation of the solid-state structures forming the basis of the analysis (82 independent templates *tripod*Co; the symbol Co is used in the pictogram to indicate a *tripod*Co entity)

	X	$CN\rightarrow BPh_3$	$NCCN(NH_2)_2$	NH_2-OCH_3		X	CH_2	NH
	Y	$CN\rightarrow BPh_3$	$NCCN(NH_2)_2$	Cl		Y	S	O
						R	H	Ph
		NABCEY	NABGIG	SUHMUD			NABCIC	NABFUR
	X	NH_2	NH	NH	NH_2	NMe_2	S	S
	Y	NH_2	NH_2	NMe_2	O	O	C=S	C=S-$Cr(CO)_5$
		KIRNOO	KIRNII	NABGOM	SUHMIR	SUHMOX	TPHCOA	TRPCOS10 2x
	X	O	O	O	O	O	O	H
	Y	O	O	O	O	O	O	H
	Z	CH	CCH_3	CCH_3	CCH_3	CCH_2Ph	CPh	BH_2
		NABHAZ	ACOPME	NABDID	PIPKEE	NABHED	NABGEC	HBPMCO01
	X	S	S	S	S	S	CH_2	CH_2
	Y	S	S	S	S	S	CH	CCH_3
	Z	CPh	COiPr	CO(2-iPr-5-Me-Ph)	$CSCH_3$	C-N-piperidine	CH_2	CH_2
		NABDUP	NABFAX	NABCUO	BEKGIH	NABFEB	YATHUW	YATJAE
	X	S	S	S	S	N of py	N of py	-N=N=N-
	Y	NH	NH	NH	NCH_3	O at C2 of py	S at C2 of py	-N=N=N-
	Z	CCH_3	$CNH_2\rightarrow BPh_3$	$CNHCH_3$	$CNHCH_3$	C2 of py	C2 of py	Co
		LETSUY	NABFIF	NABHUT	NABDAV	NABCOI	NABFOL	JANLUE
	X	O	O	S	O	S	S	SCH_3
	Y	O	O	S	O	S	S	SCH_3
	Z	SO_2	C=S	C=O	S-pTol	Co*triphos*[a]	Co*triphos*	Co*triphos*
		PMECOS	FIFBAX	BELTIV10	BACZIO	STRCOB	STRCOD	STROAC
	X	O	S	NH	O	O	S	
	Y	O	S	NH	S	NH	NH	
		LALBAB 4x	LAKZUS	LAKZAY	LAKZOM	LAKZEC	LAKZIG	FUJZEP
	X	NH	S	S	N	N	N	
	Y	NH	S	S	O	O	S	
	Z	CH_2	CH_2	CH_2	$N\rightarrow BPh_3$	$N\rightarrow BPh_3$	$N\rightarrow BPh_3$	
	R	H	H	CH_3	$(CH_2)_3CH_3$	$(CH_2)_3OH$	H	
		NABGUS	LALBEF	NABHIH	WESLUB	NABHON	NABGAY	NABDEZ
	X	P	P	P	P	P	P	$PMnCp(CO)_5$
	Y	P	P	P	P	P	$PCr(CO)_5$	$PMnCp(CO)_5$
	Z	P	PH	PCH_3	S	Se	$PCr(CO)_5$	$PMnCp(CO)_5$
		CPPECO10	JOBFUB	DUVXAT	BIHVIX	CEWGIU	CIGNIP10	GESYUY
	X	As	As	As		X	P	P
	Y	S	Se	Te		M	$HgCH_3$	AuP_3Co*triphos*
		CIGNIP10	GESYUY	GESZAF			KIKBIP	KEKGUC 2x
	X	As	P	As		X	P	P	As
	Y	S	P	As		M	$(CuBr)_6P_3$	Fe*tripod*[b]	Co*triphos*
	Z	As	S	S			Co*triphos*		
	R	CPh_2	$Pt(PPh_3)_2$	$Pt(PPh_3)_2$					
		DEWGAN	COJHOY20	CEYMUO			BEFDOF	COPFEA	ASDPCO10 2x

[a] $CH_3C(CH_2PPh_2)_3$. – [b] $CH_3C(CH_2PEt_2)_3$.

in the data set analysed (Table 1) an analogous broad variation in response patterns shown by the conformation of the *tripod* metal templates might be a reasonable expectation. The fact that this is not the case must mean that the forces within the *tripod* metal templates are much stronger than the forces acting between the ligand set L_n and the template.

