STEREOCHEMISTRY OF REACTION PATHS AT CARBONYL CENTRES

H. B. Bürgi, J. D. Dunitz, J. M. Lehn and G. Wipff

Laboratories of Inorganic and Organic Chemistry, Swiss Federal Institute of Technology (E.T.H.), 8006 Zürich, Switzerland

and

Institute of Chemistry, University Louis Pasteur, 4 rue Blaise Pascal 67008 Strasbourg, France

Abstract—Results of recent experimental and theoretical studies of nucleophilic addition to carbonyl groups are described. The reaction paths found by different methods for different nucleophiles show some striking similarities that appear to be characteristic for the reaction type.

For the last hundred years the tetrahedron has been recognized as the typical building-block of organic molecules. The structural transformations of a tetrahedral centre, i.e. the chemical reactions it undergoes, can be broken down into one or other of two primary processes.

The sequence of primary processes for a vast number of chemical reactions of organic and inorganic tetrahedral molecules has been inferred from mechanistic arguments, and the energy changes that occur during these steps have been followed in many cases from kinetic data. With carbon as central atom the above classification has led to two main families of reactions. One is that of the nucleophilic substitution (SN) reactions, an example of which is the Walden inversion.
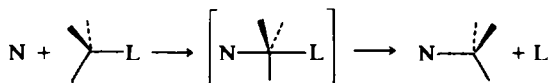
Addition of the nucleophile N is an energetically uphill process, leading to a five-coordinate trigonal bipyramidal transition-state, which corresponds to a maximum in energy and is not directly observable. The other family of reactions, nucleophilic or electrophilic addition/elimination reactions, may be uphill or downhill depending on the nature of the ligands and on many other factors. The structure of the corresponding transition states has hardly been investigated. The trigonal-tetrahedral interconversion at a carbonyl centre is a special but very important case of the ligand removal reaction in reverse.

For many classes of reaction it has been possible to construct those well known reaction coordinate *vs* energy diagrams that adorn so many publications. On the other hand, the details of the geometric changes that occur along the reaction coordinate (actually a path in a many-dimensional space) are usually passed over because of lack of knowledge.

It is indeed very difficult to obtain this kind of information because every molecule takes a different trajectory depending on the initial magnitude and direction of its momentum. A description of these trajectories requires a knowledge of the underlying energy landscape or hypersurface (Born–Oppenheimer surface). The paths that connect the energy minima on this hypersurface *via* its passes are called reaction paths. The question may now be asked: How do the bonds and angles in a molecule change as it moves along the reaction path? We have been interested in two approaches to this kind of problem. One approach uses experimental data from crystal structure analyses, the other is based on quantum mechanical calculations for simple model systems. The basic idea behind the crystal structure approach is to search for correlations between certain selected parameters describing the geometry of some sub-system (say, a tetrahedral centre) frozen in a variety of crystal or molecular environments. Each kind of environment can be regarded as a perturbation that acts on the sub-system, which will slightly modify its structure so as to minimize the potential energy of the crystal as a whole. Regardless of the nature of the perturbations, which are rarely if ever understood in detail, these structural modifications should

occur along local energy valleys in the parameter-space defining the sub-system. Accordingly, if some correlation can be found among the experimentally observed parameters then the correlation function can be taken as describing a minimum energy path in the Born–Oppenheimer surface of the sub-system in question. If the pattern of parameter changes can be identified in a general sense with that expected to occur in the course of a chemical reaction, then the path found from such experimental data should be a fair approximation to the corresponding reaction path. We assume as working hypothesis that the observed set of static structures, arranged in the right sequence, provides a picture of the changes that occur along the reaction pathway open to the subsystem. The analogy with a set of static pictures arranged and viewed as in a cinematic film is obvious. These ideas have been applied to derive the minimum energy pathways for the two primary types or reactions: ligand addition[1] ($XCdS_3 + Y \rightarrow$ $[X \ldots CdS_3 \ldots Y] \rightarrow X + CdS_3Y$) and ligand removal[2] ($SO_4^{-2} \rightarrow SO_3 + O^{-2}$, $AlCl_4^- \rightarrow AlCl_3 + Cl^-$) at noncarbon tetrahedral centres as well as to nucleophilic addition/elimination reactions at carbonyl.[3] The results obtained so far for these systems suggest that important features of the minimum energy paths are largely independent of the nature of the central atom and of the ligands but are characteristic for the reaction type.

The second approach, from quantum mechanical calculations, consists of computing the energy of the isolated reacting system for a sufficiently large number of nuclear arrangements. The most general procedure would be to calculate the complete energy surface, but for reasons of economy the calculations are usually restricted to those parts that are believed, on the basis of previous chemical knowledge or intuition, to be most relevant, usually only to a section through the surface along the assumed reaction coordinate and to a few neighbouring points. In general, Hartree–Fock SCF–LCAO–MO type calculations are reasonably well adapted to such problems since changes in electron correlation energy are expected to be small compared with the energy change in the course of a reaction, involving closed shell reactants.[4]

These two approaches, with their own intrinsic strengths and weaknesses, complement and also reinforce each other very nicely. The reaction path derived from crystal structure correlations and other chemical information can serve as a pointer to the parts of the energy surface to be covered by the computations. The computations, in turn, provide the energy variation along the reaction path, which cannot be obtained directly from the structural data alone, and may indicate how and to what extent interactions between sub-system and crystal environment may affect the experimental correlations. The computations treat the reacting system

as an isolated species *in vacuo*, whereas the strong interactions between the system and its crystal environment may be roughly analogous to those that occur between a reacting system and its environment in solution (solvent effects).

