

the similar stability of $\text{VMo}_5\text{O}_{19}^{3-}$ toward water and heat suggest that the instability of anion **1** may arise in part from relatively high negative charge density on the surface of its $\text{Mo}_5\text{O}_{18}^{6-}$ subunit. Apparently the $[(\text{C}_5\text{H}_5)\text{Ti}^{\text{IV}}]^{3+}$ group, containing a tetravalent metal center, cannot withdraw negative charge from the $\text{Mo}_5\text{O}_{18}^{6-}$ unit as effectively as the $[\text{OV}^{\text{V}}]^{3+}$ group which employs a pentavalent metal center.⁹ The negative charge density on the $\text{Mo}_5\text{O}_{18}^{6-}$ unit in anion **1** can be lowered by addition of an $\text{Mn}(\text{CO})_3^+$ unit, resulting in a hydrolytically stable adduct, anion **3**. The mechanism by which charge reduction occurs throughout the entire anion upon adduct formation is revealed by detailed comparison of the structures of anion **1** and its MoO_2Cl^+ adduct, anion **2**.

Single crystals of the $(n\text{-C}_4\text{H}_9)_4\text{N}^+$ salts of anions **1** and **2**, obtained in the form of their CH_2Cl_2 solvates from $\text{CH}_2\text{Cl}_2/(\text{C}_2\text{H}_5)_2\text{O}$ solution, were utilized for single-crystal X-ray diffraction structure determinations.¹⁰ Anion **1** (see Figure 1) is structurally related to the $\text{Mo}_6\text{O}_{19}^{2-}$ anion by substitution of a $[(\text{C}_5\text{H}_5)\text{Ti}^{\text{IV}}]^{3+}$ group for an $[\text{OMo}^{\text{VI}}]^{4+}$ unit. The $\text{Mo}_6\text{O}_{19}^{2-}$ structure has been shown to approximate full O_h symmetry in the solid state,¹¹ with small but statistically significant differences in Mo-O bond lengths constituting the major departures from idealized symmetry. Replacement of a $[\text{OMo}^{\text{VI}}]^{4+}$ unit in $\text{Mo}_6\text{O}_{19}^{2-}$ by a $[(\text{C}_5\text{H}_5)\text{Ti}^{\text{IV}}]^{3+}$ group to give **1** necessarily reduces the anion symmetry and has a pronounced structural effect on the $\text{Mo}_5\text{O}_{18}^{6-}$ moiety. Similar distortions are observed in $\text{Na}_6\text{H}_2[\text{Ce}(\text{W}_5\text{O}_{18})_2]\cdot 30\text{H}_2\text{O}$.¹² Even though **1** is not required to possess any rigorous crystallographic symmetry, it approximates rather closely C_{4v} site symmetry (if the local symmetry of the C_5H_5 ring is ignored), with O_E , Mo_5 , O_A , Ti, and the cyclopentadienyl center of gravity ideally lying on the fourfold axis. The $[(\eta^5\text{-C}_5\text{H}_5)\text{Ti}(\text{Mo}_5\text{O}_{18})\text{MoO}_2\text{Cl}]^{2-}$ anion (**2**) (see Figure 2) can be derived from **1** by bonding an MoO_2Cl^+ unit to three of its doubly bridging oxygens. Although such a species could ideally approximate C_s - m symmetry if the Cl atom was located in the pseudo mirror plane of **2** determined by Ti, Mo_5 , Mo_6 , O_A , O_{B1} , O_{B3} , O_E , and the cyclopentadienyl center of gravity, one of the oxygen atoms (O_{G1}) lies in this plane instead. Since the $\text{Mo}_6\text{-O}_{G2}$ bond has multiple bond character while the $\text{Mo}_6\text{-Cl}$ bond does not, the symmetrical bonding patterns observed in **1** will be perturbed, with those bonds closest to Mo_6 being affected the most. The bonds to the six atoms closest to each molybdenum in **2** show the normal trends of long bonds being trans to short bonds. Thus the $\text{Mo}_6\text{-O}_{C1}$ bond length is more than 0.24 Å shorter than $\text{Mo}_6\text{-O}_{C2}$ [2.021 (8) vs. 2.263 (8) Å]. These large differences in $\text{Mo}_6\text{-O}_C$ bonds would be expected to affect the remaining two Mo-O_{C2} bonds relative to the remaining Mo-O_{C1} bonds. Thus $\text{Mo}_5\text{-O}_{C1}$ is more than 0.10 Å longer than $\text{Mo}_5\text{-O}_{C2}$