The forces describing the interaction between the ligands L and the *tripod* metal templates must be analysed by molecular-modelling procedures. For the development of appropriate force fields the results of the present analysis are of basic importance: The analysis shows that the conformations are determined by the inner potential within each molecule so that the observed conformations will correspond to local minima on the energy hypersurface not seriously disturbed by lattice forces. Adequate force field parameters must, therefore, reproduce the observed conformations. Inversely, the observed conformations can serve as a basis for finding these parameters.

We gratefully acknowledge the financial support of this work by the *Deutsche Forschungsgemeinschaft* and the *Fonds der Chemischen Industrie*. We are grateful to the following coworkers of our group for making the data accessible to us before publication: *M. Fritz, A. Barth, S. Vogel, W. Imhof*. We are pleased to acknowledge the continuous interest of Prof. *J. Gasteiger* in our work. Helpful discussion with the members of the Graduiertenkollegs *Selektivität in der Organischen und Metallorganischen Synthese und Katalyse* (chemical institute) and *Modellierung und wissenschaftliches Rechnen in Mathematik und Naturwissenschaften* (IWR, Centre for Interdisciplinary Scientific Computing) warrant special mention. Two of us wish to express their gratitude to the above mentioned Graduiertenkollegs for a fellowship (S. B.) and for a membership (J. H.).

Experimental

Data Set: Table 1 summarizes the compounds analysed. Structural data were retrieved from CSD[4]; the geometrical parameters were computed using XPMA[72] and the QUEST module of the CSD software[4].

PCA: The analysis reported is based on the implementation of factor analysis in the QSAR module of SYBYL[73]. It refers to the autoscaled matrix of nine components (6 φ and 3 τ values).

Hierarchical Clustering: The analysis is based on the hierarchical clustering algorithms as implemented in the QSAR module of SYBYL[73]. The clustering has been applied to the autoscaled matrix of nine components (6 φ and 3 τ values) using the average of all pairwise distances between points in two clusters to evaluate distances between clusters.

PLS: The analyses are based on the implementation of PLS in the QSAR module of SYBYL 6.2[73]. PLS has been applied to several combinations of φ and τ columns using autoscaled data. Cross-validated runs were used to determined the best value for the number of columns necessary to describe the dependent column.

Neural Networks: The program SOM_PAK from the group of Kohonen has been used[74]. It is available via anonymous ftp at cochlea.hut.fi (130.233.168.48). The find module provides two training processes, the first for initialisation and the second for learning. The following parameters were applied:
neighbourhood function = gauss function

1. Initialisation: iteration steps = 1000; learning rate: $\alpha = 0.1$; neighbourhood radius = 5.

2. Learning: iteration steps = 300000; learning rate: $\alpha = 0.01$; neighbourhood radius = 2.

The program LVQ_PAK (Learning vector quantisation) was used applying several runs of the lvq_run module[75]. The initialisation of the reference vectors has been chosen "allocated" to the number of entries for each class in the data set. The result presented refers to the classifiers obtained on the basis of two reference vectors for the five-coordinate and one reference vector for the six-coordinate compounds.

Hardware: All calculations were carried out by using Silicon Graphics workstations Indigo^2, MIPS R4400, 200 Mhz, 64 MB RAM, 2 GB hard disk.

☆ Dedicated to *E. Weiss* on the occasion of his 70th birthday.

