Anal. ($C_{26}H_{35}NO$) C, H, N.

2(a)-(α-Hydroxybenzyl)-4-phenylpiperidine (34ay). In the same manner as for **22**, **34ay** was prepared using 0.67 g (6.0 mmol) of potassium *tert*-butoxide, 0.02 mL (0.92 mmol) of water, and 0.228 g (0.58 mmol) of **39ay** in 10 mL of *tert*-butyl alcohol. Recrystallization from ether gave 0.132 g (85%) of **34ay** as a white crystalline solid: mp 69-71 °C; ^{1}H NMR ($CDCl_3$) δ 7.38-6.90 (m, 10 H, PhH), 4.47 (d, 1 H, J = 8 Hz, CHO), 4.00 (br, 2 H, NH and OH), 3.05-2.65 (m, 2 H, NCH_2), 2.79 (m, 1 H, J = 8 Hz, NCH), 2.44 (m, 1 H, CHPh), 1.60 (m, 2 H, CH_2), 1.33 (m, 2 H, CH_2); IR (mull) 3400 (NH), 3200 (OH), 1650 (NH), 1310 cm^{-1} (CO); mass spectrum (70 eV), *m/e* (relative intensity) 267 (2, M^+), 231 (7), 161 (100), 160 (100), 144 (13), 117 (15), 104 (8), 91 (10).

Anal. ($C_{18}H_{21}NO$) C, H, N.

***threo*-2(a)-(α-Hydroxybenzyl)-4-phenylpiperidine (34ax).** In the same manner as for **22**, **34ax** was prepared using 1.57 g (14.0 mmol) of potassium *tert*-butoxide, 0.05 mL (2.3 mmol) of water, and 0.536 g (1.36 mmol) of **39x** in 10 mL of *tert*-butyl alcohol. Recrystallization from ether gave 0.303 g (83%) of **34ax** tentatively assigned the threo isomer,[26] as a white crystalline solid: mp 89-90 °C; ^{1}H NMR ($CDCl_3$) δ 7.35-6.95 (m, 10 H, PhH), 4.78 (d, 1 H, J = 11 Hz, CHO), 3.30 (br, 2 H, NH and OH), 3.10-2.75 (m, 4 H, NCH_2, CCH, and CHPh), 2.0-1.45 (m, 4 H, CH_2); IR (mull) 3280 (NH), 3200 (OH), 1615 (NH), 1280 cm^{-1} (CO); mass spectrum (70 eV), *m/e* (relative intensity) 267 (61, M^+), 231 (42), 230 (50), 161 (100), 160 (100), 144 (77), 132 (44), 117 (99), 104 (96), 91 (48).

Anal. ($C_{18}H_{21}NO$) C, H, N.

Acknowledgment. We are grateful to the National Science Foundation and the National Institutes of Health for their support of this work. We are also grateful to Dr. D. B. Reitz for the initial work on the 2,2-diethylbutanamides, to Professor E. Leete for a gift of conhydrine, and to Professor A. I. Meyers for information about his work with formamidene derivatives of piperidines.

Registry No. 4, 64712-52-3; **4** (deuterated), 88132-03-0; **5**, 79288-71-4; **6**, 79288-68-9; **7**, 88131-98-0; **8**, 88131-75-3; **9**, 88131-76-4; **10**, 88131-99-1; **11**, 79288-70-3; **12**, 88131-77-5; **13**, 88131-78-6; **14**, 88132-07-4; **14** (deuterated), 88132-13-2; **15**, 88132-10-9; **15** (deuterated), 88156-72-3; **16**, 88132-06-3; **16** (deuterated), 88132-14-3; **17**, 35354-15-5; **18**, 78986-72-8; **19**, 78986-74-0; **20**, 88131-79-7; **21**, 88131-80-0; **22**, 18259-40-0; **23**, 78986-75-1; **24**, 88132-00-7; **25**, 88132-01-8; **26**, 88131-81-1; **27**, 78986-86-4; **28** (isomer 1), 88131-82-2; **28** (isomer 2), 88132-04-1; **29**, 88132-02-9; **30**, 88156-71-2; **31** (isomer 1), 88131-83-3; **31** (isomer 2), 88131-84-4; **32**, 88131-85-5; **33**, 63401-12-7; **34ax**, 88131-86-6; **34ay**, 88131-87-7; **34ee**, 88131-88-8; **34et**, 88131-89-9; **35**, 88131-90-2; **36** (isomer 1), 88131-91-3; **36** (isomer 2), 88131-92-4; **37**, 88131-93-5; **38**, 88131-94-6; **39** (isomer 1), 88131-95-7; **39** (isomer 2), 88131-96-8; **40**, 88131-97-9; **41**, 88132-05-2; **42**, 79288-70-3; ethyl 2,2-diethylbutanoate, 34666-17-6; ethyl 2-ethylbutanoate, 2983-38-2; 2,2-diethylbutyric acid, 813-58-1; dodecyl bromide, 143-15-7; 4,6-diamino-1,3-diisopropylbenzene, 3102-71-4; 4,6-diamino-1,3-diisopropylbromobenzene, 77256-79-2; 2,6-bis(dimethylamino)-3,5-diisopropylbromobenzene, 77256-80-5; *N,N*-diethyl-2-(dimethylamino)benzamide, 88132-08-5; *N,N*-diethyl-2-(dimethylamino)-6-ethylbenzamide, 88132-09-6; 2-methoxy-3-isopropyl-6-methylbenzoic acid, 72135-27-4; *N,N*-diethyl-2-methoxy-3-isopropyl-6-methylbenzamide, 88132-11-0; *N,N*-diethyl-2-methoxy-3-isopropyl-6-ethylbenzamide, 88132-12-1; 2,4,6-triisopropylbenzoic acid, 49623-71-4; diethylamine, 109-89-7; piperidine, 110-89-4; 4-*tert*-butylpiperidine, 1882-42-4; benzaldehyde, 100-52-7; benzophenone, 119-61-9; trimethylsilyl chloride, 75-77-4; 4-phenylpiperidine, 771-99-3; butyl iodide, 542-69-8; diethylcarbamoyl chloride, 88-10-8; 2-dimethylaminobenzoic acid, 610-16-2.

Supplementary Material Available: Deuteration experiments and substituted benzamide syntheses (12 pages). Ordering information is given on any current masthead page.

Directional Hydrogen Bonding to sp^2- and sp^3-Hybridized Oxygen Atoms and Its Relevance to Ligand-Macromolecule Interactions

Peter Murray-Rust[1a] and Jenny P. Glusker*[1b]

Contribution from The Institute for Cancer Research, The Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111. *Received June 13, 1983*

Abstract: In order to analyze the directionality of hydrogen bonding to oxygen atoms the Cambridge Crystallographic Data File was searched for O···X (X = N, O) intermolecular contacts at less than 3 Å in structures containing ether, ketone, epoxide, enone, and ester groups. The results are represented as scatterplots. However, for further clarity, they are also represented as diffuse (probability) densities obtained by superposing spherical atomic electron density functions with a parameterizable temperature factor on each point in the scatterplot and contouring the maps at a height proportional to the number of data points. In all systems the largest concentration of hydrogen-bonded X groups lay in the direction commonly ascribed to lone pairs, but generally only one such X group per oxygen atom. In addition we obtained a tentative view that epoxides had a broader distribution of points for X than did ethers; in this respect the positions of the lone pairs in epoxides are more like those in ketones. The experimental plots that are described here may be used to map the stereochemistry of that part of the macromolecule (possibly of unknown structure) that binds to the ligand molecule.

