Acta Cryst. (1998). A**54**, 873–885

Structure Correlation and Chemistry

H. B. Bürgi

University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland. E-mail: hbuergi@krist.unibe.ch

(*Received 8 April 1998; accepted 4 August 1998*)

Abstract

The main goal of crystal and molecular structure determination is to provide a starting point for understanding the physical, chemical and biological properties of matter. At present, results from nearly 300 000 crystal structure studies are available in computer-readable form. Structure correlation attempts to extract knowledge and understanding from this body of information, which is not available from its parts. This article reviews some typical examples: libraries of prototypal molecular dimensions, mappings of chemical reaction pathways, correlations between structure on one hand and energy, reaction rate, catalytic activity or magnetism on the other. The knowledge gained from structure-correlation studies, together with quantum-chemical and other modeling techniques, provides conceptual and practical tools for designing molecules and materials with tailor-made properties.

1. Introduction

Meeting our needs for raw materials, for food and clothing, for pharmaceuticals, color and aroma, requires some understanding of chemical substances: their synthesis and modification, their physical and chemical properties, their biological effects, and their processing in the fabrication of marketable products. At a much more fundamental level, chemistry deals with the distributions in space of the sluggish positively charged nuclei and the whirling, negatively charged electrons. The distribution of the nuclei can be made visible through diffraction of a beam of neutrons, that of the electrons through diffraction of a beam of X-rays. The picture emerging from such experiments shows the net result of an intricate balance of attractions and repulsions between the charged atomic particles: a three-dimensional scaffold of nuclei embedded in a haze of electrons, dense in the neighborhood of nuclei and thinning out in between. Measuring and understanding these distributions is the subject matter of crystallography, of crystal and molecular structure determination.

But what is the relationship between chemistry and structure determination? Hoffmann (1997) proposed a concise yet comprehensive program to answer this question: 'There is no more basic enterprise in chemistry than the determination of the geometrical structure of a molecule. Such a determination, when it is well done, ends speculation and provides us with the starting point for the understanding of every physical, chemical and biological property of the molecule. Indeed, the chemical sciences ... are what they are today largely as the result of careful structure determination'. While this view of a leading chemist is rather flattering to crystallographers, it also signals limitations: a crystal structure determination is rarely the answer to a problem, it is a prelude to an answer. The purpose of this article is to illustrate instances of the interplay between chemistry and crystallography, where 'seeing' the architecture of matter at the atomic level has led to a better understanding of chemical phenomena.

The early development of structure determination has been aptly summarized in the leading sentences of the Editorial Preface (1948) to Volume 1 of *Acta Crystallographica*: 'Crystallography was transformed into an effectively new science by the discovery of X-ray diffraction in crystals in 1912. In the following two decades an enormous development took place, forging into a practical tool what had originally been a physical experiment, and the atomic structures of hundreds of substances of ever increasing complexity were established'. Since 1948, the number of crystal structures established by X-ray and neutron diffraction has increased by orders of magnitude and will approach ~300 000 by the year 2000! Similarly, the complexity of the structures has reached a stage where functional assemblies of proteins, DNA and RNA, built from thousands of atoms, can be 'seen' with the help of X-rays [see the articles by Holmes (1998), Yonath *et al.* (1998) and Rossmann (1998) in this Special Issue for examples].

Hans-Beat Bürgi obtained his PhD in 1969 from the Eidgenössische Technische Hochschule (ETH), Zürich, after working with Professor J. D. Dunitz. From 1969 to 1971, he was a postdoctoral associate with Professor L. S. Bartell at the University of Michigan. After completing his Habilitation at the ETH, Zürich, he was appointed Professor of X-ray crystallography at the University of Bern in 1979. His research interests are in structural chemistry (see http://iris.unibe.ch/home.html).

© 1998 International Union of Crystallography
Printed in Great Britain – all rights reserved

Several technological advances have made this development possible: today's X-ray lamps, the synchrotrons, are brighter by many orders of magnitude than the early home-made X-ray tubes [see the articles by Helliwell (1998) and by Hart & Berman (1998) in this Special Issue]; fast highly sensitive area detectors have reduced data-collection times from weeks to hours; the ever increasing power of electronic computers makes it possible to process larger and larger data sets ever more speedily. It seems that in none of these areas have the limits yet been reached. However, the basic 'physical experiment' is still essentially the same as in 1912.

2. Structural systematics

The notion of molecular and crystal structure has two aspects: *molecular* or *crystal geometry* refers to internuclear distances and angles, *electronic structure* to the distribution of electrons. Both aspects are necessary for implementing Hoffmann's program. Because of the mutual interactions of nuclei and electrons, each of the distributions carries an imprint of the other and much can be learned by considering only one of the two. In this article the emphasis is on the nuclear scaffold, especially its geometric details. Electron distributions are discussed in the paper by Coppens (1998) in this Special Issue.

The determination of crystal and molecular structures is now a routine analytical tool, widely used for characterizing newly synthesized compounds. For most classes of chemicals, large numbers of structures have accumulated. This has opened a new field for chemical crystallography, that of structural systematics. The idea is to address physical, chemical and biological questions which cannot be answered with a single crystal structure, but might be clarified by comparisons of many such structures. The field of structure correlation may be divided into two major parts. One is structure–structure correlations. It searches for interdependencies among structural descriptors such as interatomic distances and angles. The other comprises structure–property correlations. Here the goal is to establish and understand the relationships between structure and physical, chemical or biological properties. The following sections will present examples of both kinds of correlations, selected with a bias towards chemistry.

3. Chemical bonds

The notion that atoms and their neighbors are connected by chemical bonds and assume specific geometric arrangements in three-dimensional space was derived from purely chemical experiments in the second half of the 19th century, elegantly confirmed by X-ray diffraction after 1912 and interpreted by quantum mechanics in the 1920s. By 1938, when Pauling wrote the first edition of his influential book on *The Nature of the Chemical Bond*, a substantial number of bond distances had been measured (Pauling, 1960). Although the information was very limited compared to what is known today, it was sufficient to establish some empirical rules: (i) bond distances between a given pair of atoms in similar environments are the same within narrow limits, often within the standard uncertainties of the measurements, and (ii) these distances can be represented, approximately, as the sums of typical atomic dimensions, the covalent radii

$$d_o(AB) = r_o(A) + r_o(B). \tag{1}$$

This connection between the nature of the atoms (or nuclei) participating in a bond and the length of the bond is one of the simplest examples of a structure–property correlation.

Changing a typical bond distance requires energy. A reasonably accurate description of the dependence of energy on interatomic distance is given by the Morse potential (Fig. 1). It can be considered as a combination of a repulsive and an attractive part:

$$V(d - d_o) = D_o\{\exp[-2B(d - d_o)] - 2\exp[-B(d - d_o)]\}. \tag{2}$$

d_o is the equilibrium bond distance and D_o is the bond energy at the minimum of the curve; B is an empirical constant. In some cases, equilibrium distances vary significantly and continuously with environment and the concept of atomic radii becomes meaningless. In aromatic and other unsaturated hydrocarbons, for example, the distances between bonded carbon atoms range from ~1.2 to ~1.5 Å. This observation can be interpreted in terms of a variable balance between the repulsive and attractive terms of the Morse potential. To change the balance, one may multiply the attractive part with a variable factor n (or, more generally, some power