Nucleophilic addition to a carbonyl group

The carbonyl group plays a central role in organic chemistry as well as in biochemistry by the number of important reactions it may undergo, by its potential utilization for generating stereochemistry and molecular asymmetry, and by its function as a building block in the construction of natural and synthetic molecules and biopolymers.

The primary function of the trigonal carbonyl carbon is to act as an acceptor for electron-rich, nucleophilic reagents, which add to the carbon atom. The resulting tetrahedral species may, in principle, be a stable product, an intermediate, or a transition state in a reaction. In the last two cases it will react further, generally undergoing an elimination, i.e. addition in reverse, to yield a new trigonal centre. The corresponding reaction paths may present a rich ensemble of stereochemical features, both structural and stereoelectronic. The generally accepted mechanism of this type of reaction involves the following typical steps:

(1) approach of the nucleophile towards the carbonyl group;
(2) passage through a first transition state;
(3) formation of a tetrahedral intermediate;
(4) passage through a second transition state;
(5) formation of the products.

Depending on the nature of the environment, various proton transfer processes may take place along these steps (acid or base catalysis).

The mechanism and kinetics of this reaction have been the subject of many investigations, using a great variety of nucleophiles and carbonyl-substrates of chemical or biochemical significance; much of the recent work has been prompted by interest in the catalytic power and specificity of proteolytic enzymes. The wealth of data accumulated by such studies on the various reaction steps, on structural effects and on enzymic function have been reviewed elsewhere[5-8] and cannot even be summarized here. In short, they agree with the overall mechanism described above, and, most important, definite evidence for the formation of tetrahedral intermediates has been obtained by direct detection and by interpretation of kinetic data. Very recently, some details of the free energy variation along the reaction coordinate have been evaluated for several carbonyl hydration processes,[9] and a preliminary description of the reaction path has been obtained from crystal structure correlations for the addition of a tertiary amine to a carbonyl group.[3]

Here we take up several aspects of the problem and discuss them in the light of recent developments, combining, wherever possible, results from crystal structure correlations and from *ab initio* calculations for simple model systems.

The path of approach of the nucleophile

Structure correlations. It has been shown[3] from structural correlations of molecules containing a tertiary amino group and a carbonyl group with which it can interact, that approach of the nucleophilic N atom towards the electrophilic C atom is accompanied by a displacement (Δ) of the C atom out of the plane defined by its three bonded atoms (two substituents R and R', and the carbonyl O atom) towards the nucleophilic centre (Figs 1, 2). The out-of-plane displacement Δ increases as the N . . . C distance d_1 decreases, following a logarithmic relationship

$$d_1 = -1\cdot701 \log_{10} \Delta + 0\cdot867 \text{ Å} \qquad (1)$$

For $d_1 < 3$ Å the approach path of the nucleophile lies in a plane bisecting the RCR' angle and is virtually along a straight line inclined at an angle of about 107° to the C–O bond.

Equation (1) can be reformulated as

$$d_1 = -1\cdot701 \log_{10} n + 1\cdot479 \text{ Å} \qquad (2)$$

where $n = \Delta/\Delta_{max}$ and Δ_{max} is the out-of-plane displacement calculated from (1) for $d_1 = 1\cdot479$ Å, the standard C–N bond length.[10] Eq (2) has the same form as Pauling's formula[11] relating interatomic distance to bond number, and the quantity n can be identified with Pauling's bond number. At the same time the C–O bond length d_2 increases as d_1 decreases, roughly according to the relationship

$$d_2 = -0\cdot71 \log_{10} (2 - n) + 1\cdot426 \text{ Å} \qquad (3)$$

which is, however, much less well defined by the experimental data; here the multiplicative constant is taken from Pauling's formula relating C–C bond length and bond number,[11] and the additive constant is the standard C–O single bond distance.[10] Eqs (2) and (3) would imply that the sum of the C–O and C–N bond numbers equals two at all stages of the
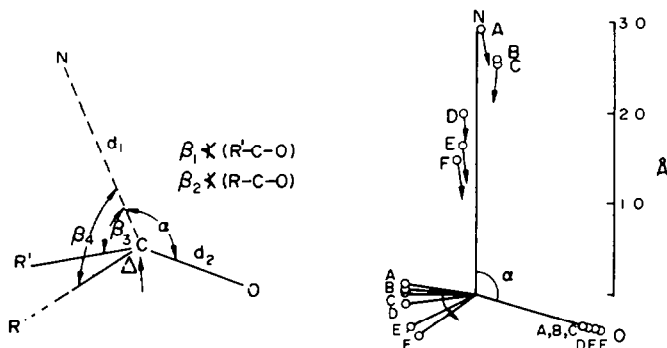

Fig 1. *Left*: definitions of structural variables. *Right*: reaction path for addition of amine to carbonyl showing nitrogen (top; arrows indicate estimated direction of lone pair), carbonyl oxygen (bottom right) and inclination of RCR' plane.
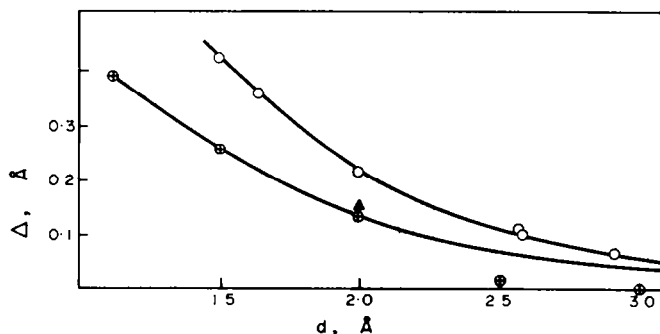

Fig 2. Plots of Δ vs d_1. Empty circles—experimental points for amine addition; hatched circles—theoretical points for hydride addition; triangle—theoretical point for amine addition. The smooth curves are those obtained from equations (1) and (4).

addition reaction. Analogous empirical relations involving the constancy of the appropriate sum of bond numbers have also been found to hold for other reactions in which bonds are broken or formed.[12] Although it is difficult to give a rigorous definition of bond number in terms of current theoretical models of chemical bonding,† it appears to be a most useful quantity for interpreting the empirical correlations that exist among the geometrical parameters describing molecular systems.