[2.083 (7) vs. 1.982 (8) Å] and $\text{Mo}_1\text{-O}_{C1}$ is more than 0.06 Å longer than $\text{Mo}_2\text{-O}_{C2}$ [2.107 (8) vs. 2.037 (9) Å]. As one progresses further from Mo_6 , the distortion in bonding parameters from idealized C_s - m symmetry is much less pronounced. Nonetheless, a systematic pattern of trans bond length alternation at molybdenum is observed for sequences of bonds to the doubly bridging OMO_2 oxygen atoms. For the O_B -type oxygens the following sequence of bond lengths, averaged in accord with approximate C_s - m symmetry, is observed: $O_{B1}\text{-Mo}_2$, 2.014 (7,1,1,2) Å;¹³ $\text{Mo}_2\text{-O}_{B2}$, 1.870 (8,7,7,2) Å; $O_{B2}\text{-Mo}_3$, 1.971 (8,15,15,2) Å; and $\text{Mo}_3\text{-O}_{B3}$, 1.919 (8,6,6,2) Å. Analogous considerations apply to bonding at O_{C3} and O_{C4} . Note, however, that bond alternation is not observed at titanium.

Acknowledgment. W.G.K. acknowledges the National Science Foundation for support of this research.

Supplementary Material Available: Tables of atomic positional and thermal parameters (10 pages). Ordering information is given on any current masthead page.

(13) The first number in parentheses following an averaged value of a bond length or angle is the root-mean-square estimated standard deviation of an individual datum. The second the third numbers, when given, are the average and maximum deviations from the averaged values, respectively. The fourth number represents the number of individual measurements which are included in the average value.

The First General Index of Molecular Complexity

Steven H. Bertz

Bell Laboratories,
Murray Hill, New Jersey 07974
Received July 18, 1980

Synthetic chemists have been defining a "complex molecule" in the way that many people define art: they know it when they see it. While the features which contribute to the complexity of a molecule have been discussed,¹⁻⁴ no unified index has been formulated which takes into account the size, symmetry, branching, rings, multiple bonds, and heteroatoms characteristic of a complex molecule. This communication shows how concepts from graph theory and information theory can be combined to create the first general index of molecular complexity.

A molecule is commonly represented by its *skeletal molecular graph*, M' , (mathematically,⁵ the union of a set of points, symbolizing atoms other than hydrogen, and a set of lines, symbolizing bonds). Its properties can then be expressed in terms of *graph-theoretical invariants*, N_{ij} , which Gordon has defined as "the number of distinct ways in which skeleton i can be cut out of skeleton j ".⁶ The simplest such invariant which takes into account both points and lines is N_{2j} , the number of ways that propane can be cut out of a saturated hydrocarbon, which has been used successfully as a branching index. A more general approach which can be extended to unsaturated systems is based upon η , the number of *connections*, defined as the number of pairs of adjacent lines. For a saturated hydrocarbon, the number of connections equals N_{2j} . Thus, ethane has no connections; propane has one; isobutane, three; and neopentane, six.

In order to include multiple bonds in this scheme, it is necessary to show how "pairs of adjacent lines" are counted in a *multigraph* (a graph with more than one line connecting a pair of points).

(1) Corey, E. J.; *Pure Appl. Chem.* **1967**, *14*, 19-37. *Q. Rev., Chem. Soc.* **1971**, *25*, 455-482.

(2) Hendrickson, J. B. *Top. Curr. Chem.* **1976**, *62*, 49-172.

(3) Woodward, R. B. "Perspectives in Organic Chemistry"; Todd, A., Ed.; Interscience: New York, 1956; p 160. *Pure Appl. Chem.* **1968**, *17*, 519-547.

(4) Ireland, R. E. "Organic Synthesis"; Prentice-Hall: Englewood Cliffs, NJ, 1969; p 100-116.

(5) Harary, F. "Graph Theory"; Addison-Wesley: Reading, MA, 1969. The symbol M is reserved for the full molecular graph, including hydrogen.

(6) Gordon, M.; Kennedy, J. W. *J. Chem. Soc., Faraday Trans. 2* **1973**, *69*, 484-504.

(9) The ¹⁷O NMR data reported in ref 2 also indicate more negative charge on OMo and OMo₂ oxygens in anion **1** relative to the corresponding oxygens in $\text{VMo}_5\text{O}_{19}^{3-}$.