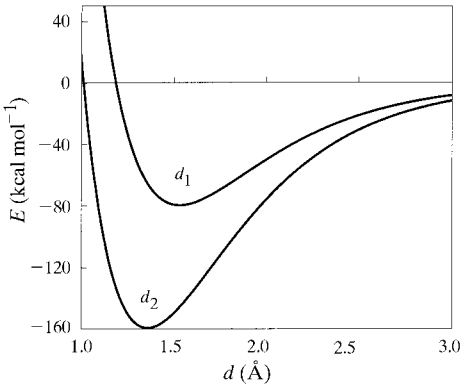


Fig. 1. Top: Morse potential for a C—C single bond [equation (2) with $d_o = d_1 = 1.5$ Å and $D_o = 80$ kcal mol^{-1}; 1 kcal mol^{-1} = 4.1868 kJ mol^{-1}]. Bottom: modified Morse potential for a C—C double bond [equation (3), with $n = 2$, $d_e(2) = d_2 = 1.3$ Å, $D(2) = 160$ kcal mol^{-1}].

of it) and leave the repulsive term unaltered (Bürgi & Dunitz, 1987):

$$V(d-d_o) = D_o\{[\exp[-2B(d-d_o)] - 2n^q[\exp -B(d-d_o)]\}. \quad (3)$$

d_o is the equilibrium distance and D_o the bond energy for $n = 1$; q and B are empirical constants. Several relationships between the main properties of chemical bonds follow. The equilibrium distance d_e varies with the modification factor n, also called the bond order, according to

$$d_e(n) = d_o - qB^{-1}\ln n = d_o - c\ln n. \quad (4)$$

This is Pauling's bond-length–bond-order equation: the higher the bond order n, the shorter the bond. The bond energy as a function of bond order is

$$D(n) = D_o n^{2q}. \quad (5)$$

Equations (4) and (5) imply high bond energy for short bonds and *vice versa*. Small distortions Δd from equilibrium require a deformation energy which, according to Hooke's law, is $\sim k\Delta d^2/2$. The stretching force constant k can also be expressed in terms of bond order

$$k(n) = 2B^2 D_o n^{2q}. \quad (6)$$

Equations (4) and (6) imply that short bonds are stiff and long bonds are soft. The constants in equations (3)–(6) can be chosen to reproduce quite a range of experimental data, at least as far as order of magnitude is concerned. With $q \simeq 1/2$, $B \simeq 2$ Å^{-1}, as found for many diatomic molecules, and with a C—C single-bond energy $D(1)$ of 83 kcal mol^{-1}, we obtain $D(2) \sim 165$, $D(3) \sim 250$ kcal mol^{-1}, $k(1) \sim 4.6$, $k(2) \sim 9.2$ and $k(3) \sim 13.8$ mdyn Å^{-1}, all of the same order of magnitude as tabulated values. The modified Morse equation and the interrelations derived from it are further examples of structure–property correlations. They express the widespread intuition that bonds become weaker and softer as they become longer. This notion has reached almost the status of a dogma. Note, however, that there may be violations of the dogma and that the general trends should be applied to specific cases with caution (for a recent discussion, see Ernst *et al.*, 1995).

There are many ways to interpret the modifying factor n^q. Pauling, who favored the valence-bond theory of chemical bonding, was arguing in terms of resonance between different valence-bond structures (Lewis formulae). Those who prefer molecular-orbital theories of chemical bonding interpret n^q in terms of aromaticity, conjugation, hyperconjugation, and many other such concepts. Both schools invoke steric repulsions between the ligands attached to the bonded atoms. This is not the place to go into details, but note that there are fairly simple schemes capable of connecting experimental observations with theoretical descriptions relevant to the concept of a chemical bond.

Since most atoms form several bonds to neighboring atoms, the question arises as to how they may be interrelated. It is often found that the sum of bond orders around an atom is nearly constant. This implies that shortening one of the bonds requires lengthening others. Bond-order conservation has been found to hold throughout the Periodic Table of the elements. A limited number of reference distances together with a single numerical value of c [equation (4)] and a bond-order conservation rule are able to rationalize the observed ranges of bond distances for over a thousand atom pairs (Brown & Altermatt, 1985). Bond-order conservation rules are useful for testing unusual coordination numbers and geometries, for locating missing atoms in crystal structures and, more generally, for modeling inorganic crystal structures and chemical reaction paths (Brown, 1994).

4. Chemical reaction paths, general considerations

In this and the following section, an attempt is made to connect two phenomena which, at first sight, seem difficult to reconcile: the static regularity of crystal structures on one hand and the dynamic disarray of reacting molecules on the other. In his review of donor–acceptor interactions, Bent (1968) discusses the structures of polyiodide ions in their crystalline salts, in particular the two interatomic distances $d(I_1I_2)$ and $d(I_2I_3)$ in the linear triiodide fragment $[I_1I_2I_3]^-$ (Fig. 2). He notes a reciprocal relationship between the two distances and hypothesizes that 'the lengthening of an intramolecular interaction with a shortening of a *trans* intermolecular interaction (until finally, perhaps, the two interactions become indistinguishable) may be a general phenomenon'. Bent offers the following interpretation for this prototypal structure–structure correlation: 'the hyperbolic like curve may be presumed to show, approximately, the changes that occur in the distances between nearest neighbors in the linear exchange reaction $I_1^- + I_2I_3 \rightleftharpoons I_1I_2 + I_3^-$. In order to understand Bent's analogy, it is first necessary to understand why the distances of the triiodide anion differ with different cations. An isolated triiodide is linear and symmetric. In a crystal lattice, it can undergo geometric deformation and charge redistribution under the influence of cations. On the one hand, this will cost intramolecular deformation (or reorganization) energy, on the other hand, intermolecular packing energy can be gained. The observed deformations correspond to maximum gain at minimal cost. It follows from the general arguments given above that the reorganization energy is kept small if as much bonding as possible is maintained during deformation, *i.e.* if weakening one bond is constantly compensated by strengthening another. Thus, assume that the sum of bond orders in triiodide remains

constant during deformation; the dependence of $d(I_1I_2)$ on $d(I_2I_3)$ becomes

$$\text{constant} = n(I_1I_2) + n(I_2I_3) = \exp[d(I_1I_2) - d_o]/c + \exp[d(I_2I_3) - d_o]/c. \quad (7)$$

This hyperbolic like curve of constant total bond order fits the experimental observations in Fig. 2 as well as can be desired (Bürgi, 1975).

A slightly more complex but chemically more interesting situation is found for linear triatomic fragments with additional ligands on the central atom. They exhibit tetrahedral, trigonal bipyramidal and intermediate coordination geometries:

Detailed analysis of crystal structure data for different YML_3X fragments have shown that bond reshuffling follows curves of constant bond order, similar to that of I_3^-, and is correlated to a high degree with the angular rearrangement of the ligands *L*. This structure–structure correlation can be considered as the reaction path for an associative ligand exchange or S_N2 reaction proceeding with Walden inversion (for a review, see Bürgi & Shklover, 1994).

How do these observations on solids relate to a chemical reaction in solution? In analogy to the solid state, a $Y\cdots ML_3X$ fragment in solution will adapt its geometry and charge distribution to the perturbations of the surrounding solvent molecules to optimize solvation energy. However, contrary to the solid state, the perturbations in solution are rapidly changing with time. Occasionally, a particular solvent arrangement will drive $Y\cdots ML_3X$ to relax to $YML_3\cdots X$, *i.e.* to undergo an exchange reaction. The concentration of molecules about to react is far too low and the rearrangement itself far too rapid to be observed by diffraction. However, if we assume that the laws governing chemical bonding and intermolecular interactions are the same in the solid and in solution, a collection of YML_3X structures in the solid can be considered as a series of snapshots, representative of an average reaction trajectory in solution.