There have been many attempts recently to understand the manner by which small molecules bind to macromolecules and how this binding can result in specific activity. The specificity of interactions between oxygen- and/or nitrogen-containing molecules in organic and biological systems suggest that hydrogen bonding and other directional constraints are important. Although there has been much interest in the geometry of the A-H···B system (A, B = oxygen, nitrogen, etc.), much less information is available on the angular distribution of proton donors around an acceptor, e.g., the H···O=C angle.[2] However, the constraints on this (and similar) systems are of vital importance to our understanding of the specificity of binding of ketones, epoxides, and related molecules to biological and other macromolecules.

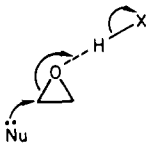
Hydrogen bonding to epoxides has a second important function. The reactivity of epoxides is linked to their potency as alkylating

(1) (a) Present address: Glaxo Group Research, Ltd., Greenford Road, Greenford, Middlesex UB6 OHE, England. This work was done while on leave from the Chemistry Department, The University of Stirling, Stirling, Scotland, FK9 4LA, UK. (b) To whom correspondence should be addressed.

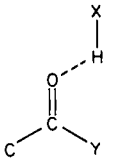
(2) Donohue, J. In "Structural Chemistry and Molecular Biology"; Rich, A., Davidson, N., Eds.; W. H. Freeman and Company: San Francisco and London, 1968, pp 443-465.

0002-7863/84/1506-1018$01.50/0

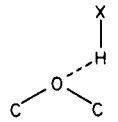
agents via an opening of the epoxide ring. This is believed to be catalyzed by hydrogen bonding to the ring oxygen,[3-5] it is effectively a partial protonation as shown in I (X = O, N, etc.; Nu = nucleophile). Presumably the most effective catalysis will be the result of an optimization of hydrogen-bond formation.



The two hydrogen-bonding systems we shall investigate here are sp^2- and sp^3-hybridized oxygen atoms in systems II and III.



II, X = O, N; Y = O, C (esters, ketones)



III, X = O, N (epoxides, ethers, esters)

The first is relevant to the binding of steroids (particularly 4-en-3-ones) and the second is relevant to the binding of epoxides. However, to calibrate our methods, we shall also consider the related esters (looking at both the C=O and C—O—C groups) and ethers.

In principle ab initio molecular orbital calculations can describe how the energy of hydrogen bonding varies with the geometry.[6] In practice, however, this is a large project and calculations have normally only been made to determine the energy and geometry of the minima in the potential energy hypersurface. Here we approach the problem by analyzing the large amount of geometrical data for these interactions that is present in the crystallographic literature (and therefore accessible through the Cambridge Crystallographic Data Files, referred to in this article as CCDF[7]).

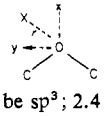
Recently a number of studies[8-24] have shown that intermolecular interactions in crystals show geometrical arrangements of patterns (or motifs) that occur particularly frequently and are presumed to correspond to low-energy areas of this energy hypersurface. With information on the stereochemistry of a large number of interactions of a particular type it should be possible to map not only the minima in the energy surface but also to get an idea of the latitude that might be allowed in deviations from the optimum geometry. Although no absolute energy scale can be attached to such diagrammatic maps they should be useful in helping us to understand the degree of geometrical precision that is important in ligand–macromolecule binding.

Analysis of the Geometrical Variation of Hydrogen Bonds to Oxygen Functionality

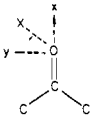
The use of the Cambridge Crystallographic Data File (CCDF)[7] for studying intermolecular interactions is now fairly well established. In the present case there are some additional features that require careful handling and are described below. In particular we have used a more powerful program (GEOSTAT)[24,25] for analyzing geometry than is available with the standard Cambridge package.

It is important to the present study that the chemical nature of small fragments that we wish to examine in the CCDF should be rigorously defined; the ether group, C–O–C, is probably the hardest to treat because it occurs not only in ethers, but also in esters, anhydrides, and ketals/acetals. These are sufficiently chemically different from each other, in spite of the presence of a C–O–C group, that we decided to exclude them from a study of ethers. There are several ways of doing this; all are somewhat laborious and inexact but we chose to restrict our survey of ethers to those with the $-CH_2-O-CH_2-$ group, effectively removing the other unwanted groups above. [Note: Compounds could still have been selected that contained *both* $-CH_2-O-CH_2-$ *and*, say, an ester group; this problem was dealt with by using GEOSTAT (below)]. The selection of this very specific group of ethers has another major advantage in that such ethers are all very similar in shape and size in the region of interest. (As we shall see later, substituents affect our analysis in a manner that makes it harder to extract quantitative results.) By limiting the ethers to those in *rings* we remove the variable of conformation, and thus can study a *sterically* very homogeneous set of compounds. Unfortunately this type of limitation cannot be made for all groups; for the epoxides, which are found fairly infrequently on the file, it is essential to use data for substituted compounds (otherwise we would have almost no data). We must therefore measure the effect that substitution has on the hydrogen-bonding geometry by

(3) Politzer, P.; Daiker, K. C.; Estes, V. M.; Baughman, M. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1978**, *5*, 291.
(4) Mezey, P.; Kari, R. E.; Denes, A. S.; Csizmadia, I. G.; Gosavi, R. K.; Strausz, O. P. *Theor. Chim. Acta* **1975**, *36*, 329–338.
(5) Hopkinson, A. C.; Lien, M. H.; Csizmadia, I. G.; Yates, K. *Theor. Chim. Acta* **1978**, *47*, 97–109.
(6) Allen, L. C. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4701.
(7) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr., Sect. B.* **1979**, *B35*, 2331–2339.
(8) Rosenfield, R. E., Jr.; Parthasarathy, R.; Dunitz, J. D. *J. Am. Chem. Soc.* **1977**, *99*, 4860–4862.
(9) Guru Row, T. N.; Parthasarathy, R. *J. Am. Chem. Soc.* **1981**, *103*, 477–479.
(10) Britton, D.; Dunitz, J. D. *Helv. Chim. Acta* **1980**, *63*, 1068–1073.
(11) Murray-Rust, P.; Motherwell, W. D. S. *J. Am. Chem. Soc.* **1979**, *101*, 4374–4376.
(12) Leiserowitz, L.; Schmidt, G. M. J. *J. Chem. Soc. A* **1969**, 2372–2382.
(13) Leiserowitz, L. *Acta Crystallogr., Sect. B* **1976**, *B32*, 775–802.
(14) Berkovitch-Yellin, Z.; Leiserowitz, L. *J. Am. Chem. Soc.* **1980**, *102*, 7677–7690.
(15) Einspahr, H.; Bugg, C. E. *Acta Crystallogr., Sect. B* **1980**, *B36*, 264–271.
(16) Einspahr, H.; Bugg, C. E. *Acta Crystallogr. Sect. B* **1981**, *B37*, 1044–1052.
(17) Burgi, H. B.; Dunitz, J. D.; Shefter, E. *Acta Crystallogr., Sect. B* **1974**, *B30*, 1517–1527.
(18) Burgi, H. B.; Dunitz, J. D.; Shefter, E. *J. Am. Chem. Soc.* **1973**, *95*, 5065.
(19) Britton, D.; Dunitz, J. D. *J. Am. Chem. Soc.* **1981**, *103*, 2971–2979.
(20) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063–5070.
(21) Chakrabarti, P.; Dunitz, J. D. *Helv. Chim. Acta* **1982**, *65*, 1482–1488.
(22) Rosenfield, R. E. Jr.; Murray-Rust, P. *J. Am. Chem. Soc.* **1982**, *104*, 5427–5430.
(23) Cody, V.; Murray-Rust, P., submitted for publication.
(24) Murray-Rust, P.; Stallings, W. C.; Monti, C. T., Preston, R. K.; Glusker, J. P. *J. Am. Chem. Soc.* **1983**, *105*, 3206–3214.
(25) GEOSTAT (described in ref 30). By use of TEST TORS the chirality of the fragment and labeling of the symmetry-equivalent atoms can be controlled so that the X atom falls in a particular asymmetric unit of the assumed point group. These point groups are $C_{2v}(x)$ for i, ii, iv, and v and $C_s(z)$ for iii, vi, vii, and viii.