(10) Space group and unit cell data for the two compounds are as follows: tetragonal, space group $I4_1cd-C_{4v}^2$ (no. 110) with $a = b = 25.162$ (9), $c = 50.380$ (10) Å; $Z = 4$ for **1**; monoclinic, space group $P2_1/c-C_{2v}^2$ (no. 14) with $a = 22.871$ (9), $b = 11.942$ (5), $c = 25.127$ (8) Å; $\beta = 110.52$ (3)°; $Z = 4$ for **2**. Intensity measurements were made on a computer-controlled Nicolet P1 autodiffractometer for both compounds using full (0.90° wide) ω scans and graphite-monochromated Mo $K\alpha$ radiation. Totals of 7412 (**1**) and 11777 (**2**) independent reflections having $2\theta_{\text{MoK}\alpha} < 50.70^\circ$ (the equivalent of 0.80 limiting Cu $K\alpha$ spheres) were measured for both compounds. The titanium and molybdenum atoms of both totally general-position asymmetric units were located by using direct methods (MULTAN); the remaining anionic, cationic, and solvent nonhydrogen atoms were located by standard difference Fourier techniques. The resulting parameters have been refined to convergence [R_1 (unweighted, based on F) = 0.064 for 2556 reflections of **1** having $2\theta_{\text{MoK}\alpha} < 43^\circ$ and $I > 3\sigma(I)$; $R_1 = 0.057$ for 6882 reflections of **2** having $2\theta_{\text{MoK}\alpha} < 50.70^\circ$ and $I > 3\sigma(I)$]. Cycles of unit-weighted full-matrix least-squares refinement for both compounds utilized anisotropic thermal parameters for all anionic and solvent nonhydrogen atoms and isotropic thermal parameters for all cationic nonhydrogen atoms. Refinement is continuing for both compounds with the more complete data sets and/or models which incorporate anisotropic thermal parameters for all nonhydrogen atoms.

(11) (a) Allcock, H. R.; Bissel, E. C.; Shawl, E. T. *Inorg. Chem.* **1973**, *12*, 2963. (b) Garner, C. D.; Howlander, N. C.; Mabbs, F. E.; McPhail, A. T.; Miller, R. W.; Onan, K. D. *J. Chem. Soc. Dalton Trans.* **1978**, 1582.

(12) Iball, J.; Low, J. N.; Weakley, T. J. R. *J. Chem. Soc. Dalton Trans.* **1974**, 2021.