5. The role of environment

S_N2 paths have been traced from structural data for a variety of central atoms *M*, including Al, Si, Ge, Sn, Zn, Cd, but not for carbon, which, at least for an organic chemist, is probably the most interesting case. Superficially, this is because there are no crystal structures of organic compounds with partial bonds to carbon analogous to those found for Si, Ge, Sn *etc*. From an energy point of view, it means that the reorganization energy needed to deform the stable arrangements $Y\cdots CL_3X$ or $YCL_3\cdots X$ to $Y\cdots CL_3\cdots X$ are far too large to be compensated by an increase in intermolecular packing energy. In solution, the situation is different. Every so often, solvent molecules with higher than average kinetic energy produce an environment which induces a $Y\cdots CL_3X$ structure to relax to $YCL_3\cdots X$, *i.e.* to undergo a substitution reaction. *Ab initio* calculations have shown that the reaction path for ligand exchange at carbon is similar geometrically to those at Si, Ge, Sn *etc*., but also that the reorganization (or activation) energy necessary to reach the intermediate $Y\cdots CL_3\cdots X$ geometry is much higher.

The interplay between the reorganization of a molecule and its interaction with environment is strikingly illustrated by the next example. Recently, structures of donor–acceptor complexes have been determined in the gas phase by microwave spectroscopy and quantum-chemical calculations, and in the solid state by X-ray crystallography (for a review, see Leopold *et al.*, 1997). The relevant data for $H-C\equiv N\cdots BF_3$ are summarized in Fig. 3. Note the dramatic decrease of the $N\cdots B$

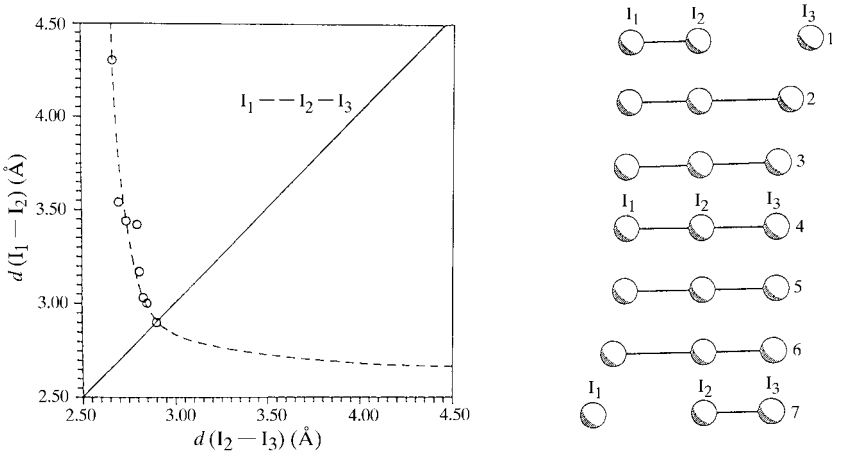


Fig. 2. Scatterplot of observed I_1-I_2 *versus* I_2-I_3 distances (left) and seven structures along the corresponding correlation curve (right) tracing the geometrical changes which may be presumed to occur along the reaction path of the linear exchange reaction $I_1^- + I_2-I_3 \rightleftharpoons I_1-I_2 + I_3^-$. Data from Bent (1968). Reprinted with permission from Ferretti *et al.* (1996). Copyright (1996) Gordon and Breach Publishers.

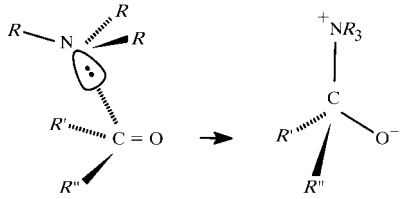
distance by ~0.8 Å and the equally dramatic increase of the N···B—F angles by 14° on going from the gas phase to the solid state; a smaller increase of ~0.05 Å is observed for the three B—F bond distances. The geometric changes are accompanied by an increase in the 'molecular' dipole moment from ~4 to ~9 debye! What is the picture in terms of energies? The sublimation of $HCN—BF_3$(s) into HCN(g) and BF_3(g) requires ~22 kcal mol^{-1} as estimated from vapor pressures. According to *ab initio* calculations, the formation of $HCN\cdots BF_3$(g) returns about −5 kcal mol^{-1} and the deformations of distances and angles to the solid-state values require ~6 kcal. This leads to a packing energy of about −23 kcal mol^{-1} for the deformed molecules with the high dipole moment. Assuming a similar arrangement for the undeformed molecules, with low dipole moment, the packing energy estimated from the ratio of the dipole moments is much smaller, by at least a factor of four. Thus, the initial investment in reorganization energy is clearly overcompensated by the return in extra packing energy.

$H—C\equiv N\cdots BF_3$ is not a singular example. Analogous but smaller structural differences have been observed for $CH_3—C\equiv N\cdots BF_3$, $(CH_3)_3N\cdots BF_3$ and $H_3N\cdots BF_3$. All structures taken together show a correlation between the shortening of the N···B distance and the pyramidalization of the BF_3 fragment. The empirical relationship is

$$d = d_o - c\ln[9\cos^2\alpha(\mathrm{FBN})] \qquad (8)$$

with d_o = 1.57 Å and c = 0.19 Å. When α is 90°, cos α is zero and the N···B distance is infinite, there is no complexation; as α approaches 109.4°, the value for a regular tetrahedron, cos α approaches 1/3 and the N···B distance becomes 1.57 Å, corresponding to complexation and subsequent deformation as described above. Note that the structure–structure correlation expressed by equation (8) has the form of a bond-order–bond-length relationship with the bond order expressed in terms of bond angles. The series of structures with various substituents at nitrogen and different molecular packing maps the reaction path for addition of a donor to BF_3.

The addition reaction of nucleophiles to carbonyl groups, *e.g.*



has many parallels with the formation of the $R_3N—BF_3$ complexes. As the N—C bond forms, the carbonyl group becomes pyramidal and the C—O bond lengthens. In solution, further steps follow but the reaction is thought to begin as indicated. The corresponding reaction path is well documented and need not be elaborated here (for a review, see Cieplak, 1994). We mention it only because the N···C═O interaction also exhibits a distinct dependence on environment. According to Leonard, the infrared absorption of the carbonyl group in the bicyclic amino ketone shown in Fig. 4 is strongly solvent dependent; it is found at 1666 cm^{-1} in the apolar solvent cyclohexane and at 1613 cm^{-1} in polar chloroform (for a review, see Leonard, 1997). The change indicates a softening and thus lengthening of the C═O bond. Similarly, the chemical shift of the carbonyl carbon atom, measured by nuclear magnetic resonance spectroscopy, is 172 p.p.m. in cyclohexane solutions, but 102 p.p.m. in chloroform (Nakashima & Maciel, 1972). The former value is close to that for an unperturbed carbonyl group, whereas the latter indicates a strong N—C interaction. Both experiments show, albeit indirectly, that an increase in solvent polarity moves the structure of the molecule towards the zwitterionic form $N^+—C—O^-$.

The relationships used in this and the preceding sections to compare molecular structures in different environments and to connect static crystal structures with the dynamics of chemical reactions have been judged 'quite archaic by modern standards', *i.e.* in comparison to the sophisticated results obtainable from quantum chemistry (Reeve *et al.*, 1993). We agree, the picture is oversimplified. At the same time, we take note

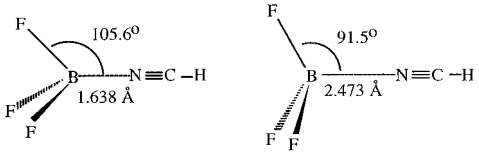


Fig. 3. Molecular structures of the donor–acceptor complex $BF_3\cdots NCH$ in the crystalline state (left) and in the gas phase (right).

