The principles of conformational analysis

Nobel Lecture, December 11, 1969

Il y a trois périodes dans l’histoire de toute découverte.
Quand elle est annoncée pour la première fois, les gens pensent que ce n’est pas vrai.
Puis, un peu plus tard, quand son exactitude leur paraît si flagrante qu’ils ne peuvent plus la nier, ils estiment que ce n’est pas important.
Après cela, si son importance devient assez manifeste, ils disent: en tout cas, ce n’est pas nouveau.

William James

The importance of conformational analysis in Chemistry became manifest during the decade immediately after the last World War. This lecture is, therefore, more an account of chemical history than of recent advances. In order to appreciate the significance of conformational analysis, a short introduction describing the development of structural theory in Organic Chemistry is necessary.

In the second half of the nineteenth century it became possible, thanks to the theories of Kekulé and others, to assign a constitution to each organic substance. The constitution is simply the specification of which atoms are bonded to which in the molecule and, in the great majority of cases, is unambiguous. A constitutional formula has no stereochemistry. The necessity to consider the stereochemistry of molecules became self-evident when two or more distinctly different substances were found to have the same constitution. It was Le Bel and, especially, Van ’t Hoff who, also in the nineteenth century, introduced the idea of configuration. One can consider that the configurations of a molecule of a given constitution represent the specification of the order of the bonds in space about the atoms, or groups in the molecule which give rise to stereoisomers. For the great majority of substances this means the specification of the order of the bonds about an asymmetric "double bond" or about a "centre of asymmetry", normally a carbon atom substituted by four different groups. The first type of stereoisomerism is called "geometrical isomerism", the latter "optical isomerism". The number of possible stereoisomers of a given configuration then becomes \(2^n \times 2^m = 2^{n+m}\) where \(n\) is the
number of "asymmetric double bonds" and \( m \) is the number of "centres of asymmetry". This formula of Van 't Hoff, the first winner of a Nobel Prize, is still of fundamental importance today, nearly a century after it was first written down. Organic chemists have been very busy in demonstrating the truth of constitutional and configurational theory. Between one and two million different organic compounds have so far been prepared and we can prepare as many more millions as are required.

Van 't Hoff had a very clear idea of the reason for the success of the \( 2^n+m \) formula. It was based on the concept of "restricted" rotation about double bonds and of "free" rotation about single bonds. The latter concept was necessary to explain why most "optical isomers" did not have a myriad of isomers themselves. We shall not be concerned in this lecture with "geometrical isomerism" about double bonds. Optical isomerism is based on the idea of chirality, or the non-superposability of object and mirror image.

The first indication that rotation about single bonds was not always free came in 1922 when Christie and Kenner were able to resolve 2,2'-dinitrodiphenyl-6,6'-dicarboxylic acid (I) into optically active forms. This resolution is possible because the four bulky groups at the 2,2'-, 6- and 6'-positions prevent rotation about the central carbon-carbon single bond. Many analogous examples of restricted rotation of this type were later investigated. It does not seem, however, that organic chemists were much worried about barriers to rotation in organic molecules in general at that time because there was no technique available to demonstrate the phenomenon experimentally.

In the decade starting in 1930, chemical physicists noted a discrepancy between the observed and calculated entropy of ethane. It was clear that this could only be explained by a barrier to free rotation about the two methyl groups, but whether this barrier was attractive or repulsive in origin with respect to the hydrogen atoms of the two methyl groups was a subject of considerable argument. However, the existence of such a barrier to free rotation in ethane implied that such barriers existed for aliphatic and alicyclic compounds in general.
The problem was clarified by studies on the simple alicyclic hydrocarbon cyclohexane (II) and on derivatives of this compound. It had been appreciated for many years that cyclohexane molecules could be constructed in two forms, the boat (IV) and the chair (III), both free from angle strain. This distinction had, however, no meaning to organic chemists since there was no reason to know which form was preferred, nor was it understood that the preference would have any chemical consequences. It was the fundamental electron-diffraction work of Hassels\textsuperscript{3,4} that established clearly that the chair conformation (III) was always preferred. In this chair conformation the hydrogen atoms are as far apart as possible and correspond to the "staggered" form of ethane and of aliphatic compound in general. Similar considerations apply to medium and large alicyclic rings\textsuperscript{5}. One can conclude, therefore, that the barrier to rotation in such substances is repulsive rather than attractive in character.