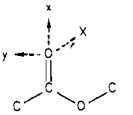


i, C's must be sp^3; $2.4 < r < 3.0$; X constrained to lie in positive z (using TEST TORS)



ii, iii, v, and viii, X constrained to have +y, +z; $2.4 < r < 3.0$

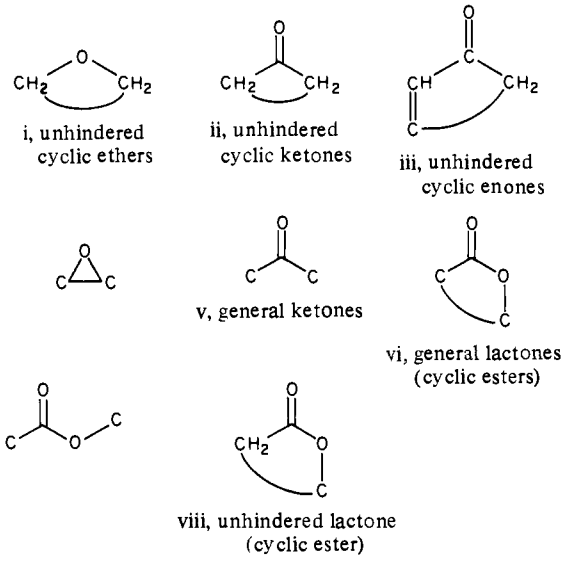
iv, as for i



vi and vii, $2.4 < r < 3.0$; $|\tau|(CCOC) > 150°$ or $< 30°$ (depending on whether the ester is cis or trans); X constrained to be +z

comparing two data sets containing the same oxygen functionality (substituted in one but not in the other).

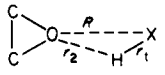
Accordingly we carried out searches to find eight series of crystal structures that contained one (or more) potential hydrogen donors (X–H in I = O–H or N–H) and, respectively, the groups below:



The curved line represents a cyclic compound and C a carbon that may have any number of any type of substituent groups; where H is indicated it may not be replaced by any other group.

For i, ii, iii, iv, and viii there were fairly few data and all entries were retrieved. For v, vi and vii (the "general" ketones, esters, and lactones) there were well over 1000 entries of each, and, to cut this down to a manageable number, numeric data screening (RETRIEVE) was employed. Only data entries fulfilling all of the following criteria were used: (1) Hydrogen atom positions should be reported. (2) No atom heavier than potassium should be present in the structure (so that C, O, N, and H were well located). (3) No reported disorder should be present. (4) The estimated standard deviation of a C–C bond, σ(C–C), should be less than 0.01 Å. (5) The R factor (the agreement between the observed data and those calculated from the proposed model) should be less than 0.10. An application of these criteria resulted in a manageable number of structures.

In this study we had to analyze several thousand crystal structures for potential hydrogen bonds. As far as possible an attempt was made to have this done automatically and accurately. Our search algorithm had to be such that as many hydrogen bonds as possible were retrieved from the data file (CCDF) but that contacts not corresponding to hydrogen bonds ("noise") were excluded. One possible way to do this would have been to search for fragments of the sort shown in IV with constraints on R, r_1



IV

and r_2.[26] There are several drawbacks to this procedure, however. In a number of structures (particularly those determined several years ago) coordinates for hydrogen atoms (particularly those on hydroxyl groups) are not reported, although the structure was described as containing a hydrogen bond. To reject these might cut down the number of retrieved structures to a level where useful conclusions are impossible. Moreover, searching for the fragment above is not trivial, since a large number of symmetry-related molecules must be considered. For a molecule such as a steroid (containing more than 40 atoms including hydrogen atoms) the introduction of symmetry-related molecules can make the problem too large for the present version of GEOSTAT. Accordingly we have assumed that an intermolecular O···X (X = N, O) contact less than 3 Å represents a hydrogen bond.[27] At first sight this method of searching for a hydrogen bond might seem to be very dangerous, since the conventional van der Waals radii of oxygen and nitrogen atoms are 1.4 and 1.5 Å, respectively. In fact, however, relative to hydrogen bonds, nonbonded intermolecular O···O contacts of 2.8 Å in organic structures are very infrequent indeed (and correspondingly even more infrequent for O···N). We have therefore checked manually through 100 of our retrieved O···O and O···N "hydrogen bonds" to see whether they really were interpretable as O···HX bonds. In only one case (which was in any case manifested as an obvious outlier in a scatterplot) was this not so. Since we shall be concerned with a general, deliberately smeared, picture, undetected noise at approximately 1% or 2% will not affect our conclusions at all and therefore we feel that this is a valid method of searching for hydrogen bonds in a massive data set. (In any case we automatically examine any outliers in considerable detail. As a further aid in checking hydrogen bonds we can fairly quickly examine, using the LABel option in GEOSTAT, the chemical nature of the putative hydrogen donor; this procedure will show attached hydrogen atoms if reported). Additionally most hydrogen bonds are fairly linear[28] (85% of hydrates have O···H-O greater than 150°); thus, for an O···H–X group (O···X = 2.8 Å) with an O···H–X angle of 150° (i.e., fairly bent), the angle between the O···H and O···X vectors is only 10° (Figure 1), and the distance of the hydrogen atom from the O···O vector is only about 0.3 Å. Since we shall deliberately "smear" the oxygen atom positions by at least this amount, the error can readily be tolerated. Also, the O–H bond length determined by X-ray diffraction is usually systematically shorter (by about 0.1 Å) than the true internuclear distance, so that the calculated O···H–X geometry is of dubious use (particularly since the creators of the CCDF find many hydrogen atom coordinates incorrectly reported).

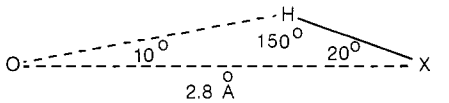


Figure 1. Demonstration that for strong O···H–X bonds the O···X vector is fairly close to the O···H vector.