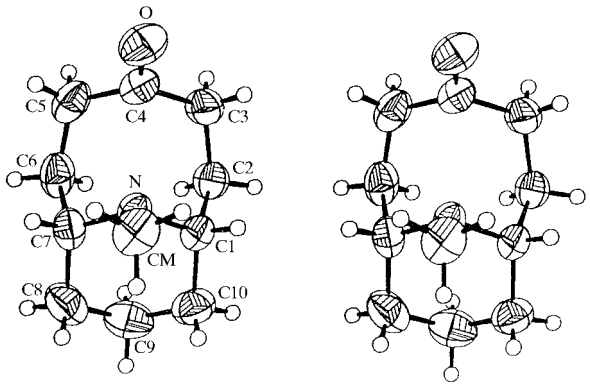


Fig. 4. Stereoscopic view of 11-methyl-11-azabicyclo[5.3.1]undecane-4-one. The distance corresponding to the donor–acceptor interaction N···C4 is 2.46 Å. Reprinted with permission from Kaftory & Dunitz (1975).

of the continuation of the passage (remembering that it refers to N···B complexes): 'Nonetheless, they (the archaic relationships) are worth noting as they appear to provide a remarkably accurate representation of the relationships between R(NB) (bond distances) and α (bond angles) across a wide range of structures'. Of course, molecules undergo deformations other than those along reaction paths or curves of constant bond order. However, for a given amount of deformation energy, alternative deformations tend to be less pronounced.

6. Crystal packing effects

The variations in structure interpreted in chemical terms in the preceding sections may also be viewed as more or less dramatic manifestations of what is often shrugged off as 'crystal packing effects'. Packing effects are always there. When they are large, they provide interesting chemical information. When they are small, they are not necessarily unimportant. This is because small geometrical variations may be associated with much more important differences of the electron distribution. Methods to measure and compare electron distributions directly are discussed in the article by Coppens (1998) in this Special Issue.

Packing effects on distances and angles are difficult to pin down when comparable gas-phase structures are not available, or when they are similar in magnitude to experimental errors. They can be estimated, however, at least in a statistical sense, by looking at special classes of crystal structures: polymorphic forms, *i.e.* different packings of the same molecular building block; crystal structures containing several crystallographically unrelated copies of the same molecule; or structures built from molecules whose molecular symmetry is higher than that required by the crystal lattice (Martin & Orpen, 1996). In each such case, an average molecular structure and deviations from the mean are calculated. The resulting distribution of deviations and thus its variance σ^2 reflect effects of packing and of experimental error,

$$\sigma^2 = \sigma^2_{\rm packing} + \sigma^2_{\rm exp}. \qquad (9)$$

An estimate of $\sigma^2_{\rm exp}$ is available either directly from the experimental standard uncertainties or indirectly from the crystallographic R factors. The desired quantity, $\sigma_{\rm packing} = \sigma^2 - \sigma^2_{\rm exp}$ has been found to be typically 0.01–0.02 Å for metal–metal and metal–ligand distances and 1–2° for valence angles at metal atoms. Packing effects on torsion angles may be much larger and will be discussed below.

The Cambridge Crystallographic Data Center has compiled extensive tables of 'Typical Interatomic Distances' which also include 'standard deviations' (Allen *et al.*, 1992; Orpen *et al.*, 1992, 1994). The meaning of these quantities is slightly different. An additional term, $\sigma^2_{\rm intra}$, needs to be considered. It accounts for the fact that the same type of bond can occur in different molecules and be affected by *intramolecular* interactions in addition to *intermolecular* ones. The variance has to be interpreted as

$$\sigma^2 = \sigma^2_{\rm intra} + \sigma^2_{\rm packing} + \sigma^2_{\rm exp}. \qquad (10)$$

At present, detailed information on the separate contributions is not available from the tables.

7. Transition-state structure and catalysis

A molecule traveling from reactant to product structure moves up along an energy valley, traverses a pass and descends into another valley of its energy landscape. It dashes through the divide at the top of the pass, the transition state, in a time corresponding to a vibrational period ($\sim 10^{-12}$ s). This is too short to be observed by X-ray diffraction and most other experiments [except for some very fast electron spectroscopies, see Polanyi & Zewail (1995)]. It has been shown above how to map the general geometric characteristics of reaction paths by looking at series of related crystal structures. For a given molecule, the journey starts and ends at specific points on this path, the observable equilibrium structures of reactant and product molecules, and follows the path closely in between. It is less clear where along the path the transition state is located; is it early, *i.e.* close to the reactant structure, or late, *i.e.* close to the product structure? A correlation of structural data alone says nothing about the location and energy of the transition state relative to the reactants and products. Activation energies, which are accessible through kinetic studies, can fill this gap. The general framework for combining structural and kinetic data is as follows: (i) Energy surfaces vary smoothly; there are usually no crevices, precipices and similar discontinuities. (ii) In the neighborhood of equilibrium structures the change in energy along the reaction path can be approximated by Hooke's law, $E = k\Delta d^2/2$; the magnitude of the corresponding force constant k is usually known, at least approximately. (iii) For a given type of reaction, the activation energy for an early transition state is smaller than that for a late one. These three points summarize the well known Marcus theory of chemical reactivity (Marcus, 1968). After a slight modification, needed to exploit the structural differences between similar reactants (Bürgi, 1992; Ferretti *et al.*, 1996), two useful correlations between structure and reactivity can be obtained. They are illustrated here for bond-breaking reactions.

The first result concerns the location of transition states. It can be shown that the lengthening of a scissile bond between ground and transition state is a simple

function of the ratio between the activation energy and the bond stretching force constant,

$$d^{\neq} - d_o = f(\Delta E_o^{\neq}/k). \qquad (11)$$

For a diverse collection of bonds including C—O, B—N, Co^{III}—N, Ni^{II}—N and Ni^{II}—O, $d^{\neq} - d_o$ was found to lie between 0.3 and 0.8 Å. The smaller values are associated with reactions for which solution experiments have indicated an early transition state. Conversely, the larger values are found when solution experiments suggest a late transition state. Thus, a combination of kinetic data, force constants and structure correlation allow the estimation of the point along a reaction path where the transition state can be expected. Such estimates are complementary to calculations by quantum-chemical methods performed on isolated reactants or taking limited account of the environment.

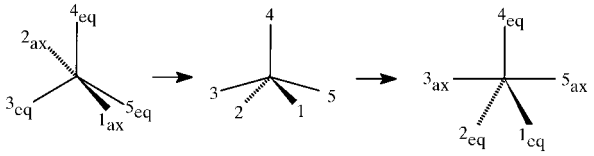
The second result addresses the difference in reactivity of related molecules undergoing the same type of reaction. Consider two reactant molecules with slightly different equilibrium distances of the scissile bond, d_1 and d_2, and different activation energies, $\Delta E_1^{\neq}$ and $\Delta E_2^{\neq}$. For any reasonable representation of the energy surface [points (i) and (ii) above], it is found that

$$\Delta E_2^{\neq} - \Delta E_1^{\neq} = g(\Delta E_1^{\neq} \cdot k)(d_2 - d_1), \qquad (12)$$

where g is a simple function of the product of $\Delta E_1^{\neq}$ and the stretching force constant k. The exact form of g depends somewhat on the analytical representation of the energy surface. The difference in activation energy is proportional to the difference in reactant structure. Values obtained for g are in the range of 300–30 kcal mol^{-1} $Å^{-1}$. The significance of the higher value, typical for a C—O bond, is as follows: an intra- or intermolecular perturbation lengthening a C—O bond by as little as 0.03 Å reduces the activation energy for bond breaking by 9 kcal mol^{-1}. This corresponds to an increase in rate constant by a factor of 10^6–10^7 at room temperature! This value is typical for the acceleration brought about by an enzyme and higher than those for most catalytic antibodies (Mader & Bartlett, 1997). It seems fair to say that understanding small structural differences of intra- and intermolecular origin has implications also for the understanding of reactivity and catalysis.