Very similar deformations are observed when nucleophilic oxygen atoms approach the C atoms of carbonyl groups.[12] The relevant data from crystal structure analyses are much more extensive than in the N . . . C=O case but they are also much more hetereogeneous in that the nucleophilic O atom may belong to a variety of chemical types (ether, carbonyl, carboxylic acid, ester or anhydride, nitro group, etc.) and similarly for the electrophilic C atom. Nevertheless, the data here also show a clear trend for the out-of-plane displacement Δ at the carbonyl C atom to increase as the nucleophile approaches. For a given O . . . C=O approach distance d_1 the mean displacement is only about a third of that found in the N . . . C=O examples, suggesting that the O . . . C interaction is about ten times weaker than the N . . . C interaction, in agreement with expectation that alcoholic or ketonic oxygen is a much weaker nucleophile than amino nitrogen.

Theoretical results

A set of SCF–LCGO–MO calculations has been carried out for various nuclear arrangements of the simple model system $CH_2O + H^- \rightarrow CH_3O^-$, corresponding to nucleophilic addition of hydride ion to formaldehyde to yield methanolate ion.[13] The reaction path obtained shows some striking similarities to the one derived from structural correlations for amine addition. For d (H⁻ . . . C) < 2 Å, the out-of-plane displacement Δ at the carbon atom is given by a logarithmic relationship (4) similar to (1):

$$d(H^- \ldots . C) = -1.805 \log_{10} \Delta + 0.415 \text{ Å} \quad (4)$$

The C–O distance also increases as the H⁻ ion approaches, from 1.203 Å (in formaldehyde) to 1.405 Å (in methanolate ion).

The experimental and theoretical curves are shown in Fig 2 and one may well be astonished at the similarity between them. The upper curve refers to addition of an amino group and was obtained from experimental data on interatomic distances in crystals. The lower one refers to addition of hydride ion and was obtained from the nonempirical calculations for the isolated system *in vacuo*, as

described above. The differences between the two curves largely reflect the difference between the N–C and H–C bond lengths, 1.479 and 1.12 Å, respectively, in the two tetrahedral products. Until evidence to the contrary is obtained it seems a reasonable extrapolation to assume that a general similarity exists between corresponding curves describing addition of any nucleophile to a symmetrically substituted carbonyl group. However, even small changes in the structure of the reacting systems and in the properties of the environment may affect the details.

For example, a set of SCF–LCGO–MO calculations for the H_3N . . H_2CO system at N . . . C separation of 2.0 Å yields a Δ value about 30% smaller than that observed for the intramolecular, transannular N . . . C=O interaction in clivorine (N . . . C = 1.99 Å).[14] This discrepancy could easily result from the different microenvironments in the two cases (much more polar in the crystal, thus favouring $N^{-\delta} \ldots C^{+\delta} = 0$ interaction) as well as from steric differences between the model system and that encountered in the crystal.

The entire course of the minimum energy path for addition of hydride ion to formaldehyde (shown in Fig 3) can be broken down into three main sections, each characterized by a certain type of interaction.

(*a*) At large distances (d(H⁻ . . . C) > 3Å) the anion approaches the formaldehyde molecule in the molecular plane along the HCH bisector. This section is under electrostatic control, the anion following the line of maximum electrostatic attraction to the molecular dipole. The total energy of the system changes slowly along the path, but at 3 Å distance the stabilization with respect to infinite separation is already about 19 kcal/mole, roughly 40% of the total energy drop (48 kcal/mole) along the entire path.

(*b*) For a distance d(H⁻ . . . C) = 3 Å the H⁻ . . . H separation is about 2.7 Å. At this distance, repulsive interactions between the anion and the two formaldehyde hydrogens, which oppose the electrostatic attraction, start to come into play. As a consequence, the reaction path curves out of the molecular plane. In the section of the path from $d = 3.0$ Å to 2.5 Å the anion glides over the formaldehyde hydrogen atoms until it senses the optimal direction for its attack on the carbon atom. As the H⁻ . . . C=O angle changes from 180° to about 125°, the H⁻ . . . H distance changes by only 0.05 Å, the formaldehyde molecule stays essentially planar (see Fig 3), and the energy drop is less than 2 kcal/mole.

(*c*) As d(H⁻ . . . C) decreases below about 2.5 Å, the potential energy valley starts to become steeper and narrower (Fig 3; right) as orbital interactions (see below) become more and more important. In the highest occupied molecular orbital, for example, the 1 *s* orbital of H⁻ is involved in an in-phase