GEOSTAT was therefore used in this way to find all cases in data sets i–viii for which there was a potential hydrogen bond, defined as an O···X contact less than 3.0 Å. The control for GEOSTAT was formulated so that the proton donor fell in the appropriate asymmetric unit of the particular point group of the fragment [i.e., C_{2v} for i, ii, iv, and v and C_s for iii, vi, vii, and viii]. Unfortunately very few hydrogen-bonded fragments of type viii were found and this group was abandoned. Of the remaining esters, almost all formed hydrogen bonds to the carbonyl group rather than to the C–O–C group, presumably because of their relative proton donor powers. All discussion of ester hydrogen bonding is restricted to those that use the carbonyl group.

Accordingly, therefore, the appropriately hydrogen-bonded fragments in sets i–vii were retrieved and the geometrical parameters describing hydrogen bonding were compared. Although the results can be given as a series of tables of appropriate proton distances and angles, a much more graphic presentation[27] of the results is given if all the fragments are superimposed (with common axes) to form a giant "molecule" (which can be displayed with standard graphics software). Appropriate symmetry operations are applied, so that the "molecule" has the same symmetry as the functional group. Our results are shown in Figure 2i–vii (two views of each scatterplot). As can be seen the distribution of points is never isotropic and shows that in all cases hydrogen bonding occurs *preferentially in certain directions*. On the other hand, this

(26) Henceforward we shall use X to denote O or N, the atom in the proton donor. Distances: R = O···X, r_1 = X–H, r_2 = O···H where X = N, O.

(27) Pauling, L. "The Nature of the Chemical Bond", 3rd ed.; Cornell University Press: Ithaca, NY, 1960.

(28) Olovsson, I.; Jönsson, P.-G. in "The Hydrogen Bond"; Schuster, P., Zundel, G., Sandorfy, C., Eds.; North-Holland Publishing Company: Amsterdam, New York, Oxford, 1976; Vol. II, Chapter 8, pp 393–456.

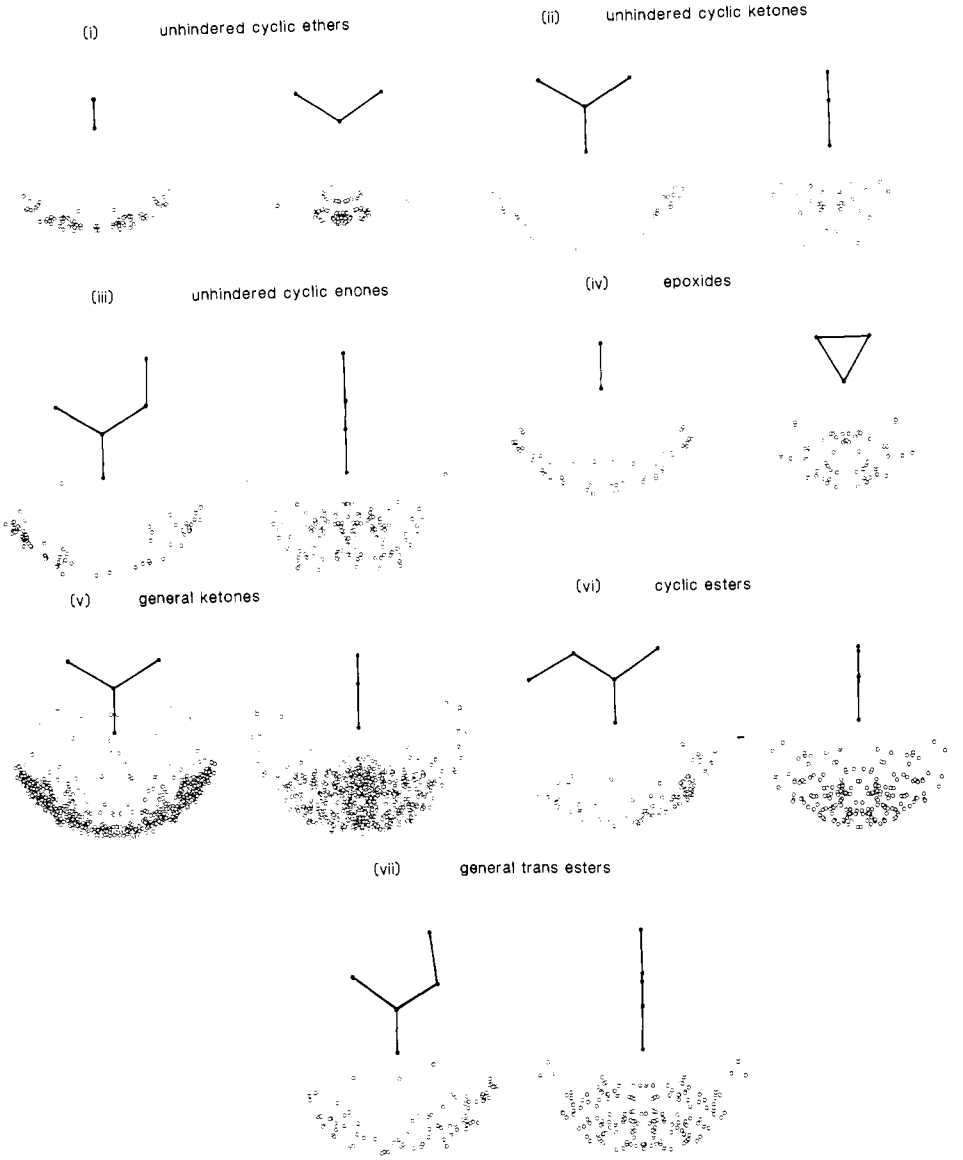


Figure 2. Scatterplots (viewed in two perpendicular directions) illustrating the positions of hydrogen donor atoms X (in H–X) involved in hydrogen bonding to the functional oxygen atom that is illustrated: (i) unhindred cyclic ethers, (ii) unhindered cyclic ketones, (iii) unhindered cyclic enones, (iv) epoxides, (v) general ketones, (vi) cyclic esters, and (vii) general trans esters.

anisotropy is not always so pronounced that the distribution consists of isolated, tight clusters (as it did for halide ions about $RNMe_3^+$).[22]

It is difficult to compare Figure 2i–vii objectively, since they contain different numbers of points. Moreover, there are problems of perception, since the eye tends to give undue prominence to outliers, particularly if they might seem to link large clusters. We have therefore devised a technique to enhance the importance of the clustering in these distributions, and to attempt to represent this in a semiquantitative manner, and to reduce the importance of rare, isolated, outliers. We transform a scatterplot such as Figure 2 into a continuous distribution by convoluting point atomic density with a diffuse (probability) density.[29] For convenience we have used spherical atomic electron density functions with a parameterizable temperature factor. The resulting three-dimensional pseudodensity function is then contoured by standard methods (e.g., "chicken-wire" plots[30]) to give a three-dimensional contour plot. When viewed on a graphics device in stereo this gives a powerful way of representing the distribution (Figure 3). Unfortunately it is difficult to represent these quantitatively in a paper and we have therefore used two-dimensional projections of the density (which are very similar to sections through a plot, see Figures 4 and 5). The contour lines represent contours at different positional heights in space with respect to the plane of view (shown in stereo pairs for ketones, ethers, and epoxides in Figure 3 and as the superposition of three levels of contouring in Figure 4). We use a temperature factor of 10, which for point atoms would mean that the density between two atoms 0.7 Å apart was very roughly constant; for two atoms 1.5 Å away the density drops considerably at the midpoint. Though this value might be refined in later studies it gives reasonable results here (comparison of Figures 2 and 4 shows it to be very effective in removing outliers). Before the contouring is done, the electron density values are divided by the number of data points (O···H—X interactions) that contributed to their calculation. Thus all plots are scaled to approximately the same contour level. We use three contours that seem to work quite well: (a) 0.5 × electron density (in e/Å^3) × number of contributing structures (data points), (b) 0.5 × (a), and (c) 0.25 × (a). These contours are superimposed as shown in Figure 4 for the systems under question and as shown in simplified versions in Figure 5.