8. From bond-angle correlations to catalyst design

It is only a small step from distance correlations to generalized structure–structure correlations between any type and number of structural descriptors and, thus, to reaction paths for any kind of chemical process. Muetterties & Guggenberger (1974) have analyzed the bond angles of five-coordinated inorganic molecules. The ligands often assume a trigonal bipyramidal geometry, sometimes a square pyramidal one and other times neither one nor the other. Berry (1960) had recognized earlier that a modest reorganization of the ligands takes a trigonal bipyramid into a square pyramid and, eventually, into another trigonal bipyramid:



Guided by Berry's finding, Muetterties & Guggenberger arranged the coordination geometries found in seven crystal structures into sequential order (Fig. 5). Their plot maps, step by step, the first half of the Berry pathway, or as Hoffmann (1997) likes to view the pictures: 'Cut them out, mount them on cards, flip the cards – you have a movie of the Berry pseudorotation'.

The structures of pentacoordinated phosphorus molecules, arranged in order, also lie along a Berry coordinate. This naturally settles an apparent inconsistency between the equilibrium structure of PF_5 and its fluorine NMR spectrum. The former is trigonal bipyramidal, whereas the latter shows a single signal instead of the expected two, one for the axial fluorines and one for the equatorial ones. Each transit along the Berry pathway exchanges axial and equatorial fluorines. Such exchanges, if they are fast enough, are known to collapse different NMR signals into an averaged one (for a review, see Holmes, 1980).

Much later it was noticed that an obvious extension of the Berry pathway had been missed, one that had implications in catalysis. The story starts in polymer chemistry where it was observed that chiral relatives of Cp_2ZrL_3 complexes catalyze the stereospecific polymerization of α-olefins (Brintzinger *et al.*, 1995; see Fig. 6 for the analogous Cp_2TaH_3). Chemists have reasons to consider the cyclopentadienyl ligand (Cp) as a kind of voluminous atom with electronic properties similar to the much smaller triply bonded ligands $RC\equiv$ or $RN\equiv$. Viewed in this way, Cp_2ZrL_3 is also five-coordinated. Its molecular structure features a bent Cp—Zr—Cp fragment and a coplanar arrangement of the three ligands L with two acute $L—M—L$ angles which can be as small as ~50°. The structure does not resemble any of the forms in Fig. 5, and yet it fits logically into the Berry pathway (Fig. 6): start at the square pyramid, push the axial ligands towards the trigonal bipyramid and continue pushing until, eventually, the structure at the far right is reached. It is called an 'edge-bridged tetrahedron', since the two Cp's and the two outer ligands L form a distorted tetrahedron and the middle ligand L bridges one of its edges. Brintzinger has proposed that the stereospecificity of the catalyst is related to the crowding of the three ligands L. It was therefore of interest to understand the electronic and steric factors responsible for edge-bridged tetrahedral geometries and, on this

basis, to identify other chemical compounds with similar structures and, hopefully, catalytic activities (Ward *et al.*, 1997).

A molecular-orbital analysis showed that the trigonal bipyramid is an unstable structure for d^0 metal centers coordinated to five σ-donating ligands like alkyl groups. It distorts to *either* a square pyramidal *or* an edge-bridged tetrahedral structure. A way to favor one over the other is to play with the equatorial ligands. Replacing the single bond of a σ-donating ligand by the double or triple bond of a σ- *and* π-donating one pushes the axial ligands away from the multiple bond towards a square-pyramidal structure. Introducing two ligands with multiple bonds pushes the axial ligands away from both, *i.e.* towards the desired edge-bridged tetrahedral structure. Examples of multiply bonded ligands are, apart from Cp, $RR'C{=}$, $RC{\equiv}$, $RN{=}$ and $RN{\equiv}$. The larger size of Cp ligands as compared to $RC{\equiv}$ or $RN{\equiv}$ ligands provides an additional steric driving force in the direction of the edge-bridged tetrahedron. This simple concept was confirmed by many structures found in the Cambridge Structural Database. Now the problem is back on the chemist's drawing board. The goal is to design and synthesize new molecules with a distinctly edge-bridged tetrahedral geometry, to characterize their structures and to test their catalytic properties.

9. Molecular structure and magnetic properties

Many organic radicals and complexes of first-row transition metals behave like tiny magnets, because the spins of some of their electrons are not paired up. In the crystalline state, these molecules interact in one of two ways. The magnetic moments of the individual spins either add up or they cancel each other, corresponding to ferromagnetic or antiferromagnetic coupling, respectively. If the couplings can be controlled, materials with custom-made macroscopic magnetic properties can be tailored (Kahn, 1993). According to Hoffmann's (1997) program, such control has a structural basis.

A simple example is provided by binuclear copper(II) complexes (Fig. 7; Crawford *et al.*, 1976). The two metal ions have one unpaired electron each and are coupled through two bridging ligand atoms. The two unpaired electrons show parallel or antiparallel spins depending on the Cu—O—Cu angles α: if α is smaller than 98°, the spins are antiparallel, if it is larger they are parallel, but barely. The energy difference J between the two states, the singlet–triplet splitting energy, is small, a few hundred wavenumbers. The interesting point, however, is that the energy J correlates with the bond angle α according to

$$J\ (\mathrm{cm}^{-1}) = -74\alpha\,(\mathrm{deg}) + 7270. \qquad (13)$$

Contrary to most other topics discussed so far, it is necessary here to understand the interdependence of geometric and electronic structure in considerable detail. Apart from molecular orbitals, consideration of electron–electron repulsion, configuration interaction and correlation energy is important, topics which are beyond the scope of this paper. Although this magneto-structural correlation was found some 20 years ago, it is rarely cited in the crystallographic literature. More recently, an analogous correlation has been established for the family of linear and bent oxo-bridged transition-metal dimers, including both homo- and heteronuclear species (Weihe & Güdel, 1998). Molecular structure–property correlations form a basis for investigating

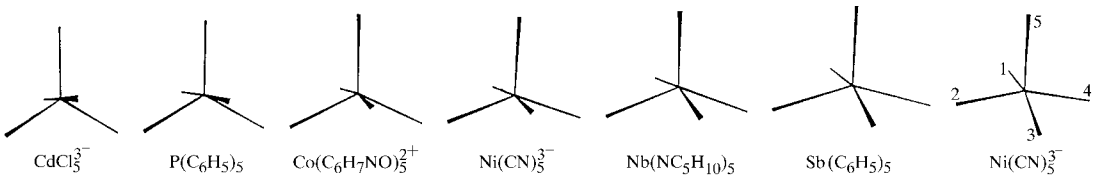


Fig. 5. Seven five-coordinated structures tracing the Berry pathway from trigonal bipyramidal to square-pyramidal coordination geometry. Reprinted with permission from Muetterties & Guggenberger (1974). Copyright (1974) American Chemical Society.

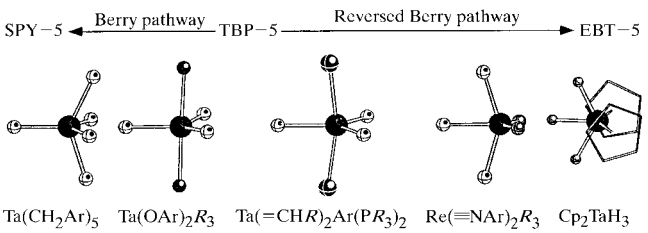


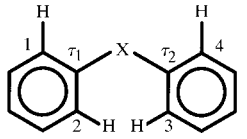
Fig. 6. Mapping of the reversed Berry pathway with (pseudo-)five-coordinated d^0 complexes incorporating two π-donor and three σ-donor ligands. (SPY: square pyramidal; TBP: trigonal bipyramidal; EBT: edge-bridged tetrahedral. Reprinted with permission from Ward *et al.* (1997). Copyright (1997) American Chemical Society.

coordination polymers in which *all* metal ions in the crystal are coupled by ligands. Coordination polymers have a potential of being interesting magnetic materials, if they can be tailor-made (Coronado *et al.*, 1996).