---

†If the bond number *n* is identified with a factor that multiplies the attractive part of the Morse curve, i.e. $V(r) = D(1 - 2n e^{-br} + e^{-2br})$, then the desired logarithmic dependence on *n* can easily be derived.
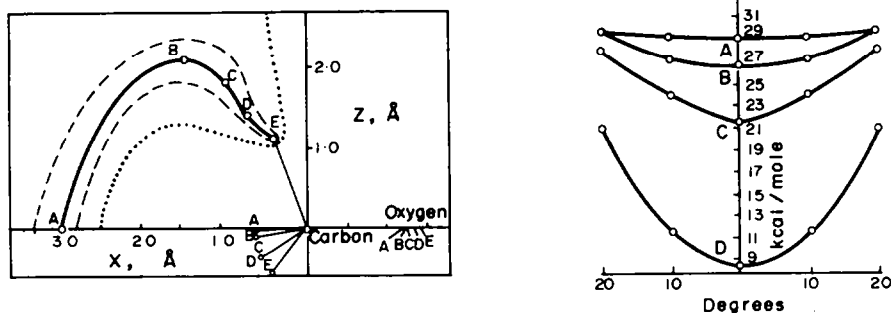
Fig 3. *Left*: Minimum energy path for addition of hydride to formaldehyde. The points A, B, C, D, E correspond to H⁻ . . . C distances of 3·0, 2·5, 2·0, 1·5 and 1·12 Å for which the calculated binding energies (relative to infinitely separated H⁻ + CH₂O) are 19·4, 21·2, 26·7, 39·9 and 48·4 kcal/mole. The dashed and dotted curves show paths that are 0·6 and 6·0 kcal/mole higher than the minimum energy path. *Right*: Energy profiles for lateral angular displacements out of the *XZ* plane.

combination with the π-orbital of formaldehyde and in an out-of-phase combination with the π^*-orbital. The linear combination obtained from this perturbative treatment and the electron distribution of the corresponding MO calculated for $d(\text{H}^- \ldots \text{C}) = 2$ Å are depicted in Fig 4. It is only in this final part of the reaction path that the geometrical changes discussed in detail above occur.

The energy along the reaction path is continuously decreasing and reaches a minimum for methanolate anion without passing through a transition state. Nevertheless, the general features of the energy surface can be fitted with concepts invoked to understand the high kinetic acceleration of enzymic or intramolecular reactions: togetherness[15] or proximity[16] and orbital steering[16] or orientation effects.[17-19] The first two terms, togetherness and proximity, are understood to indicate the loss of translational freedom or decrease in translational entropy i.e. they are thermodynamic factors which in the present example oppose the gain in energy found in part (a) of our surface. Orbital steering or orientation effects may have two interpretations: one of them, in terms of entropy, emphasizes the loss of rotational freedom. This corresponds to part (b) of the surface where the hydride anion tries to find the optimal direction of attack. The other interpretation is in terms of a dependence of the energy of transition states on their geometries. The question of its possible importance in the case of nucleophilic addition has to be postponed until structure and flexibility of the relevant transition states are established. *Ab initio* calculations on the addition of ammonia to formaldehyde plus an additional water molecule introduced to catalyze proton transfer will hopefully yield a model of such a transition state.

Interpretation of kinetic data by Storm and Koshland[17] indicate that molecules with intramolecular O . . . C=O angles α of about 98° show the highest rate of intramolecular lactonization,

whereas both the calculations (part c) and experimental results show that the N . . . C=O or O . . . C=O angles α are about $105 \pm 5°$ for all distances between nucleophile and electrophile smaller than about 2·5 Å (see also next section). The kinetic and structural evidence taken together suggest that the transition state of a nucleophilic addition reaction will also show an angle α between 100 and 110°.

Analogous sequences of events should occur in the course of addition of other nucleophiles e.g. hydroxide, amino nitrogen, etc., to a carbonyl group in ketal or acetal formation, ester or amide formation or hydrolysis. In the case of weaker nucleophile–carbonyl interactions, the energy valley may become shallower, with less stringent orientation requirements on the path of approach, and in the case of large systems all sorts of interactions between other parts of the two reacting species may come into play. Nevertheless, even in such cases the path shown in Fig 3 should correctly represent the bare nucleophilic addition process, to which the other interactions can be added as perturbations.

Tetrahedral intermediate

General considerations. Intermediates occurring in carbonyl hydration, oxygen exchange, ester and amide formation and hydrolysis reactions contain tetrahedral carbon atoms with two or more heteroatom substituents bearing lone pairs. As has been recognized for some time, especially through investigations of the "anomeric" effect in saccharides, such tetrahedral groupings show specific variations in electronic structure,[20,21] spectroscopic properties,[22-25] molecular geometry[26,27] and reactivity.[18,19,28]

α-Chloro-ethers show a marked decrease of the ³⁵Cl nuclear quadrupole resonance frequency if the C–Cl bond is *trans*-antiperiplanar (app) to a lone pair on oxygen. This was interpreted by Lucken in terms of interaction (through-bond) between a
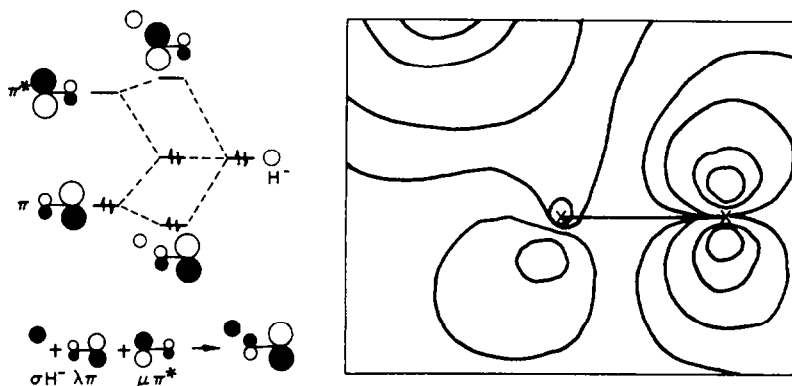
Fig 4.   *Left*: Diagram showing interactions between the $1s$ (H$^-$) AO and the π(C–O) and π^*(C–O) MO's to give the HOMO of the (H$^-$ ... CH$_2$O) system (bottom left). *Right*: Calculated electron density distribution (in XZ plane) of the HOMO at H$^-$ ... C distance of 2·0 Å.

p-type oxygen lone pair and the antibonding σ^*- orbital of the polar C–Cl bond.[22-25] The characteristic, conformation-dependent differences in bond length observed in crystal structures of many sugar derivatives and α-chloro ethers can be interpreted in similar terms.[20,26,27] The preference for syn-clinal (*sc*) rather than *app* conformations of ClCH$_2$–OCH$_3$[29] and FCH$_2$–OH[30,31] may be interpreted in terms of through-bond interactions (see above) and of through-space interactions between the lone pairs of the electronegative atoms.