It is clear that contour maps of type (c) give a reasonable idea of the general spread of points. Isolated outliers are not normally seen, except for the lowest contour in cases where the number of

(29) Carrell, H. L. GENAT and GENMAP are programs from the Institute for Cancer Research, Philadelphia, PA 19111, 1982.
(30) Murray-Rust, P.; Rosenfield, R. E., Jr.; Meyer, E.; Carrell, H. L., submitted for publication.

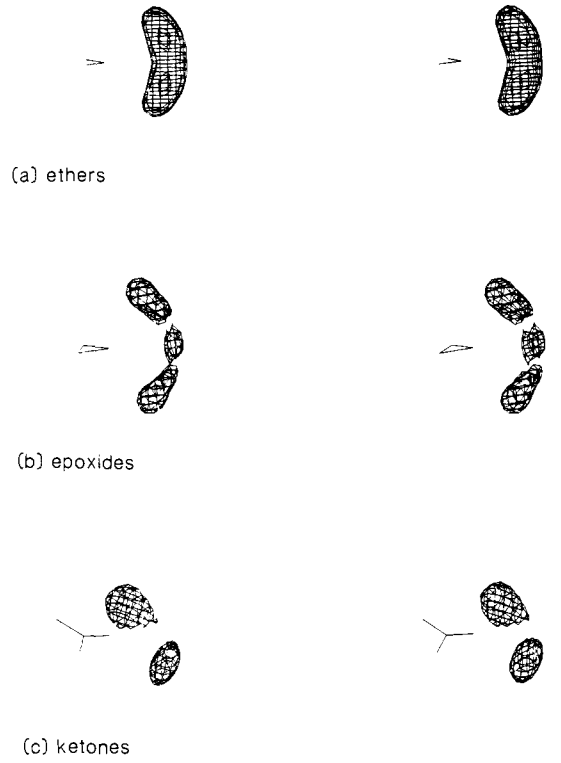


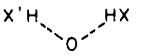
Figure 3. Stereo pairs of three-dimensional "chicken-wire" contours (at 0.25 e/Å^3) for X–H···O interactions in (a) ethers, (b) epoxides, and (c) ketones, showing that the maxima lie near directions expected for lone pairs on oxygen atoms.

data points is relatively small. By contrast contour a represents a very "sharp" picture of the most concentrated regions of the scatterplot (Figure 2). (Because of the sharpness it is always possible that random noise might produce a small peak above the contour level, although we see no evidence here.) Contours (b) seem to give a possible semiquantitative measure of differences in shape between different distributions. For a very compact distribution [as for the halides[22] round RNH_3^+ or ether···HX in the present case (Figures 2i, 4i, 5i)] there is very little difference among i to iii. For a diffuse distribution, such as the epoxides [Figures 2iv, 4iv, 5iv] the different emphasis of each contour is striking.

Discussion

All systems show the largest concentration of hydrogen-bonded H–X lying in the directions that we conventionally draw the lone pair orbitals. For ether (i) and epoxides (iv) this is in a plane perpendicular to the C–O–C plane, whilst for ketones (ii, iii, v) and esters (vi, vii) this is in the CC(=O)Y (Y = C, O) plane. Whilst perhaps not too unexpected this has generally been a supposition rather than a hard experimental fact. The consistency of the diagrams is especially gratifying; in particular the similarity between i and ii and among iii, iv, and v is illustrated. We can conclude that *there are directional features in the approach of proton donors to lone pairs and that crystallographic studies can reveal these.*

It is more difficult to assess the importance of the *variation* in angular geometry (both *within* and *between* systems). For any particular crystal the resultant structure is a complex compromise among many independent intermolecular forces (and often some intramolecular distortion as well). Here we assume (axiomatically) that the set of structures studied are different enough that there is not a common molecular feature which biases structures to pack in a certain way.[31] On this assumption, then, the deviations from the ideal geometry for an intermolecular contact are distributed randomly. Elsewhere[31] we have postulated the same principle for "soft" intramolecular distortions and found it to fit the data well. By calibrating probabilities of distortion with known energies, we suggest[32] the formula: probability of distortion = exp (–energy of distortion/0.15 kcal mol^{-1}). We have no energy calibration in the present case, but should this become available, then it ought to be possible to put energy values on the contours in Figure 4. If a group forms more than one hydrogen bond there can be steric hindrance between the proton donors (XH, X'H) that will restrict the space available to each. Because of this steric hindrance it may be almost impossible for either of the groups to coordinate along the C_2 axis and hence a somewhat spurious "lone-pair-like" geometry, V, could be produced. In fact very few of the fragments



V

correspond to doubly coordinated groups: 0/26 for the epoxides; 0/40 for the ethers; 18 pairs/286 for the ketones; 5 pairs/96 for cis cyclic esters, 2 pairs/67 for enones, 2 pairs/81 for trans cyclic esters. We do not, therefore regard this effect as seriously contaminating our results.

The Importance of Neighboring Substituents

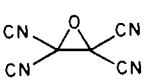
If there are substituents close to the proton acceptor they may be bulky enough to interfere with the proton donor (or *its* close substituents). The number and size of the substituents varies considerably and is not really quantifiable. This variation will result in "smearing" of the distribution and is clearly shown by the comparison of the unhindered ketones (ii and the larger, general set (v). The latter is more diffuse, presumably because optimal contacts can be made much less frequently. It is useful to note that, despite this smearing the peaks of the distribution are observable in the same positions as in ii (where the effect is amost unobservable in the scatterplot).

A similar effect can be seen in the distribution of proton donors hydrogen bonded to the C=O group of esters (vi, vii). The variation in conformation clearly affects the average steric hindrance and hence the distribution of X. Here again, however, the positions of the peaks can still be observed in the highest contour level.

Ethers and Epoxides

As expected, for ethers and epoxides the proton donor density corresponds closely to the direction of the lone pairs. The densities in these directions, however, do not seem to be clearly separated and, for the ethers, we get a picture of almost continuous density. Although the epoxide density is affected by steric smearing there is some indication of local concentrations of this density at about ±60° off the C_2 axis. As shown in Figure 6, tentative evidence (based on our present data sets) suggests that the distribution of proton donors, X, hydrogen bonded to ethers (with a C–O–C bond angle of 110–120°) is tighter than the distribution of donors bonded to epoxides (with a C–O–C angle of 60° and a distribution that is broader both in- and out-of-plane and more like that to ketones). The idea that the angle between the lone pairs is somewhat larger in epoxides than in ethers[33] seems to be supported by our data.

In a classic investigation of valence electron density, the epoxide lone pair region for tetracyanoethylene epoxide, TCNE, VI, was



VI

mapped by X–N methods.[34,35] This is one of the earliest analyses

(31) This would not be true, for example, if all our structures were, say, planar molecules, or differed only by being linear homologues. There is not an obvious feature of molecular shape in our data set which might force a recurrent type of packing.