Rather than use the vague terms boat and chair forms of cyclohexane, it is convenient to have a general term. In fact, the appropriate word "conformation" had already been used in sugar chemistry by W. N. Haworth\textsuperscript{6}. The most general definition of conformation is as follows\textsuperscript{7}: "the conformations of a molecule (of defined constitution and configuration) are those arrangements in space of the atoms of the molecule which are not superposable upon each other". Such a definition includes arrangements of atoms in which angle strain has been introduced, as well as bond extension and compression. It replaces an earlier definition\textsuperscript{8,9} which excluded angle strain and bond extension or compression. Thus all molecules have theoretically an infinite number of conformations. It is fortunate that the complexities which might arise from such a definition are minimised by the fact that, in general, only a few of the possible conformations are energetically preferred. One may therefore consider "chair", "boat" and "twist-boat" (a conformation half-way between two boats; see ref. 10) conformations of cyclohexane all of which are free from angle strain. The stability order is chair>twist-boat>boat.

It is obvious that in the chair conformation of cyclohexane two geometrically distinct types of carbon-hydrogen bonds are present\textsuperscript{11}. Six of the C-H bonds are parallel to the three-fold axis of symmetry (V) and are called <axial> (ref. 12). The other six (VI) are approximately in an equatorial belt around the three-fold axis and hence are called <equatorial>. As soon as a substituent is introduced into a cyclohexane ring the molecule may adopt a preferred chair conformation with the substituent either axial or equatorial. Owing to repulsive non-bonded interactions between axial groups the equatorial conforma-
tion is, in general favored. In the case of multiply substituted cyclo-
hexanes the preferred conformation is, provided dipolar interactions are not
dominant, that with the maximum possible number of substituents equatorial.

When two cyclohexane rings are fused together as in the configurational
isomers trans- (VII) and cis- (VIII) decalin, a unique two-chair confor-
mation, (IX) and (X) respectively, can be written for both molecules. Bastiansen
and Hassel showed that, as expected from consideration of non-bonded
interactions, these two conformations were indeed the preferred ones.

At about this time the first semi-empirical, semi-quantitative calculations
of non-bonded interactions were appearing in the literature. Application of
these methods to ethane, to cyclohexane and to the trans- and cis-deca
gave results in qualitative agreement with the findings of Hassel and others
mentioned above. In order to carry out these calculations, which at that time,
in the absence of computers, were exceedingly arduous, special models were
constructed which later proved to be very useful in working out the prin-
ciples of conformational analysis. The same models were also useful in under-
standing, in conformational terms, the dissociation constants of the tricarbox-
ylic acid (XI) obtained by the oxidative degradation of abieric acid (XII). This
was, in fact, an early example of the use of conformational analysis.

The stage was now set for a much fuller appreciation of the meaning of
conformational analysis. At the time (1950) when our paper in Experientia
was written, steroid chemistry was already a major branch of science which
had just received a strong additional stimulus from the discovery of the utility of cortisone. There was an enormous literature of stereochemical fact which had not been interpreted properly in its three-dimensional aspects. Most steroids contain three six-membered rings fused trans to a five-membered ring and the two commonest configurational arrangements can be represented as in (XIII) and (XIV). Accepting that the three six-membered rings will in both cases adopt the preferred and unique three-chair conformation then (XII) may be represented in three dimensions as (XV) and (XIV) as (XVI).

In steroid chemistry it is convenient to designate substituents on the same side of the molecule as the two methyl groups as β-oriented. Those on the opposite side of the molecule are then said to be α-oriented. Substituents attached to the steroid nucleus thus have a configuration which is α or β and which can be determined by the classical methods of stereochemistry (ring formation, ring fission, etc.). The key to the application of conformational analysis is that the ring fusions of the steroid nucleus fix the conformation of the whole molecule such that a substituent necessarily has both a configuration (α or β) and a conformation (equatorial or axial). Since, at a given carbon atom in the steroid nucleus, a substituent will be more stable equatorial than axial it follows that one can at once predict the more stable configuration between a pair of isomers. Thus a 3β-substituted (equatorial) trans A/ B steroid (XVII) should be more stable than the corresponding 3α-substituted (axial) compound.
(XVIII). This is in agreement with experiment. The same argument applies to all the other substitutable positions in the steroid nucleus in the six-membered rings, and, in general, good agreement with experiment is seen. The same applies for all molecules, for example triterpenoids, where fused all-chair conformations are present.

The relative stability of substituents is determined by repulsive non-bonded interactions (steric compression). Therefore, not only can one predict which isomer (α- or β-) will be formed in a chemical reaction which for mechanistic reasons gives the more stable product, but also any reaction which involves steric compression can be understood better. Thus in the alkaline hydrolysis of esters where the transition state for the reaction is more space-demanding than the initial state, one can predict that an equatorial isomer will hydrolyse faster than an axial isomer attached to the same carbon atom. This is, in general, true, and the principle aids in the prediction of selective hydrolysis reactions.