Conformation effects of this kind may also have a decisive role in the reactivity of such tetrahedral groupings. In the decomposition of conformationally rigid hemi-ortho-esters[18] and hemi-amide-acetals[19] in solution it appears that the bond most readily broken is the one *app* to two lone pairs. Ozonization of acetals shows similar selectivity.[28] For tetrahedral intermediates close in energy to transition states leading to their formation or decomposition, the paths describing these processes should likewise be strongly dependent on the stereoelectronic factors that operate in these intermediates.

It is therefore of interest to study how the strength of bonds (as expressed by bond lengths and overlap populations (OP's)) changes with conformation and to try to correlate these changes with selectivity of bond cleavage. Bond lengths may be obtained from structural data or from quantum mechanical calculations in which the energy of a molecule is minimized with respect to its geometrical parameters. The OP's can be obtained from Mulliken population analysis of the calculated wave functions.

Structural correlations

Many crystal structure analyses have been carried out of compounds containing tetrahedral carbon atoms with two oxygen substituents. These compounds include diols, ketals, hemiketals, acetals and hemiacetals, and can be regarded as cases where the nucleophilic addition reaction has gone to completion to give a stable product.

A significant correlation is found[12] between the quantities $\Delta d = d_1 - d_2$ and $\Delta\beta = \beta_1 + \beta_2 - \beta_3 - \beta_4$ (Figs 1 and 5) which correspond to anti-symmetric stretching and bending distortions of a tetrahedron
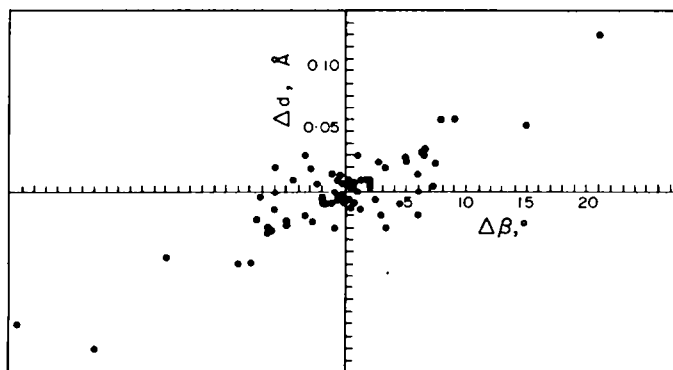

Fig 5.   Scatter plot of $\Delta d = d_1 - d_2$ against $\Delta\beta = \beta_1 + \beta_2 - \beta_3 - \beta_4$ (see Fig 1 for definitions of symbols) for tetrahedral carbon centres carrying two oxygen substituents.

with C_{2v} symmetry. The relation obtained by linear regression on the observed differences is

$$\Delta\beta(°) = 173\ \Delta d(\text{Å}) \text{ (correlation coefficient } 0\cdot84).$$

As one C–O bond lengthens and the other contracts the coupling between the antisymmetric stretching and bending deformations comes into play to bring the shorter bond towards the RR'C plane and the longer bond away from it. This strongly suggests that in the decomposition of a tetrahedral intermediate RR'CXY the leaving of the C–X bond but rather that the line of departure is continually adjusted to maintain a XCY angle of about 110°. The same kind of angle constancy was found in the preferred direction of approach of the nucleophile for N . . . C=O interactions and also, though less pronounced and over a smaller range of distances, in the calculated reaction path for hydride addition (Fig 3). The present result merely confirms that with oxygen as nucleophile, the reverse decomposition follows the same reaction path.

The structural data available for tetrahedral carbon atoms RCXYZ with three electronegative substituents (ortho-esters, amide acetals, etc.) are sparse and not very accurate. However, they seem to support the idea that here too there is a similar coupling between antisymmetric bending and stretching deformations that may be relevant to the decomposition mode.

There is also a correlation between the C–O bond lengths in such tetrahedral units and the torsion angles about neighbouring bonds, i.e. in the R–O–C–O' system the length of the C–O' bond seems to depend on the torsion angle about the C–O bond. For example, there is a definite tendency for very short C–O' bonds to be *app* to O-alkyl or O–H bonds, whereas long C–O' bonds tend to be *sc* [20]. An alternative way of looking at this result is to say that a lone pair on O, which is *app* to O' shortens the C–O bond or lengthens the C–O' bond or does both.