(32) Murray-Rust, P. in "Molecular Structure and Biological Activity"; Sutton, L. E., Truter, M. R., Eds.; Elsevier Biomedical: New York, Amsterdam, Oxford, 1982; pp 117–133.

(33) Bowers, M. T.; Goldwhite, H.; Vertal, L. E.; Douglas, J. E.; Kollman, P. A.; Kenyon, G. L. *J. Am. Chem. Soc.* **1980**, *102*, 5151–5157.

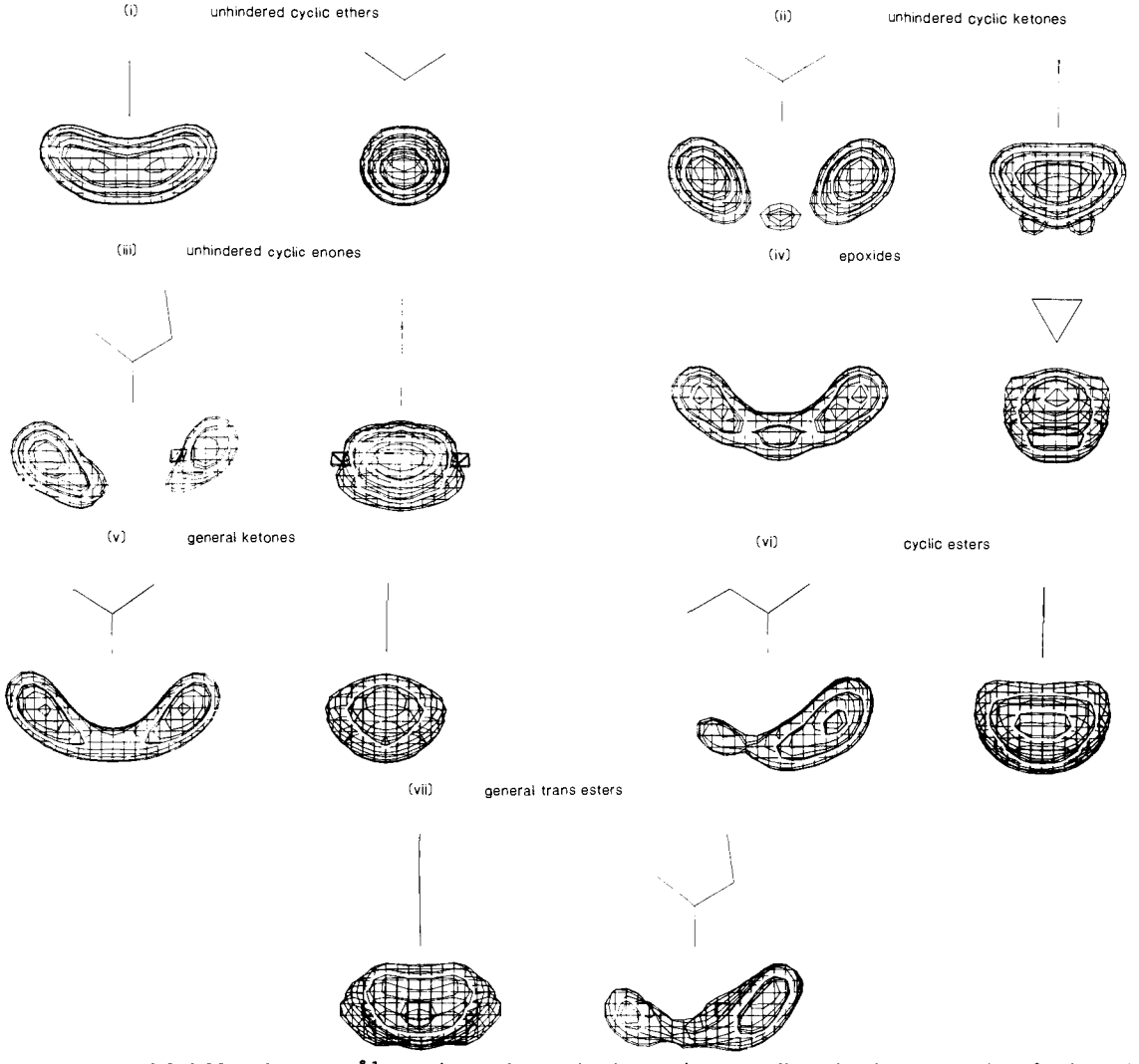
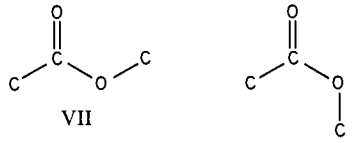


Figure 4. Contour maps, at 0.5, 0.25, and 0.125 e/Å^3, superimposed on each other to give a two-dimensional representation of a three-dimensional plot (such as is shown for one contour in Figure 3).

and is not as well resolved as modern X–N and X–X maps,[36-42] but the lone pair density is clearly seen to be continuous (rather than as two isolated lone pairs) and looks extremely similar to Figure 4.

Ketones and Esters

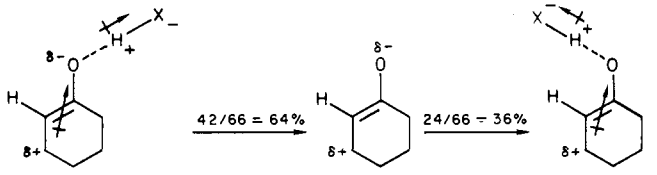
There are two general conformations of esters, shown in VII and VIII. In all cases (ketones and esters) we find the same



VII

VIII

general picture as that found for the sp^2 oxygen atoms, a distribution concentrated in the plane of the presumed lone pairs and [despite steric smearing in v–vii] with concentrations of density along the sp^2 vectors. The lone pairs in the unhindered ketones (ii and iii) are strikingly well resolved. These latter (unhindered cyclic enones) are of particular interest to use since it is a presumed pharmacophore in the A ring of the sex hormones (progesterone, testosterone, etc.). This shows a somewhat unexpected asymmetry with nearly twice as many contacts (42 contacts) on the CH_2 side than the CH side (24 contacts). The difference is significant at the 98% confidence level and seems to reflect a genuine preferred direction of hydrogen bonding. This is not easily explained in steric terms (since the shortest X···H contacts to the ring substituents are about 2.8 Å and the CH group has a similar bulk to that of the CH_2 group). It is possible that the dipole (or more likely the induced dipole) of the pharmacophore is substantially asymmetric, being represented by structures such as VIII and IX. Presumably



VIII

the more nearly the two dipoles are aligned, the more favored the interaction is; thus the angle between dipoles in VIII is more nearly 180° that in IX. [Note (a word of caution): All except one of the compounds contributing to this effect are steroids. It seems very unlikely, however, that this selection of data is causing this effect, since there is great variability in the packing of steroids.][43]

(34) Matthews, D. A.; Swanson, J.; Mueller, M. H.; Stucky, G. D. *J. Am. Chem. Soc.* **1971**, *93*, 5945–5955.
(35) Matthews, D. A.; Stucky, G. D. *J. Am. Chem. Soc.* **1971**, *93*, 5955.
(36) Helmholdt, R. B.; Vos, A. *Acta Crystallogr., Sect. A* **1977**, *A33*, 456–465.
(37) Stevens, E. D. *Acta Crystallogr., Sect. B* **1978**, *B34*, 544–551.
(38) Tanaka, K. *Acta Crystallogr., Sect. B* **1978**, *B34*, 2487–2494.
(39) Kvick, A.; Koetzle, T. F.; Stevens, E. D. *J. Chem. Phys.* **1979**, *71*, 173–179.
(40) Savariault, J.-M.; Lehmann, M. S. *J. Am. Chem. Soc.* **1980**, *102*, 1298–1303.
(41) Stevens, E. D.; Rys, J.; Coppens, P. *J. Am. Chem. Soc.* **1978**, *100*, 2324–2328.
(42) Stevens, E. D.; Coppens, P. *Acta Crystallogr., Sect. B* **1980**, *B36*, 1864–1876.