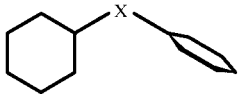
10. Conformational libraries, molecular machinery and a movie of protein motion

The world of macromolecules displays a diversity of shapes: the tangles of polyethylene, the winding stairs of double-helical DNA, the α-helices and twisted sheets in proteins. This diversity of shape becomes possible whenever two parts of a molecule are connected by a single bond and can be rotated, one with respect to the other, at practically no cost in energy. Molecules with several such bonds are therefore an ideal playground for structure–structure correlations and studies of packing effects.

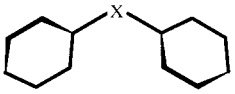
For a simple example, consider molecules with two geminal phenyl groups. The planar structure (I) has two hydrogen atoms which are too close to each other.



(I)



(III)



(II)

$X = CH_2, C=CR_2, C=O,$
NH, O, S, SO_2, etc

Ways of avoiding this are shown in (II) and (III). Are there others? And how do the phenyl torsion angles depend on the nature of the bridging group X? Such questions are easily answered with the help of the Cambridge Structural Database (Allen & Kennard, 1993). The torsion angles τ_1 and τ_2 for the molecules or molecular fragments of interest are plotted as a function of each other. The points in the scatterplot for $X = O$ are not uniformly distributed (Fig. 8). There are no points corresponding to the planar molecules (I) with coordinates (0°, 0°). There are a few points near (0°, 90°) which correspond to the situation (II) with the two rings perpendicular. There are many points at (45°, 45°) representative of conformation (III), and there is almost a continuum of points between the two groups: as one torsion angle increases the other one decreases. This indicates that a continuous rotation of the two rings in opposite directions can occur and does not lead to steric overcrowding as in (I) (Rappoport & Biali, 1997). Conformational scatter plots for the other groups X and for molecules with substituents in positions 1 to 4 may look quite different (Klebe, 1994). Libraries of such scatterplots are no substitute for computational studies, but serve as a guide for designing molecules which are not only isosteric to known active substances but also have the desired conformational flexibility.

With some stretch of the imagination, the two phenyl groups of (I) can be considered as meshed cogwheels, somewhat special cogwheels with only two cogs and two notches. Nevertheless, the analogy has inspired chemists to synthesize molecules with a more direct morphological relationship to macroscopic gears: two triptycenes coupled through O, CH_2, C═O or CHOH groups (Fig. 9). The conformational scatterplot is analogous to Fig. 8. The barrier to cogwheeling motion is much too small to be observed by NMR; it was calculated to be less than 1 kcal mol^{-1}! By comparison, the barrier to gear slippage (...krrtch) was calculated at 30 kcal mol^{-1}, in fair agreement with an experimentally determined value. The extension to multiple gear systems is obvious. The possibility that macromolecular chains of securely meshed chemical gears might be used to transmit information over long distances on the atomic scale has been mentioned by Mislow (1989) in a critical comparison of chemical and mechanical gears, well worth reading.

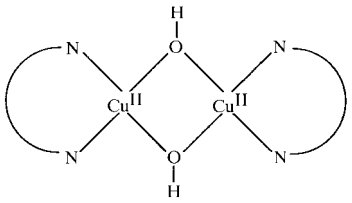


Fig. 7. Schematic representation of binuclear Cu^{II} complexes with two hydroxyl groups as bridging ligands. The Cu—O—Cu angle varies between ~95 and ~104° depending on the nature of the bidentate nitrogen ligand N—N.

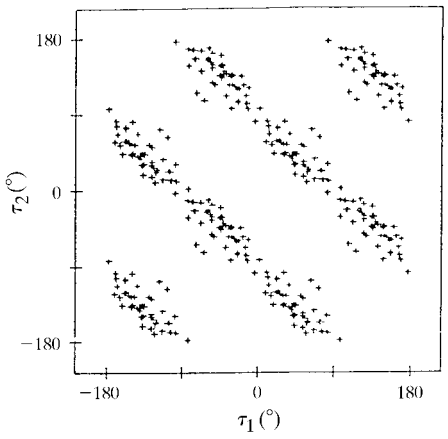


Fig. 8. Conformational map of the two phenyl groups in diphenyl ether derivatives. (I; $X = O$, with hydrogen atoms in the *ortho* positions 1–4). Reprinted with permission from Klebe (1994). Copyright (1994) Elsevier Science.

To conclude this section, we return to protein structure and motion. There are 17 crystal structures of nucleoside monophosphate kinases. They show large conformational changes upon binding substrates. Aligning the main polypeptide chains of the 17 structures revealed two domains in various conformational states. The states span the range between a 'closed' and a less well defined 'open' conformation. The structures were visually sorted into a more or less evenly spaced series of domain states and supplemented somewhat by inserting interpolated structures between the observed ones. The ordered experimental structures were used as still pictures for a movie outlining the closing motions of the enzyme when the substrates bind. The paper of Vonrhein *et al.* (1995) provides two series of 40 pictures, to be copied, stapled together and flipped through. The resulting animation of protein motion can be viewed directly and more easily on the World Wide Web (http://bio5.chemie.uni-freiburg.de/local/ak_movie/ak_movie.html).

11. The nightmare of *n* dimensions

The gear system and the protein example presented in the previous section illustrate two different ways of comparing molecular structures. For the kinases, the positions of corresponding atoms in the various protein structures are matched as closely as possible; the proteins are aligned in their physical three-dimensional space and then sorted 'visually'. For the gear system, two torsion angles are calculated from the positional coordinates first and plotted, one as a function of the other, in a so-called conformational space. Each point in this space represents the conformation of one molecule. The scatterplot in Fig. 8 sorts the conformations without resorting to 'visual' comparison. If two molecules are represented by the same point, they are equal as far as conformation is concerned; if not, their conformational difference is uniquely defined by the length and direction of the distance vector in conformational space. Conformational scatterplots are therefore an objective way of comparing conformations. If one had applied this method to the kinase structures, there would have been a problem. The conformations of such molecules are characterized by a large number of torsion angles. The conformational space has a correspondingly large number of dimensions, one for every torsion angle. The distribution of scatter points can no longer be grasped visually, but must be analyzed painstakingly in terms of the vectors between representative points. This is the nightmare of *n* dimensions. The problem is a general one, it occurs for any structure correlation between more than three structural parameters and thus for any configuration space with more than three dimensions.

There is an additional complication. In a protein, each amino-acid residue and thus each atom can be uniquely labeled. As a result, the atoms of two different conformations of the same protein have a unique one-to-one correspondence which is the basis for the alignment procedure. In diphenyl ether, the labeling of atoms is less straightforward. Apart from the numbering shown in (I), (1234), there are seven equally valid alternatives: (1243), (2134), (2143), (3412), (3421) (4312) and (4321). The values assigned to the torsion angles τ_1 and τ_2 depend on the numbering chosen. This apparent ambiguity is a consequence of the permutational symmetry of the molecular graph or line formula. In the analysis, all eight permutations must be taken into account [plus eight permutation inversions; for a review, see Dunitz & Bürgi (1994)]. The symmetric appearance of the scatterplot, Fig. 8, is an expression of this. Symmetry analysis is not a mathematical gewgaw. It is a prerequisite for avoiding arbitrary atomic numberings, for obtaining unique scatterplots and for unraveling systematically and reproducibly the information hidden in the *n*-dimensional thicket.