Theoretical results

Here we consider some results of SCF-LCGO-MO calculations, using a moderately extended basis set, for staggered conformations of $CH_2(OH)_2$,[20,32] $CH(OH)_3$,[32] $C(OH)_4$[32] and $CH(NH_2)$-$(OH)_2$,[33] which can be regarded as the simplest models of tetrahedral intermediates. For a given species, changes in correlation energy with confor-

mation are probably negligible and should not affect the *relative* bond lengths and OP's.[34]

For all these systems the computed bond lengths and OP's are found to depend on the number of electronegative substituents and on the molecular conformation in a manner that can be expressed in terms of simple additivity rules. As the number of electronegative groups increases the average C–O distance becomes shorter than in methanol $(d(\text{C–O}) = 1\cdot45\ \text{Å}$, calculated†)—by 0·03 Å for one additional OH group, by 0·05 Å for two, and by 0·04 Å for the combined effect of NH_2 and OH. Bond lengths observed in fluoromethane show a similar trend, being 1·385, 1·358, 1·332 and 1·317 in CH_3F,[35] CH_2F_2,[36] CHF_3[37] and CF_4,[38] respectively.‡

Superimposed on this electronegativity effect is another effect that is conformation dependent. In a given HO–C–O'H' fragment, a lone pair on O that is *app* to C–O' lengthens $d(\text{C–O'})$ by about 0·02 Å and shortens $d(\text{O–C})$ by about 0·01 Å—similar changes occur when either of the OH groups is replaced by NH_2. Analysis of the OP's yields analogous conclusions—the OP of a C–O or C–N bond is reduced by about 0·04–0·06 e (i.e. the bond is weakened) through interaction with an *app* lone pair; the OP of the bond to the atom that carries the lone pair is increased, by 0·02 e for C–O and 0·05 e for C–N.

These variations in bond length and OP with conformation can be interpreted in terms of qualitative MO perturbation theory.[39,40] For a given HO–C–O'H' fragment, interaction of the occupied lone pair orbital (*lp*) on O with the unoccupied $\sigma^*(\text{C–O'})$ orbital gives a modified occupied orbital, $lp + \lambda\sigma^*(\text{C–O'})$, that is bonding with respect to O–C (which is thereby strengthened and shortened) and antibonding with respect to C–O' (which is weakened and lengthened). The same applies, *mutatis mutandis*, to a HO–C–NH$_2$ fragment.

A series of computations[32] for the *sc–app* conformation of $CH_2(OH)_2$ confirms that the differences in bond length and OP of the C–O bonds are also reflected in their stretching force constants. The calculated values are: 5·8 mdyne/Å for the weaker, longer bond $(d(\text{C–O}) = 1\cdot437\ \text{Å})$ and 6·0 mdyne/Å for the stronger, shorter one $(d(\text{C–O}) = 1\cdot408\ \text{Å})$. The stretch-stretch interaction constant is +0·3 mdyne/Å. Thus, for small deviations from equilibrium, the long C–O bond is more easily stretched than the short one, so that a reaction leading to cleavage of the long bond is lower in energy than one leading to cleavage of the short bond, at least in the initial stage.

Internal rotation in methanetriol

A series of computations[33] has been carried out for $CH(OH)_3$ to investigate the changes that occur as a hydroxyl hydrogen is rotated about its C–O bond (*a*) with the other two OH groups (*b*, *c*) held

---

†The basis set used in this work yields a C–O bond length in methanol that is 0·02 Å longer than the experimental value, but this should not affect the results described in this paragraph.

‡In these molecules the bond shortening is presumably due more to the large electronegativity difference than to the influence of *app* lone pairs.

in the + *sc*, + *sc* conformation—the one that corresponds to the stable conformation of the diol. The dependence of the total energy and of the OP's of the three C–O bonds on φ, the rotation angle, are shown in Fig 6. There are two energy minima, a main one at about $\varphi = 250°$ and a subsidiary one at $\varphi = 60°$, separated by maxima at $\varphi = 120°$ and $360°$.
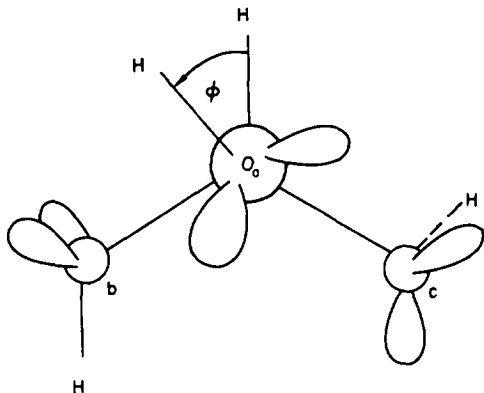
The changes in OP cover a range of more than 0·1 e, the strongest C–O bond being C–O$_a$ at $\varphi = 180°$ (OP = 0·438), the weakest C–O$_c$ at $\varphi = 300°$ (OP = 0·331). For φ angles of 240–300°, i.e. around the stable conformation, there is a large difference in OP (more than 0·05 e) between the weak C–O$_c$ bond and the comparatively strong C–O$_a$ and C–O$_b$ bonds (OP ~ 0·400). Thus, in this conformation of the triol, selective cleavage of C–O$_c$, which has two *app* lone pairs, is expected. For the less stable conformation around $\varphi = 60°$ the C–O$_b$ bond has two *app* lone pairs, but the differences in bond properties are less pronounced. In the arrangement with $\varphi = 180°$ (which does not correspond to an energy minimum) the differences between C–O$_b$ and C–O$_c$, which both have two *app* lone pairs, are small.

The C–H *bond in methanetriol*

Fig 6 shows that the OP of the C–H bond in methanetriol varies approximately as cos φ. However, the contribution of the highest occupied

molecular orbital (HOMO) to this OP varies approximately as $- \cos 2\varphi$ and is a maximum for $\varphi = 90°$ and $270°$, the latter angle being close to the stable conformation of the triol. For these arrangements the C–H bond has two *app* lone pairs, but more important perhaps, it is exactly parallel to the *p*-type lone pair of O$_a$. This orbital (*lp*) interacts with the occupied σ(C–H) orbital to give two modified occupied orbitals, σ(C–H) + λlp and $lp - \mu\sigma$(C–H), the latter corresponding to the HOMO with an appreciable σ(C–H) character.†

The variation in σ(C–H) character of the HOMO with change in conformation hints at a possible conformational specificity in attack of C–H bonds by electrophiles and may provide a rationale for discussing "orbital orientation effects" observed in ozonolysis of acetals.[28]

The CH$_2$(OH)O⁻ *species*

Base-catalyzed decomposition of a diol involves proton transfer to the base, followed or accompanied by loss of hydroxide ion. Some preliminary calculations for the CH$_2$(OH)O⁻ species illustrate some of the factors that may be involved during this process.