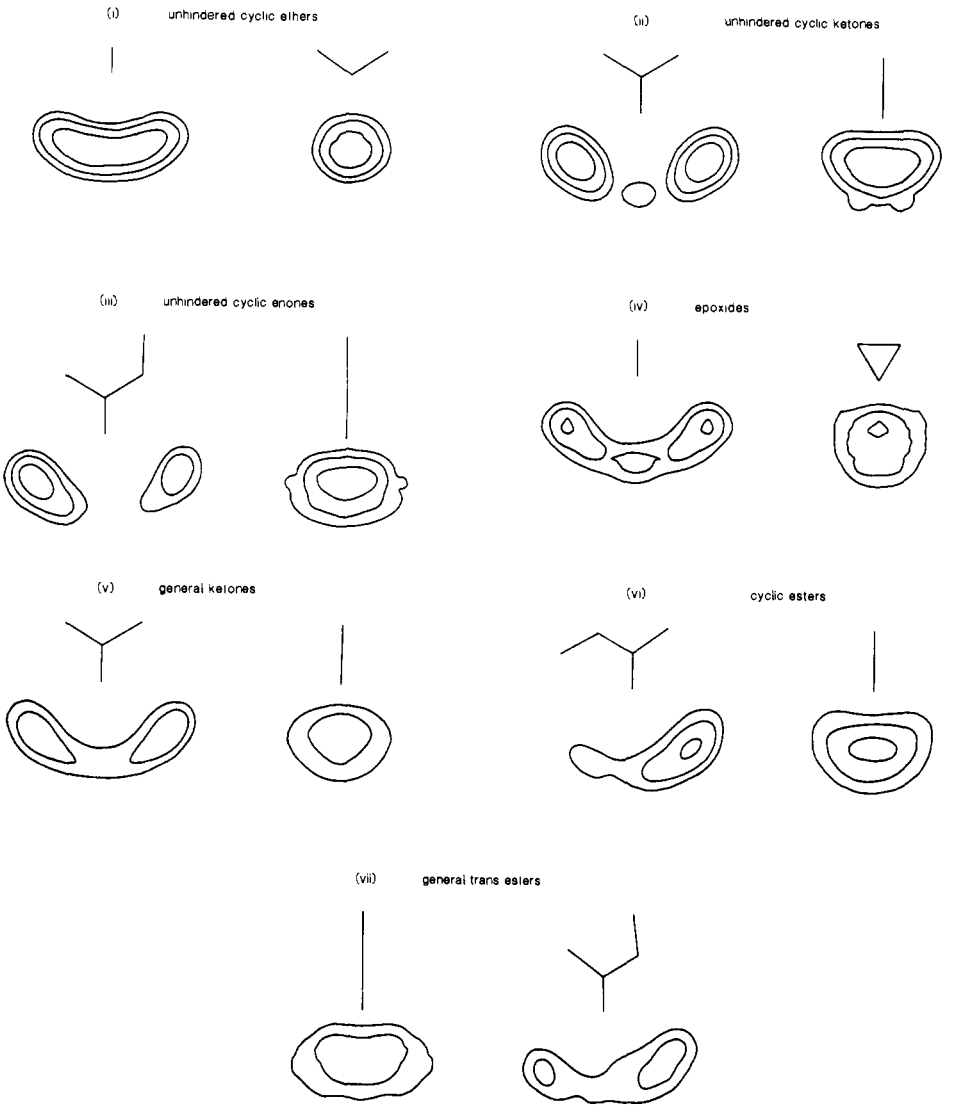


Figure 5. Simplification of Figure 4 to show approximate two-dimensional contours of the density of the hydrogen bond donor, X, around a functional oxygen atom.

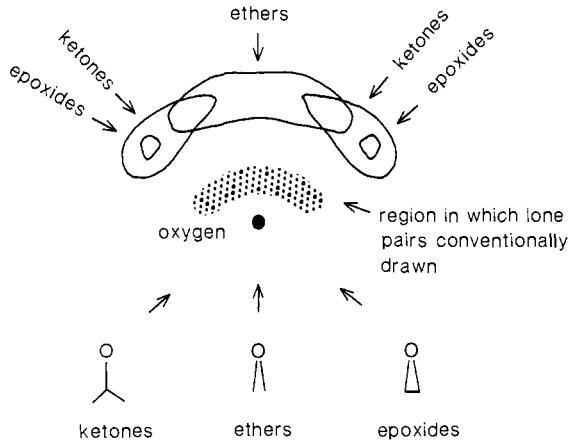


Figure 6. Relative distribution of proton donors, X, hydrogen bonded to ethers, ketones, and epoxides showing the tighter distribution for ethers from experimental X-ray crystallographic data from the CCDF.

Use of Intermolecular Contacts for Modeling Ligand Binding: The Antiphore and Its Biological Implications

Much effort is being put into trying to find common features of a group of (often fairly dissimilar) molecules that are presumed to bind to a common receptor (usually of unknown structure). For example, our interests that led to these studies of interactions was in chemical carcinogenesis, particularly by polycyclic aromatic hydrocarbons *via* their epoxides and diol epoxides,[44,45] and their interactions with nucleic acids and metabolizing enzymes. Most of this effort at finding common features has so far gone into trying to evaluate the similarities in molecular shape, particularly the van der Waals envelope and the common (or excluded) volume. The addition of information on intermolecular interactions, such as that provided here, is clearly valuable and some previous studies have tried to map the common features of the crystal environment of a set of related biologically active molecules.[22,24] In the absence of the structure of a known receptor/binding macromolecule we can try, using these known intermolecular patterns, to map the part of the macromolecule that binds the pharmacophore.[46,47]

(43) Duax, W. L., Norton, D. A., Eds. "Altlas of Steroid Structure"; Plenum: New York, Washington, London, 1975; Vol. 1.

(44) Glusker, J. P. in "Polycyclic Aromatic Hydrocarbons and Cancer"; Ts'o, P. O. P., Gelboin, H., Eds.; Academic Press: New York, 1982; Vol. 3, p 61.

(45) Sims, P.; Grover, P. L. *Nature (London)* **1974**, *252*, 326.

(46) The pharmacophore is defined as the three-dimensional arrangement of functional groups on a substrate/antigen/hormone/drug that are essential for both recognition (by binding) and activation of the enzyme/antibody/receptor/receptor. Thus, by analogy to the word chromophore, it consists of the significant parts of a molecule involved in pharmacological activity. It is necessary to identify these essential function groups and also to determine the three-dimensional arrangement of these groups that leads to activity. For more information and examples see ref 47.

(47) Olson, E. C., Christoffersen, R. E., Eds. *ACS Symp. Ser.* **1979**, 112.