Every chemical reaction has a transition state. Many transition states have a well-defined preferred geometrical requirement. Thus in an E₂-type reaction, where two substituents attached to α-carbons are eliminated simultaneously by the attack of a reagent, the preferred geometry is that where the two carbons and the two substituents (X and Y) are coplanar. The two possible arrangements are anti (XIX) and syn (XX). The latter geometry is not available in steroids without distortion of a chair conformation. The former geometry (XIX) is, however, inevitably present in trans-1:2-diaxially substituted compounds (e.g. XXI), but not present in the corresponding trans-1:2-diequatorially substituted isomer (e.g. XXII). Such geometrical relationships are clearly shown if one looks along the C₅-C₆ bond. Thus in (XXI) the projection (XXIII) is seen and in (XXII) the projection (XXIV).
Thus for a bimolecular E
\(_2\) reaction of the two bromine atoms induced by iodide ion (see XXV)
\(^{20}\), it could be predicted that the dibromide (XXI) would eliminate much faster than dibromide (XXII). Fortunately both dibromides could be prepared and their configurations determined
\(^{21}\). As anticipated, the rate of elimination of bromines for the dibromide (XXI) was several powers of ten faster than the rate of the corresponding elimination from the dibromide (XXII)
\(^{21}\). This was the first example of a phenomenon that was demonstrated later to be quite general for dibromides and many other types of eliminatable functions
\(^{22,23}\). In general the geometrical requirements of the transition states of all chemical reactions are advantageously examined in conformational terms using steroids or other molecules with locked conformations.

The phenomenon of neighbouring group participation demands a conformational interpretation (diaxial participation) which is well exemplified in steroids
\(^{24}\). Similarly, the opening of small membered rings like the halonium ion
\(^{24,25}\) or the epoxide group
\(^{9,26}\) gives predominantly diaxial rather than diequatorial products.

When the diaxial dibromide (XXI) is kept in solution at room temperature it rearranges spontaneously to an equilibrium with the more stable diequatorial dibromide (XXII). In effect two axial C-Br bonds are exchanged for two equatorial C-Br bonds. This is a general reaction
\(^{22,23}\) for 1,2-dibromides. We conceived that this rearrangement should be part of a generalised diaxial\(\leftrightarrow\) diequatorial rearrangement process as in the scheme (XXVI)\(\leftrightarrow\) (XXVII). By an appropriate choice of X and Y the truth of the proposition was demonstrated
\(^{25,26}\). In practise, this reaction is a convenient route for shifting an oxygen function from one carbon atom to the adjacent carbon.

In the above discussion we have mainly correlated asymmetric centres of known configuration with their predicted conformations. The argument can, of course, be reversed and then provides a powerful method for deducing the configurations of compounds where a preferred all-chair conformation can be postulated. The first serious applications were in triterpenoid chemistry.
The natural compound oleanolic acid (XXVIII) has 8 centres of asymmetry and can exist in principle in $2^8 = 128$ racemic configurations. Only one configuration [(+) or (-)] is synthesised by Nature. Because the analysis of conformations is easier with saturated six-membered rings we studied the saturated oleananolic acid (XXIX) which corresponds [(+) or (-)] to one racemate from a possible $2^8 = 256$ racemic configurations. Such is the power of the conformational method that the problem of configuration was reduced by chemical procedures only to a choice between two configurations (XXX and XXXI). The latter was shown to be correct by a later X-ray crystallography study. It corresponds to the planar formula (XXIX).

By conformational analysis configurations could be assigned to typical triterpenoids like lanosterol (XXXII), euphol (XXXIII), cycloartenol...
(XXXIV) and onocerin (XXXV) at the same time as their constitutions were determined. Nowadays, of course, X-ray crystallographic analyses are done so easily and speedily that there is no special merit in the conformational method of determination of configuration. But it was important in the early 1950's and is useful today when the investigator does not have ready access to X-ray facilities. The X-ray method has the overall-advantage that it determines constitution, configuration and preferred (in the crystalline state) conformation all at the same time. In general the preferred conformation in the crystalline state is that which would be predicted by the principles of conformational analysis.