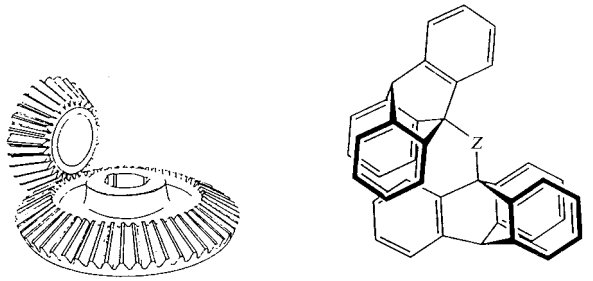


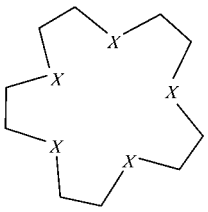
Fig. 9. Mechanical and molecular gear systems: the gearing motion of bistriptycylmethane ($Z = CH_2$) is practically unhindered, with an estimated activation barrier of < 1 kcal mol^{-1}. Reprinted with permission from Mislow (1989). Copyright (1989) Data Trace Publishing Company.

12. Principal components, cluster analysis and neural networks

There are three main methods to analyse distributions of data points in high-dimensional configuration spaces. First, one tries to reduce the number of dimensions to the relevant ones. In the case of diphenyl ethers, we could have plotted the sum of torsion angles *versus* their differences. We would have found that the sums are all close to -270, -90, 90 or 270°, whereas the differences vary continuously, implying that the scatterpoints concentrate along lines. Indeed, the difference between τ_1 and τ_2 is an adequate, albeit approximate, one-dimensional descriptor of the conformational diversity of diphenyl ethers. The simple example can be generalized into any number of dimensions. The corresponding mathematical procedure is called principal component analysis and is available in most program packages for statistical analysis (for a review, see Taylor & Allen, 1994).

Correlations are not always linear (see Fig. 2) and the number of the most important principal components may still exceed three. In such cases, it is advantageous to cluster points which are close in configuration space into groups. There are various criteria for measuring the similarity of representative points and for deciding whether or not to merge them in a cluster. Nevertheless, cluster analysis is a powerful tool for analyzing scatterplots in many dimensions. Algorithms for cluster analysis are also part of statistical program packages (for a review, see Taylor & Allen, 1994).

Conformational analysis of large ring structures provides an example of the potential of cluster analysis (Raithby *et al.*, 1997). Consider the 15-membered ring of 1,4,7,10,13-pentaoxacyclopentadecane:



X = O, S

There are 15 torsion angles, three complicated nonlinear dependencies between them and 20 symmetry-equivalent ways of labeling the atoms. And there are the sulfur and mixed analogs. The problem is to find the prototypal conformations of these molecules and their relative energies; not a trivial task by most standards. Symmetry-modified clustering of 130 unique sets of observed torsion angles has revealed more than 15 stable conformations whose conformational energies were estimated with the help of a force-field model; the seven most stable conformations are quite different structurally, but their energies differ by less than 3 kcal mol^{-1}. This type of study is now fairly automated. It quickly provides libraries of molecular structures which may be used as starting points for a rational design of new compounds and for predicting their properties.

In passing, we note that such libraries are used routinely in macromolecular structure determination. Because of limited experimental resolution, the electron density of proteins, nucleic acids and their complexes is often blurred and does not show resolved atoms. By fitting standardized peptide groups, amino-acid side chains, purine and pyrimidine bases, sugar and phosphate fragments into the poorly resolved electron density function, one can retrieve the desired three-dimensional information. The necessary libraries or dictionaries of fragment structures are extracted from accurate crystal structure determinations of individual amino-acid molecules, nucleobases, sugars, their simple derivatives and of molecules built from a limited number of these fragments (for a recent example, see Gelbin *et al.*, 1996).

The third method to analyze multidimensional datasets is based on computer simulations of neural networks. A general introduction and applications to diverse chemical problems can be found elsewhere (Gasteiger & Zupan, 1993). A specific type of neural network, the so-called self-organizing mapping networks, project data from a space of high dimension to a space of significantly lower dimension, *e.g.* a plane. The mappings maintain the topological or neighborhood relations of the primary data. As far as structure correlation is concerned, there seems to be only one such application. It is part of a research program in stereospecific catalysis which involves: (i) the synthesis of a library of closely related metal complexes; (ii) measuring, understanding and modeling their shapes; and (iii) correlating shape to results of a catalytic reaction (Beyreuther *et al.*, 1996). A set of 82 *tripod* $Co^{II}L_2$ and *tripod* $Co^{II}L_3$ complexes was retrieved from the Cambridge Structural Database [*tripod* = $CH_3C(CH_2P\{C_6H_5\}_2)_3$, see Fig. 10]. The twisting of the $C(CH_2P)_3Co$ cage and the conformations of the six phenyl rings were analyzed in a seven-dimensional space. The neural network found a path of correlated changes in all seven observables.

It is known that neural networks are the most efficient method to solve the problem of the traveling salesman. The challenge is to find the shortest route between a number of cities in such a way that it passes through every city once. In the context of structure correlation, this translates to the following procedure: First, a 'population density' map must be obtained from the available structural data. This involves constructing the distribution of data points in the appropriate configuration space with proper account of permutation symmetry. Second, the major 'cities', their 'suburbs' and the 'villages' in between need to be located with the help of cluster analysis. The 'cities' are the densely populated clusters, the 'suburbs' are their extensions, and the 'villages' are the smaller clusters between the major ones. Finally, the paths between major clusters are traced through the populated regions with the help of neural networks. To our knowledge, such a combined analysis has not yet been attempted, but may well help to find linear and nonlinear correlations, the winding

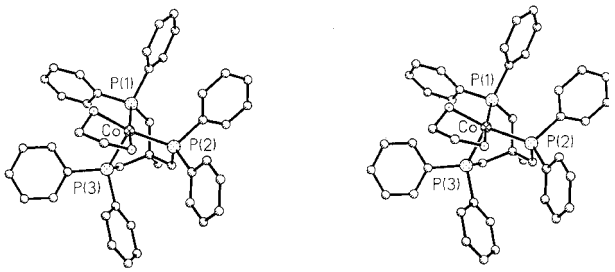


Fig. 10. Stereoscopic view of $CH_3C(CH_2PPh_2)_3Co^{II}(NH_2CH_2CH_2NH_2)$, an example of a *tripod* ML_2 complex.

roads in multidimensional configuration space. In any case, the full potential of neural networks in structure correlation studies remains to be explored more fully.

13. Further reading and conclusions

The computer-readable crystallographic databases and the associated software for retrieval, analysis and visualization of information have become the most essential tool for studies in structure correlation [see the article in this Special Issue by Allen (1998)]. The ideas and methods of structure correlation also apply to the analysis of noncovalent interactions, which govern the packing of molecules in crystals, or the binding of substrates to proteins [see the articles in this Special Issue by Nangia & Desiraju (1998) and Saenger & Steiner (1998)]. Increasingly, the results of structure correlation studies are made available through knowledge bases (Allen, 1998). Because of limitations of space, most references in this article are to recent surveys. Citations of original papers are found in a number of specialized reviews (Ferretti *et al.*, 1996; Bürgi & Dunitz, 1994; Orpen, 1993; Bürgi, 1992).

References

Allen, F. H. (1998). *Acta Cryst.* A**54**, 758–771.

Allen, F. H. & Kennard, O. (1993). *Chem. Design Autom. News*, **8**, 31–37.