[1] J. P. Collman, L. S. Hegedus, J. R. Norton, R. G. Finke, *Principles and Applications of Organotransition Metal Chemistry*, University Science Books, Mill Valley, California, **1987**.
[2] B. C. Gates, *Catalytic Chemistry*, Wiley, New York, **1992**.
[3] H. Brunner, W. Zettlmeier, *Handbook of Enantioselective Catalysis with Transition Metal Compounds*, vol. 1 & 2, VCH, Weinheim, **1993**.
[4] F. H. Allen, O. Kennard, *Chemical Design Automation News*, **1993**, *8*, 1 & 31–37.
[5] O. Walter, T. Klein, G. Huttner, L. Zsolnai, *J. Organomet. Chem.* **1993**, *458*, 63.
[6] H. Heidel, G. Huttner, G. Helmchen, *Z. Naturforsch. Part B*, **1993**, *48*, 1681.
[7] A. Muth, O. Walter, G. Huttner, A. Asam, L. Zsolnai, C. Emmerich, *J. Organomet. Chem.* **1994**, *468*, 149.
[8] T. Seitz, A. Muth, G. Huttner, T. Klein, O. Walter, M. Fritz, L. Zsolnai, *J. Organomet. Chem.* **1994**, *469*, 155–162.
[9] A. Muth, A. Asam, G. Huttner, A. Barth, L. Zsolnai, *Chem. Ber.* **1994**, *127*, 305.
[10] B. C. Janssen, A. Asam, G. Huttner, V. Sernau, L. Zsolnai, *Chem. Ber.* **1994**, *127*, 501.
[11] H. Heidel, G. Huttner, R. Vogel, G. Helmchen, *Chem. Ber.* **1994**, *127*, 271.
[12] A. Muth, G. Reinhard, G. Huttner, T. Seitz, T. Klein, L. Zsolnai, *Z. Naturforsch., Part B*, **1994**, *49*, 889.
[13] T. Seitz, G. Huttner, M. Büchner, *Z. Naturforsch., Part B*, **1994**, *49*, 1813.
[14] T. Seitz, A. Muth, G. Huttner, *Chem. Ber.* **1994**, *127*, 1837.
[15] B. C. Janssen, V. Sernau, G. Huttner, A. Asam, O. Walter, M. Büchner, L. Zsolnai, *Chem. Ber.* **1995**, *28*, 63.
[16] H. Heidel, G. Huttner, L. Zsolnai, *Z. Naturforsch., Part B*, **1995**, *50*, 729.
[17] H. Heidel, J. Scherer, A. Asam, G. Huttner, O. Walter, L. Zsolnai, *Chem. Ber.* **1995**, *128*, 293.
[18] T. Seitz, A. Asam, G. Huttner, O. Walter, L. Zsolnai, *Z. Naturforsch., Part B*, **1995**, *50*, 1287.
[19] T. Seitz, A. Muth, G. Huttner, *Z. Naturforsch., Part B*, **1995**, *50*, 1045.
[20] G. Reinhard, R. Soltek, G. Huttner, A. Barth, O. Walter, L. Zsolnai, *Chem. Ber.* **1996**, *129*, 97–108.
[21] S. C. Tsai, H. E. Wang, S. T. Huang, L. M. Yiin, S. T. Liu, *Chem. Ber.* **1995**, *128*, 151–155.
[22] S. T. Liu, K. J. Liu, *Inorg. Chem.* **1990**, *29*, 4576–4579.
[23] S. T. Liu, C. L. Tsao, M. C. Cheng, S. M. Peng, *Polyhedron* **1990**, *9*, 2579–2584.
[24] S. T. Liu, H. E. Wang, M. C. Cheng, S. M. Peng, *J. Organomet. Chem.* **1989**, *376*, 333–342.
[25] H. E. Wang, C. Y. Liu, M. C. Cheng, S. M. Peng, S. T. Liu, *Phosphorus, Sulfur, Silicon Related Elements* **1992**, *69*, 201–211.
[26] M. J. Burk, R. L. Harlow, *Angew. Chem.* **1990**, *102*, 1511–1513; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 1467.
[27] M. J. Burk, J. E. Feaster, R. L. Harlow, *Tetrahedron: Asymmetry* **1991**, *2*, 569–592.
[28] A. Mezzetti, L. M. Venanzi, *A Novel Homochiral, Tripod-like Phosphine Ligand and Some of its Ruthetium(II) Complexes*, FECHEM conference on organometallic chemistry, Greece, **1993**.
[29] C. Bianchini, A. Meli, M. Peruzzini, F. Vizza, P. Frediani, J. A. Ramirez, *Organometallics* **1990**, *9*, 226–240.