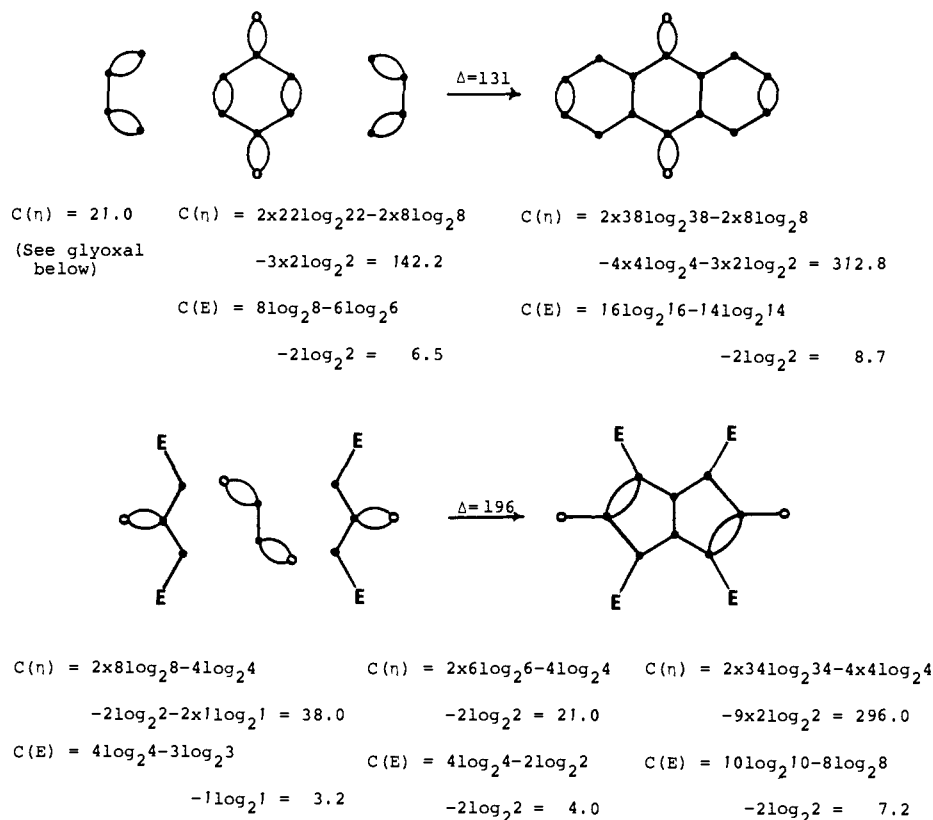


Figure 1. A comparison of a typical Diels–Alder reaction (butadiene + *p*-benzoquinone) with a typical Weiss reaction (dimethyl 3-ketogutarate + glyoxal). The stereochemistry of the products is implicit in the assignment of symmetry, i.e., equivalent connections. The Diels–Alder product is *cis-anti-cis* (C_{2h}); the Weiss product is *cis*, with *exo* ester groups (C_2). The complexity within the ester groups has been ignored, since they do not undergo change.

By noting that η is equal to the number of lines in the line graph of M' ,⁵ it is concluded that ethylene has one connection and acetylene, three. For substituted multiple bonds the connections to the substituents are counted in the usual way (i.e., the number of propanes⁶). In 1-butene the bond from the methyl group to the methylene makes one connection with the bond from the methylene to the double bond. This bond in turn makes one connection with each of the two bonds of the double bond, which contains one connection, for a total of four. The presence of branching, rings, and multiple bonds in a molecule are complicating factors, and η increases with all of them.

Treatises on synthesis invariably emphasize the importance of using the symmetry present in the target molecule (or intermediates on the way to it) as a means of simplification.^{1,4} Concepts from information theory⁷ can be used to elaborate η into an index that reflects symmetry (and therefore chirality and stereochemistry) by considering the distribution of connections among equivalence classes. Rashevsky⁸ first calculated the "information content" (per point) of a graph using Shannon's formula⁷

$$I = - \sum_i p_i \log_2 p_i \quad (1)$$

where $p_i = n_i/n$; n_i was taken to be the number of points in the *i*th set of equivalent points and n , the total number of points. Trucco⁹ noted that eq 1 can be applied to lines as well as points (indeed, it can be applied to any invariant). Mowshowitz¹⁰ defined the n_i as "the set of orbits of the automorphism group of a graph" and proposed eq 1 as a measure of "the relative complexity of graphs". This function is well suited to account for the symmetry of a graph, because it increases monotonically from zero, when all the invariants are equivalent, to its maximum value, when all

the invariants are nonequivalent. Multiplying eq 1 by n converts it from a "per invariant" to a "per graph" basis in eq 2, which

$$nI = n \log_2 n - \sum_i n_i \log_2 n_i \quad (2)$$

has been applied to the distances in a graph by Bonchev and Trinajstić.¹¹ While eq 2 accounts for the symmetry of a graph in terms of the invariant used, it does not properly reflect the size (number of invariants), since $I = 0$ when all the invariants are equivalent no matter how large the graph. Reasoning that complexity should increase with the number of invariants, when they are all equivalent, in the same way it does when they are all nonequivalent, $n \log_2 n$ is added to eq 2 to give a new definition of the complexity of a graph as a function of the invariant n :

$$C(n) = 2n \log_2 n - \sum_i n_i \log_2 n_i \quad (3)$$

Although n can represent any graph-theoretical invariant, the choice is critically important if chemically meaningful results are to be obtained. For example, if n is taken to symbolize points, cyclobutane and tetrahedrane are assigned equal complexities [$C(\text{points}) = 8.00$]; if n is taken to represent lines, cyclohexane and tetrahedrane have the same value [$C(\text{lines}) = 15.51$]. Since it takes into account branching as well as size and symmetry, $C(\text{connections})$ gives the order tetrahedrane (43.02) < cyclohexane (15.51) < cyclobutane (8.00), which makes sense considering that tetrahedrane contains three cyclobutane subgraphs in addition to four cyclopropane subgraphs. Henceforth $n = \eta$, giving the index $C(\eta)$.

The complexity of a molecule increases with the number (and the number of kinds) of heteroatoms, and the symmetric disposition of heteroatoms is a simplifying factor just as it is for carbon atoms. Therefore an equation analogous to (2) is appropriate for treating the complexity due to heteroatoms:

(7) Shannon, C. A.; Weaver, W. "The Mathematical Theory of Communication"; University of Illinois Press: Urbana, 1949; p 18–21.