Allen, F. H., Kennard, O., Watson, D. G., Brammer, L., Orpen, A. G. & Taylor, R. (1992). *International Tables for Crystallography*, Vol. C, edited by A. J. C. Wilson, pp. 685–706. Dordrecht: Kluwer Academic Publishers.

Bent, H. A. (1968). *Chem. Rev.* **68**, 587–648.

Berry, R. S. (1960). *J. Chem. Phys.* **32**, 933–938.

Beyreuther, S., Hunger, J., Huttner, G., Mann, S. & Zsolnai, L. (1996). *Chem. Ber.* **129**, 745–757.

Brintzinger, H. H., Fischer, D., Mühlhaupt, R., Rieger, B. & Waymouth, R. M. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 1134.

Brown, I. D. (1994). *Structure Correlation*, edited by H. B. Bürgi & J. D. Dunitz, pp. 405–429. Weinheim: Verlag Chemie.

Brown, I. D. & Altermatt, D. (1985). *Acta Cryst.* B**41**, 244–247.

Bürgi, H. B. (1975). *Angew. Chem. Int. Ed. Engl.* **14**, 460–473.

Bürgi, H. B. (1992). In *Perspectives in Coordination Chemistry*, edited by A. F. Williams, C. Floriani, A. E. Merbach, pp. 1–29. Basel: Verlag Helvetica Acta.

Bürgi, H. B. & Dunitz, J. D. (1987). *J. Am. Chem. Soc.* **109**, 2924–2926.

Bürgi, H. B. & Dunitz, J. D. (1994). *Structure Correlation*, Vols. 1, 2. Weinheim: Verlag Chemie.

Bürgi, H. B. & Shklover, V. (1994). *Structure Correlation*, edited by H. B. Bürgi & J. D. Dunitz, pp. 303–335. Weinheim: Verlag Chemie.

Cieplak, A. S. (1994). *Structure Correlation*, edited by H. B. Bürgi & J. D. Dunitz, pp. 206–302. Weinheim: Verlag Chemie.

Coppens, P. (1998). *Acta Cryst.* A**54**, 779–788.

Coronado, E., Delhaes, P., Gatteschi, D. & Miller, J. S. (1996). *Molecular Magnetism: from Molecular Assemblies to the Devices*, Vol. 321. NATO ASI Series. Dordrecht: Kluwer Academic Publishers.

Crawford, V. H., Richardson, H. W., Watson, J. R., Hodgson, D. J. & Hatfield, W. E. (1976). *Inorg. Chem.* pp. 2107–2110.

Dunitz, J. D. & Bürgi, H. B. (1994). *Structure Correlation*, edited by H. B. Bürgi & J. D. Dunitz, pp. 23–70. Weinheim: Verlag Chemie.

Editorial Preface (1948). *Acta Cryst.* **1**, 1–2.

Ernst, R. D., Freeman, J. W., Stahl, L., Wilson, D. R., Arif, A. M., Nuber, B. & Ziegler, M. L. (1995). *J. Am. Chem. Soc.* **117**, 5075–5081.

Ferretti, V., Gilli, P., Bertolasi, V. & Gilli, G. (1996). *Crystallogr. Rev.* **5**, 3–98.

Gasteiger, J. & Zupan, J. (1993). *Neural Networks in Chemistry.* Weinheim: Verlag Chemie.

Gelbin, A., Schneider, B., Clowney, A., Hsieh, S.-H., Olson, W. K. & Berman, H. M. (1996). *J. Am. Chem. Soc.* **118**, 519–529.

Hart, M. & Berman, L. (1998). *Acta Cryst.* A**54**, 850–858.

Helliwell, J. R. (1998). *Acta Cryst.* A**54**, 738–749.

Hoffmann, R. (1997). *Am. Sci.* **86**, 15–17.

Holmes, K. C. (1998). *Acta Cryst.* A**54**, 789–797.

Holmes, R. R. (1980). *Pentacoordinated Phosphorus. Am. Chem. Soc. Monogr.* No. 175/176. Washington, DC: American Chemical Society.

Kaftory, M. & Dunitz, J. D. (1975). *Acta Cryst.* B**31**, 2914–2916.

Kahn, O. (1993). *Molecular Magnetism.* Weinheim: Verlag Chemie.

Klebe, G. (1994). *J. Mol. Struct. (Theochem.)*, **308**, 53–89.

Leonard, N. J. (1997). *Tetrahedron*, **53**, 2323–2355.

Leopold, K. R., Canagaratna, M. & Phillips, J. A. (1997). *Acc. Chem. Res.* **30**, 57–64.

Mader, M. M. & Bartlett, P. A. (1997). *Chem. Rev.* **97**, 1281–1301.

Marcus, R. A. (1968). *J. Phys. Chem.* **72**, 891–899.

Martin, A. & Orpen, A. G. (1996). *J. Am. Chem. Soc.* **118**, 1464–1470.

Mislow, K. (1989). *Chemtracts Org. Chem.* **2,** 151–174.

Muetterties, E. L. & Guggenberger, L. J. (1974). *J. Am. Chem. Soc.* **96**, 1748–1756.

Nakashima, T. T. & Maciel, G. E. (1972). *Org. Magn. Reson.* **4**, 321–326.

Nangia, A. & Desiraju, G. R. (1998). *Acta Cryst.* A**54**, 934–944.

Orpen, A. G., Brammer, L., Allen, F. H., Kennard, O., Watson, D. G. & Taylor, R. (1992). *International Tables for Crystallography*, Vol. C, edited by A. J. C. Wilson, pp. 707–791. Dordrecht: Kluwer Academic Publishers.

Orpen, A. G., Brammer, L., Allen, F. H., Kennard, O., Watson, D. G. & Taylor, R. (1994). *Structure Correlation*, edited by H. B. Bürgi & J. D. Dunitz, Appendix A, pp. 751–858. Weinheim: Verlag Chemie.

Orpen, G. A. (1993). *Chem. Soc. Rev.* pp. 91–197.

Pauling, L. (1960). *The Nature of the Chemical Bond*, 3rd ed. Ithaca, New York: Cornell University Press.

Polanyi, J. C. & Zewail, A. H. (1995). *Acc. Chem. Res.* **28**, 119–132.

Raithby, P. R., Shields, G. P. & Allen, F. H. (1997). *Acta Cryst.* A**53**, 476–489.

Rappoport, Z. & Biali, S. E. (1997). *Acc. Chem. Res.* **30**, 307–314.

Reeve, S. W., Burns, W. A., Lovas, F. J., Suenram, R. D. & Leopold, K. R. (1993). *J. Phys. Chem.* **97**, 10630–10637.
Rossmann, M. (1998). *Acta Cryst.* A**54**, 716–728.
Saenger, W. & Steiner, T. (1998). *Acta Cryst.* A**54**, 798–805
Taylor, R. & Allen, F. H. (1994). *Structure Correlation*, edited by H. B. Bürgi & J. D. Dunitz, pp. 111–161. Weinheim: Verlag Chemie.
Vonrhein, C., Schlauderer, G. J. & Schulz, G. E. (1995). *Structure*, **3**, 483–490.
Ward, T. R., Bürgi, H. B., Gilardoni, F. & Weber, J. (1997). *J. Am. Chem. Soc.* **119**, 11974–11985.
Weihe, H. & Güdel, H. U. (1998). *J. Am. Chem. Soc.* **120**, 2870–2879.
Yonath, A., Harms, J., Hansen, H. A. S., Bashan, A., Schlünzen, F., Levin, I., Koelln, I., Tocilj, A., Agmon, I., Peretz, M., Bartels, H., Bennett, W. S., Krumbholz, S., Janell, D., Weinstein, S., Auerbach, T., Avila, H., Piolleti, M., Morlang, S. & Franceschi, F. (1998). *Acta Cryst.* A**54**, 945–955.