[30] V. Sernau, G. Huttner, M. Fritz, L. Zsolnai, O. Walter, *J. Organomet. Chem.* **1993**, *453*, C23–C29.
[31] T. R. Ward, L. M. Venanzi, A. Albinati, F. Fianza, T. Gerfin, V. Gramlich, G. M. Ramos-Tombo, *Helv. Chim. Acta* **1991**, *74*, 983.
[32] C. Bianchini, A. Meli, M. Peruzzini, F. Vizza, F. Zanobini, *Coord. Chem. Rev.* **1992**, *120*, 193.
[33] H.-B. Bürgi, J. D. Dunitz, *Structure Correlation*, vol. 1 & 2, VCH, Weinheim, **1994**.
[34] B. J. Dunne, R. B. Morris, A. G. Orpen, *J. Chem. Soc., Dalton Trans.* **1991**, 653–661.
[35] D. A. V. Morton, A. G. Orpen, *J. Chem. Soc., Dalton Trans.* **1992**, *4*, 641–653.
[36] S. E. Garner, A. G. Orpen, *J. Chem. Soc., Dalton Trans.* **1993**, *4*, 533–541.
[37] A. G. Orpen, *Chem. Soc. Rev.* **1993**, *22*, *3*, 191–197.
[38] E. R. Malinovsky, *Factor Analysis in Chemistry*, Wiley, New York, **1991**.
[39] R. Henrion, G. Henrion, *Multivariate Datenanalyse*, Springer, Heidelberg, **1995**.
[40] *SYBYL 6.2-Ligand-Based Design Manual*, Tripos, St. Louis, **1995**, p. 215–253.
[41] J. Gasteiger, J. Zupan, *Neural Networks in Chemistry*, VCH, Weinheim, **1993**.
[42] A. Zell, *Simulationen Neuronaler Netze*, Addison-Wesley, Bonn, **1994**.
[43] K. P. Kratzer, *Neuronale Netze – Grundlagen und Anwendungen*, Hanser, Hamburg, **1993**.
[44] H. Ritter, T. Martinez, K. Schulten, *Neuronale Netze – Eine Einführung in die Neuroinformatik selbst-organisierender Netzwerke*, Addison-Wesley, Bonn, **1994**.
[45] The more conventional definition of these torsional parameters would of course be the Prelog-Dunitz-Bürgi torsion angle as defined by the sequence of atoms C_{CH_2}–P–$C_{i\ Phenyl}$–$C_{ortho\ Phenyl}$ and is an equivalent to the one used here and within the small range of backbone torsional angles τ; the parameters as defined by the two different methods are almost exactly linearly correlated. The choice of the definition used within this work is a matter of visualisation: The geometrical meaning of the torsion angles φ is immediately evident in any projection parallel to the P1P2P3 plane (Figure 2a).
[46] E. L. Muetterties, L. J. Guggenberger, *J. Am. Chem. Soc.* **1974**, *96*, 1748–1756.
[47] R. R. Holmes, *Progr. Inorg. Chem.* **1984**, *32*, 119.
[48] J. S. Wood, *Prog. Inorg. Chem.* **1972**, *16*, 227.
[49] Since any general ligand position L is just one of the six positions equivalent under this conformational symmetry every compound *tripod*ML_3 gives rise to 18 data points.
[50] M. D. Vaira, L. Sacconi, *Angew. Chem.* **1982**, *94*, 338–351; *Angew. Chem. Int. Ed. Engl.* **1982**, *21*, 330.
[51] At a finer resolution these maxima are found to be composed of two maxima one to the left and one to the right of the bisectors (χ = 60°, 180°, 300°). This reflects the fact that the *tripod*M fragment has inherent C_3 and not C_{3v} symmetry. By its helicity it may well distort the idealised octahedron pushing the coligands out from an idealised mirror symmetric location. At a coarse resolution this splitting is not apparent (Figure 6b).