The calculations were made for three conformations, with torsion angle τ(H–O–C–O⁻) held at $0°(sp)$, $60°(sc)$ and $180°(ap)$. Their relative energies were calculated as 0, 7 and 13 kcal/mole, respectively. Since the stable conformation of the diol is *sc*, *sc*, proton transfer to the base is accompanied by a rotation of the remaining hydroxyl group to give the stable *sp* conformation of the anion. The calculated net charges, OP's and C–O bond lengths for this conformation (tetrahedral angles assumed) are shown below. The C–O⁻ bond is short and strong ($d = 1·355$ Å, OP = 0·671 e) while the C–OH bond is long and weak ($d = 1·477$ Å, OP = 0·285 e) compared with the C–O bonds in the stable conformation of methanediol ($d = 1·437$ Å, OP = 0·331 e). Thus, although cleavage of methanediol is facilitated by base catalysis, the effect of ionization on

---
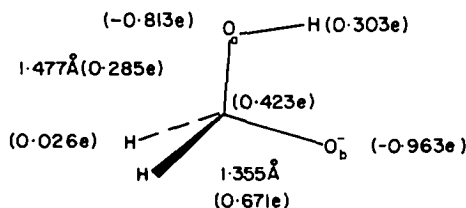
†It is much easier to use qualitative perturbation MO theory *a posteriori* than *a priori*! It would have been difficult to know in advance whether the main interaction of *lp* is with σ(C–H) or σ*(C–H).
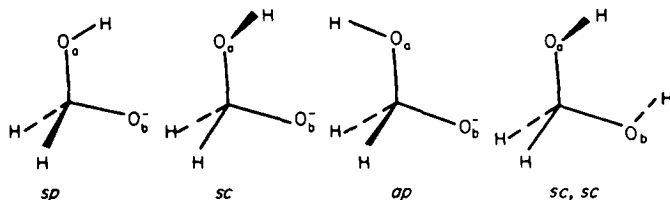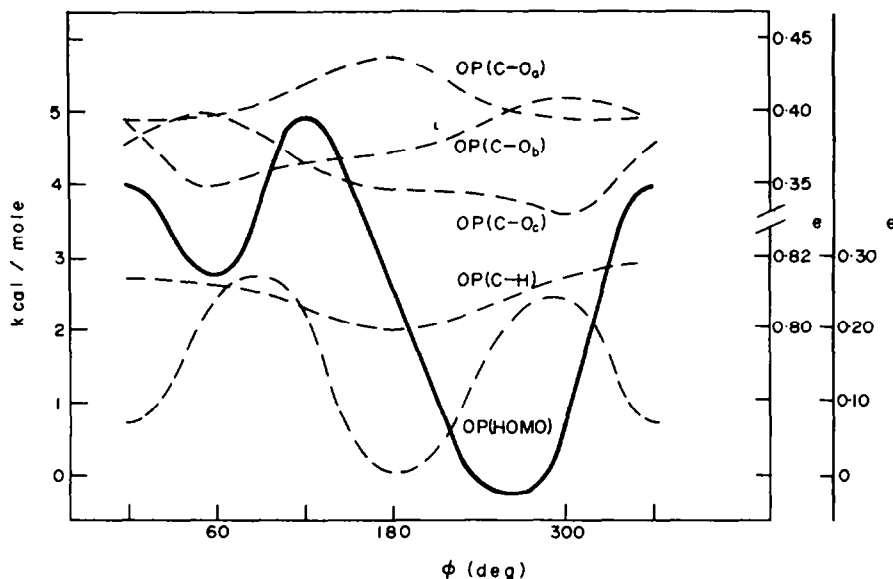
Fig 6. Methane triol. The thick line shows the variation in total energy as function of the rotation angle φ (see text diagram for definition). The other curves show the corresponding variations in total OP's of C–O and C–H bonds and also the contribution of the highest occupied MO (HOMO) to the total OP(C–H).

the length and OP of the bond to be broken is only marginally larger than the effect of an *app* lone pair (discussed in an earlier section).

The isolated species $CH_2(OH)O^-$ is calculated to be about 60 kcal/mole more stable than the separated species $CH_2O + OH^-$. Nucleophilic addition of OH^- to formaldehyde is thus an even more downhill process than addition of hydride ion (stabilization 48 kcal/mole), even though the adduct $CH_2(OH)O^-$ seems less advanced along the addition reaction path than CH_3O^-, judging from the C–O⁻ bond lengths (1·355 Å in $CH_2(OH)O^-$, 1·405 Å in CH_3O^-).

The effect of acid-catalyzed decomposition remains to be investigated.