Until 1957 there were no exceptions to the rule of preferred chair conformations in molecules where the configurations of the asymmetric centres permitted a choice between boat and chair to be made. In the course of studies of the bromination of lanostenone (XXXVI) two bromo-ketones (XXXVII) and (XXXVIII) were obtained in 95% and 5% yields. Normally these would have been assigned the 2α- (XXXVII) and 2β- (XXXVIII) configurations respectively, both based on a ring A chair conformation. However, by both infrared and ultraviolet spectroscopy it was shown that both of these compounds had their bromine equatorial. Further chemical investigations then demonstrated that the 2β-bromo-compound has, in fact, the boat (or more correctly twist-boat) conformation (XXXIX). This first exception to the normal conformational preference is due to the large methyl-methyl-bromine 1:3-diaxial
interactions in the conformation (XXXVIII) and to the fact that the A ring contains one trigonal atom (the carbon of the carbonyl group). Later this conformational anomaly was extensively investigated\(^{37}\) and its reality has been further confirmed by nuclear magnetic resonance studies\(^{38}\).

Conformational analysis was put on a quantitative basis by Winstein and Holness\(^ {39}\) and especially by Eliel\(^ {40}\). The latter author, a recognised authority in the field, has made a thorough study\(^ {41}\) of the differences in free energy between axial and equatorial substituents in six-membered rings. There remain, however, certain subtle aspects of the conformations of molecules such as steroids and triterpenoids which still demand an adequate explanation from quantitative theory. Thus we have shown\(^ {23,42}\) that if benzaldehyde condenses with lanostenone (XXXVI) under mildly basic conditions at a rate of (say) 100 to give the 2-benzylidene derivative (XL), then the simple non-polar derivative lanostanone (XLI) condenses at a rate of 55 and the isomeric olefin (XLII) at a rate of only 17. Similarly, cholestanone (XLIII, \(R = \text{C}_8\text{H}_{17}\)) give the 2-benzylidene derivative (XLIV) at a rate of 182. Simple side-chain derivatives (see XLIII) condense at the same rate. However, the isomeric olefins (XLV) and (XLVI) condense at rates of 645 and 43, a difference of nearly thirty-fold for only a change in position of a (relatively) remote double bond. We have attributed these long-range effects to "conformational transmission" implying a distortion of bond angles by substituents that is transmitted through molecules to much greater distances than hitherto suspected. Such phenomena are beginning to receive adequate explanation, at least in qualitative and semi-quantitative terms\(^ {43}\).
As already mentioned above, conformation preferences can be calculated by semi-empirical methods. Now that computers have taken away the arduous arithmetic involved it has been possible to make rapid progress. For example, preferred conformations have been calculated for alicyclic rings larger than six-membered. These calculations provide valuable clues to an understanding of the chemistry of such systems. Undoubtedly it will be possible soon to calculate fine details of conformation, and at that point long-range effects will also be calculable. At that time also optical activity, optical rotatory dispersion and circular dichroism will be understandable in their quantitative magnitudes.

Although the principles of conformational analysis are most clearly demonstrated in saturated six-membered cyclohexane ring systems, nevertheless the same basic approach is useful in understanding the reactions of unsaturated and of heterocyclic compounds. For example, cyclohexene can be assigned the conformation (XLVII) with equatorial and axial hydrogens as indicated. The hydrogens marked with a prime can then be called quasi-equatorial and quasi-axial. The conformation of cyclohexene is more easily deformed than...
that of cyclohexane even when fixed to other ring systems. Nevertheless the symbol (XLVII) has found general favour in conformational analysis as an expression of reality.

![Image of molecular structures]

The introduction of heteroatoms into a cyclohexane ring as in piperidine (XLVIII) or pyran (XLIX) raises conformational problems of considerable interest and sophistication. Thus one must consider if pairs of p-electrons have bulk or not and if so, is it greater or less than the hydrogen when attached to the heteroatom. In this way a new field of conformational analysis has rapidly developed.

As has already been discussed, the choice of a preferred conformation was originally made on the basis of inference from the electron-diffraction work on simple compounds and from other physical evidence. X-Ray crystallography is nowadays an accurate and rapid method of determining conformation in the crystal lattice, which conformation usually corresponds to the preferred conformation in solution. There is, however, another physical method, nuclear magnetic resonance, which in the last decade has become with every justification-predominant in the determination of conformation in solution. In many cases extremely detailed conformational analysis can be carried out. A simple example, which had great consequences for carbohydrate chemists, is the work of Lemieux and his colleague.

An enzymatic reaction involves a large molecule - the enzyme - and a relatively small molecule - the substrate. Any complete understanding of the mode of action of an enzyme will require a knowledge of the conformations of the substrate and of the enzyme, as well as of the functional group reactivity. We are now in a position with most substrates to specify the preferred conformation involved in the reaction. In the case of steroidal substrates the conformation can be described in a detail which will soon be quite exact. This knowledge must soon have important consequences in biology.
Conformational analysis may be said to have come of age in that two excellent monographs have now appeared\textsuperscript{41,49} which review present knowledge in detail. It is interesting to observe how an acorn of hypothesis can become a tree of knowledge.