By analogy with the term anticodon we suggest calling this the *antipharmacophore* (shorted by convenience to *antiphore*). Thus Figure 3 is (low resolution) maps of estimates of the distribution of XH groups in the antiphores of molecules binding the appropriate groups such as ketones, epoxides, ethers, and esters. If we assume that the geometry of the receptor–ligand interaction is optimized when the ligand is bound to the most biologically active ligand, then we can suggest that the highest contour levels in the maps in Figures 4 and 5 represent the preferred directions of hydrogen bonding from the biological receptor (as hydrogen-bond donor) to the ligand.

Conclusion

For most common oxygen-containing functional groups there are hundreds of reported structures on file (CCDF). It is straightforward to compare their hydrogen-bonding patterns, both in terms of frequency and in terms of distance and angular distribution, as shown for selected examples in Figure 6. Beside giving a useful, practical addition to our understanding of the geometry of ligand–macromolecule binding, they provide maps of lone pair density that can be compared with those produced by X-ray crystallography (at high resolution using, for example, X–N maps) or by ab initio calculations. As more of the types of maps illustrated in this article become available, further variations in the stereochemical specificity of binding of differing functional groups should be quantifiable.

Acknowledgment. We thank Drs. H. L. Carrell, Philip Coppens, and Richard E. Rosenfield, Jr. for helpful discussions. Diagrams were drawn with the computer programs VIEW,[48] DOCK,[49] GENAT,[29] and GENMAP.[29] This work was supported by Grants BC-242 from the American Cancer Society and CA-10925, CA-22780, CA-06927, and RR-05539 from the National Institutes of Health and by an appropriation from the Commonwealth of Pennsylvania.

Registry No. I, 75-21-8; VI, 3189-43-3; cyclohexanone, 108-94-1; oxygen, 7782-44-7.

Supplementary Material Available: Tables of references to the structures that contain the O···H–X interactions described in this article (46 pages). Ordering information is given on any current masthead page.

(48) Carrell, H. L. VIEW is a program from the Institute for Cancer Research, Philadelphia, PA 19111, 1977.
(49) Badler, N.; Stodola, R. K.; Wood, W. DOCK is a program from the Institute for Cancer Research, Philadelphia, PA 19111, 1982.

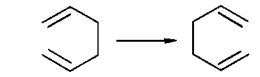
The Anionic Oxy-Cope Rearrangement: Structural Effects in the Gas Phase and in Solution

Melvin D. Rozeboom,[1] Jeffrey P. Kiplinger, and John E. Bartmess*

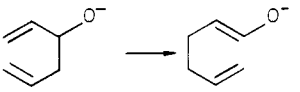
Contribution from the Department of Chemistry, Indiana University, Bloomington, Indiana 47405. *Received May 25, 1983*

Abstract: The Cope rearrangement of 1,5-dienes with an anionic oxygen on C-3 has been shown to proceed in the gas phase, using ICR spectrometry, for the case where the precursor is a tertiary alcohol. The secondary alcohol, if it reacts at all, is slower than the ICR time scale. A similar rate difference is observed in THF or Me_2SO solvents. The rate variation is ascribed to an intrinsic structural effect, not to differential ion pairing or solvation effects.

The Cope rearrangement (eq 1) has been shown to have a



(1)



(2)

half-life on a geological time scale at room temperature.[2] If an anionic substituent is present on C-3, however, the rearrangement is accelerated by a factor of 10^{10}–10^{17},[3] so as to proceed in a few minutes to hours. As a result, this reaction has been widely used in synthetic transformations[4] and is the basis for the remarkable acceleration probed by various methods.[5] Reactions of this type can be analyzed in terms of the bond-breaking or bond-making extremes, as detailed in Figure 1, in order to shed further light on the mechanism. In the case of the anionic oxy-Cope rearrangement, simple inspection reveals no reason to expect special acceleration for the bond-making pathway, since a β-alkoxide functionality should not have any specific interaction with the radical site. Goddard and co-workers have examined the thermochemistry of the bond-breaking pathway BB_1 using molecular orbital calculations and have stated that the alkoxide functionality should reduce the C–C bond strength considerably in such species.[5] The calculated decrease of the C–H bond strength in $H–CH_2O^-$ relative to $H–CH_2OH$ of 16.5 kcal/mol was found to be comparable to the ca. 18-kcal/mol reduction in the energy of activation experimentally observed for reaction 2.[3,4f]

There is another possible bond-breaking pathway that must be considered, BB_2 in Figure 1, where the electron resides on the allyl moiety rather than the enone. The electron affinity (EA) of the allyl radical in the gas phase is known to be 12.7 kcal/mol.[6] While no experimental value for the EA of acrolein is available, its radical anion is bound,[7a] but undoubtedly less so than for cinnamaldehyde with EA = 19 kcal/mol.[7b] This bracketed value of 10 ± 10 kcal/mol is consistent with the anion not being observed in various

(1) Current Address: Department of Chemistry, University of Nebraska, Lincoln, NB 68588.
(2) (a) Hurd, C. D.; Polack, M. A. *J. Org. Chem.* **1938**, *3*, 550; (b) Levy, H.; Cope, A. C. *J. Am. Chem. Soc.* **1944**, *66*, 1684. (c) Doering, W. von E.; Toscano, V. G.; Beasely, G. H *Tetrahedron* **1971**, *27*, 5299.
(3) Evans, D. A.; Golub, A. M. *J. Am. Chem. Soc.* **1975**, *97*, 4765.
(4) (a) Wilson, S. R.; Mao, D. T.; Jernberg, K. M.; Ezmirly, S. T. *Tetrahedron Lett.* **1977**, *2559*. (b) Still, W. C. *J. Am. Chem. Soc.* **1977**, *99*, 4186. (c) Evans, D. A.; Baillargeon, D. J.; Nelson, J. V. *Ibid.* **1978**, *100*, 2242. (d) Evans, D. A.; Golob, A. M; Mandel, N. S.; Mandel, G. S. *Ibid.* **1978**, *100*, 8170. (e) Wilson, R. M.; Rekers, J. W.; Packard, A. B.; Elder, R. C. *Ibid.* **1980**, *102*, 1633. (f) Paquette, L. A.; Crouse, G. D.; Sharma, A. K. *Ibid.* **1980**, *102*, 3792. (g) Tice, C. M.; Heathcock, C. H. *J. Org. Chem.* **1981**, *46*, 9. (h) Levine, S. G.; McDaniel, R. L. *Ibid.* **1981**, *46*, 2199. (i) Mikami, K.; Taya, S.; Nakai, T.; Fujita, Y. *Ibid.* **1981**, *46*, 5445. (j) Clive, D. L. J.; Russell, C. G.; Suri, S. C. *Ibid.* **1982**, *47*, 1632. (k) Martin, S. F.; White, J. B.; Wagner, R. *Ibid.* **1982**, *47*, 3190.
(5) Steigerwald, M. L.; Goddard, W. A., III; Evans, D. A. *J. Am. Chem. Soc.* **1979**, *101*, 1994.
(6) Zimmerman, A. H.; Brauman, J. I. *J. Am. Chem Soc.* **1977**, *99*, 3565.
(7) (a) Jordan, K. D.; Burrow, P. D. *Acc. Chem. Res.* **1978**, *11*, 341. (b) Wentworth, W. E.; Kao, L. W.; Becker, R. S. *J. Phys. Chem.* **1975**, *79*, 1161.

0002-7863/84/1506-1025$01.50/0