Summary and outlook

In this contribution we have described how a general pathway for nucleophilic addition/elimination reactions has been inferred by analysing crystal structure data and confirmed by *ab initio* SCF–LCGO–MO calculations. Stereoelectronic effects accounting for subtle details of molecular structure also shed light on the conformational specificity associated with certain decompositions of acetals and amide acetals. The role of acid and base catalysis, as well as the structure and energy of transition states, remain subjects for future investigations. Nevertheless, we cherish the hope that the present results may contribute to a better understanding of enzyme specificity, in much the same way as information about the structure of

simple peptides was helpful in elucidating the secondary and tertiary structure of proteins. It is encouraging that the pathway for nucleophilic addition proposed here seems perfectly consistent with recent results of Huber,[41] on the relative positions of the serine-Oγ-atom in the active sites of several proteolytic enzymes and enzyme-inhibitor complexes.

REFERENCES

[1] H. B. Bürgi, *Inorg. Chem.* 12, 2321 (1973)
[2] H. B. Bürgi, J. D. Dunitz and P. Murray-Rust, unpublished work
[3] H. B. Bürgi, J. D. Dunitz and E. Shefter, *J. Am. Chem. Soc.* 95, 5065 (1973)
[4] L. C. Snyder and H. Basch, *Ibid.* 91, 2189 (1969); L. C. Snyder, *Robert A. Welch Foundation Res. Bull.*, No. 29 (1971)
[5] W. J. Jencks, *Catalysis in Chemistry and Enzymology* McGraw Hill, New York, N.Y. (1969)
[6] M. L. Bender, *Mechanisms of Homogeneous Catalysis from Protons to Proteins* Wiley-Interscience, New York, N.Y. (1971)
[7] R. Henderson and J. H. Wang, *Ann. Rev. Biophys. Bioeng.* 1, 1 (1972)
[8] J. F. Kirsch, *Ann. Rev. Biochem.* 42, 205 (1973)
[9] J. P. Guthrie, *J. Am. Chem. Soc.* 95, 6999 (1973) and refs therein
[10] *Interatomic Distances Supplement*, Chem. Soc. Spec. Publ., No. 18 (Edited by L. E. Sutton). Chem. Soc., London (1965)
[11] L. Pauling, *The Nature of the Chemical Bond*, (3rd Edition) p. 246. Cornell University Press, Ithaca, N.Y. (1963)

[12]H. B. Bürgi, J. D. Dunitz and E. Shefter, *Acta Cryst.*, in press

[13]H. B. Bürgi, J. M. Lehn and G. Wipff, *J. Am. Chem. Soc.*, in press

[14]K. B. Birnbaum, *Acta Cryst.* **B28**, 2825 (1972)

[15]M. I. Page and W. P. Jencks, *Proc. Nat. Acad. Sci. USA* **68**, 1678 (1971)

[16]A. Dafforn and D. E. Koshland, Jr., *Biochem. Biophys. Res. Comm.* **52**, 779 (1973) and refs therein

[17]D. R. Storm and D. E. Koshland, Jr., *J. Am. Chem. Soc.* **94**, 5815 (1972)

[18]P. Deslongchamps, P. Atlani, D. Fréhel and A. Malaval, *Can. J. Chem.* **50**, 3405 (1972)

[19]P. Deslongchamps, C. Lebreux and R. Taillefer, *Ibid.* **51**, 1665 (1973)

[20]G. A. Jeffrey, J. A. Pople and L. Radom, *Carbohyd. Res.* **25**, 117 (1972)

[21]S. David, O. Eisenstein, W. J. Hehre, L. Salem and R. Hoffmann, *J. Am. Chem. Soc.* **95**, 3806 (1973)

[22]E. A. C. Lucken, *J. Chem. Soc.* 2954 (1959)

[23]P. Linscheid and E. A. C. Lucken, *Chem. Comm.* 425 (1970)

[24]Z. Ardalan and E. A. C. Lucken, *Helv. Chim. Acta* **56**, 1715 (1973)

[25]L. Guibé, J. Augé, S. David and O. Eisenstein, *J. Chem. Phys.* **58**, 5579 (1973)

[26]M. Sundaralingam, *Biopolymers* **6**, 189 (1968)

[27]C. Romers, C. Altona, H. R. Buys and E. Havinga, *Topics Stereochem.* (Edited by E. L. Eliel, N. L. Allinger) Vol 4, p. 39. Wiley-Interscience (1969)

[28]P. Deslongchamps, C. Moreau, D. Fréhel and P. Atlani, *Can. J. Chem.* **50**, 3402 (1972)

[29]M. C. Planje, L. H. Toneman and G. Dallinga, *Rec. Trav. Chim.* **84**, 232 (1965)

[30]L. Radom, W. J. Hehre and J. A. Pople, *J. Am. Chem. Soc.* **94**, 2371 (1972)

[31]S. Wolfe, A. Rauk, L. M. Tel and I. G. Csizmadia, *J. Chem. Soc.* (B), 136 (1971)

[32]J. M. Lehn, G. Wipff and H. B. Bürgi, *Helv. Chim. Acta* **57**, 493 (1974)

[33]J. M. Lehn and G. Wipff, unpublished results

[34]J. M. Lehn, *Ab initio Conformational Analysis* in *Conformational Analysis*, (Edited by G. Chiurdoglu) p. 129. Academic Press, New York (1971)

[35]C. C. Costain, *J. Chem. Phys.* **29**, 864 (1958)

[36]M. K. Lo and W. H. Flygare, *J. Mol. Spectr.* **25**, 365 (1968)

[37]S. N. Ghosh, R. Trambarulo and W. Gordy, *J. Chem. Phys.* **20**, 605 (1952)

[38]C. W. W. Hoffman and R. L. Livingston, *Ibid.* **21**, 565 (1953)

[39]R. Hoffmann, *Acc. Chem. Res.* **4**, 1 (1971)

[40]W. C. Jorgensen and L. Salem, *The organic Chemist's book of orbitals.* Academic Press, New York (1973)

[41]R. Huber, private communication