[52] While the computational result of cluster analysis is uniquely determined, the pictorial result using dendrograms depends on the sequence of compounds as fed into the analysis. The sequence underlying the dendrogram (Figure 8) is arranged such as to best represent the sequence obtained by neural network analysis; it nevertheless nicely reproduces the classification inferred from PCA.
[53] K. Mislow, *Acc. Chem. Res.* **1976**, *9*, 26–33.
[54] H. B. Bürgi, *Acc. Chem. Res.* **1983**, *16*, 153–161.
[55] H. B. Bürgi, *Acta Crystallogr., Sect. B*, **1988**, *44*, 445–448.
[56] T. Tilli, *Mustererkennung mit Fuzzy-Logik*, Franzis, München, **1993**.
[57] B. Kosko, S. Isaka, *Scientific American* **1993**, *7*, 62.
[58] T. Kohonen, J. Kangas, J. Laaksonen, *SOM PAK, The Self-Organizing Map Program Package*, Version 1.2., Manual, University of Technology, Helsinki, **1992**.
[59] T. Kohonen, *Self-Organization and Associative Memory*, Springer, Heidelberg, **1984**.
[60] T. Kohonen, *Biol. Cybernetics* **1982**, *43*, 59–69.
[61] N. Quian, T. J. Sejnowski, *J. Mol. Biol.* **1988**, *202*, 865–884.
[62] L. H. Holley, M. Karplus, *Proc. Natl. Acad. Sci.* **1989**, *86*, 152–156.
[63] D. G. Kneller, F. E. Cohen, R. Langridge, *J. Mol. Biol.* **1990**, *214*, 171–182.
[64] H. Bohr, J. Bohr, S. Brunak, R. M. J. Cotterill, B. Lautrup, L. Norskov, O. H. Olsen, S. B. Petersen, *FEBS Lett.* **1988**, *24*, 223–228.
[65] H. Bohr, J. Bohr, S. Brunak, R. M. J. Cotterill, F. Fredholm, B. Lautrup, O. H. Olsen, S. B. Petersen, *FEBS Lett.* **1990**, *261*, 43–46.
[66] H. Andreassen, H. Bohr, J. Bohr, S. Brunak, T. Bugge, J. Cotterill, C. Jacobsen, P. Kusk, B. Lautrup, *J. Acquired Immune Defic. Syndr.* **1990**, *3*, 615–622.
[67] S. Brunak, J. Engelbrecht, S. Knudsen, *Nucleic Acid Res.* **1990**, *18*, 4797–4801.
[68] P. Y. Chou, G. D. Fasman, *Biochemistry* **1974**, *13*, 211–222.
[69] P. Y. Chou, G. D. Fasman, *Biochemistry* **1974**, *13*, 222–241.
[70] T. Kohonen, J. Kangas, J. Laaksonen, K. Torkkola, *The Learning Vector Quantization Program Package*, Version 2.1., Manual, University of Technology, Helsinki, **1992**.
[71] See data basis given in Table 1. Additional Refcodes of *tripod* metal templates with M ≠ Co, which were retrieved from CSD under the restriction of 3D parameter definitions: BUFREZ, BUFREZ10, CAWFIP, COGWOK, CLNUS10, DEPNER, DIVIDIV, DUMSUZ, FONMUQ, GINLEU, JEDAUO, JEFYAU, PMENSE, BAVRIZ, BAVROF, BAVRO, BISWEF10, BUSLAC, BUSLEG, BUTSIS10, CAFSOR, CASCOO, COVGAV10, CUFTIG10, DAHFIB, DEGZUK, DEYTEG, DOJPIB, DUHZUB, FABJUN, FATXON, FAZFUH, FIDLUZ, FOHMOE, GAGVIT, GINMIZ, GINMOF, JITZIV, KATGIV, KEDBIE, KEKHAJ, KISHID, LAKYUR, PAKGOX, PEFYOO, PPMEIR, PPMERH, SACKAI, TAMGOD, VEKRIM, VITZON10.
[72] L. Zsolnai, G. Huttner, *XPMA*, Universität Heidelberg, Heidelberg, **1995**.
[73] SYBYL 6.2, Tripos, St. Louis, **1995**.
[74] T. Kohonen, J. Kangas, J. Laaksonen, *SOM PAK, The Self-Organizing Map Program Package*, Version 1.2., University of Technology, Helsinki, **1992**.
[75] T. Kohonen, J. Kangas, J. Laaksonen, K. Torkkola, *The Learning Vector Quantization Program Package*, Version 2.1., University of Technology, Helsinki, **1992**.

[95219]