(8) Rashevsky, N. *Bull. Math. Biophys.* **1955**, *17*, 229–235.

(9) Trucco, E. *Bull. Math. Biophys.* **1956**, *18*, 129–135.

(10) Mowshowitz, A. *Bull. Math. Biophys.* **1968**, *30*, 175–204.

(11) Bonchev, D.; Trinajstić, N. *J. Chem. Phys.* **1977**, *67*, 4517–4533.

$$C(E) = E \log_2 E - \sum_i E_i \log_2 E_i \quad (4)$$

where the E_i are the number of atoms of element i , and E is the total number of atoms. The total complexity is $C_T = C(\eta) + C(E)$. When all the atoms are the same, $C(E) = 0$ and the total complexity, C_T , is equal to $C(\eta)$, the complexity due to connectivity. Although a number of "branching" indexes and "topological" indexes have been advanced,^{8,11,12} C_T is the first measure of molecular complexity that is completely general.

Now that it is possible to calculate a number for any molecule which measures its complexity, it is possible to calculate the change in complexity, ΔC_T (hereafter symbolized as Δ), upon going from reactant to product in the course of a chemical reaction. The increases in complexity for the Diels-Alder reaction between butadiene and *p*-benzoquinone and the Weiss reaction¹³ of glyoxal with dimethyl 3-ketoglutarate are calculated in Figure 1. This example shows that it is still possible to invent powerful synthetic reactions¹⁴ and the calculation of complexity can aid in recognizing them. For a *functional group interchange*¹ such as the conversion of cyclohexene to cyclohexanol ($\Delta = 1.1$), the change in complexity is small. The process of calculating Δ can be repeated for all the steps in a synthetic sequence, thus providing a means to gauge progress toward a complex target molecule.

Thus by equating the complexity of a molecule with that of its molecular graph, formulating a new equation to measure this quantity, and applying this equation to the appropriate graph-theoretical invariant (connections), the first general index of molecular complexity has been constructed and shown to be applicable to synthetic analysis.

Acknowledgment. I thank D. Slepian and G. Smolinsky for their helpful advice.

(12) Hosoya, H. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332-2339. Merrifield, R. E.; Simmons, H. E. *Theor. Chim. Acta* **1980**, *55*, 55-75. Randić, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609-6615. Wiener, H. *Ibid.* **1947**, *69*, 17-20. Randić, M.; Brissey, G. M.; Spencer, R. B.; Wilkins, C. L. *Comput. Chem.* **1979**, *3*, 5-13.

(13) Weiss, U.; Edwards, J. M. *Tetrahedron Lett.* **1968**, 4885-4887. Mitschka, R.; Cook, J. M.; Weiss, U. *J. Am. Chem. Soc.* **1978**, *100*, 3973. Bertz, S. H.; Cookey, J. M.; Gawish, A.; Weiss, U. *Org. Synth.* in press.

(14) A Weiss reaction is the key step in a synthesis of modhephen under way in Professor J. M. Cook's laboratories. Another Weiss reaction has been used to synthesize isocoumarin; see: Dauben, W. G.; Walker, D. M. *J. Org. Chem.* **1981**, *46*, 1103-1108.

Selective Epoxidation of Olefins by Molybdenum Porphyrin Catalyzed Peroxy-Bond Heterolysis

Henry J. Ledon,* Patrick Durbut,† and François Varescon

Centre National de la Recherche Scientifique
Institut de Recherches sur la Catalyse
69626 Villeurbanne Cédex, France

Received July 8, 1980

Revised Manuscript Received April 9, 1981

The selective epoxidation of olefins with tertiary alkyl hydroperoxides, catalyzed by molybdenum complexes, is a synthetic reaction of great significance which is used to produce industrial organic chemicals.¹ However, the nature of the actual oxidizing species is not clearly established. In particular, the requirement for olefin activation through its coordination to the metal center prior to the oxygen-transfer step is still a matter of controversy.^{1,2}

* On leave from Université de Liège, Laboratoire de Chimie Macromoléculaire et de Catalyse Organique, Sart-Tilman B 4000, Liège, Belgique.

(1) For recent reviews, see: (a) Sheldon, R. A.; Kochi, J. K. *Adv. Catal.* **1976**, *25*, 272-413. (b) Lyons, J. E. *Aspects Homogeneous Catal.* **1977**, *3*, 1-136. (c) Parshall, G. W. *J. Mol. Catal.* **1978**, *4*, 243-270. (d) Sheldon, R. A. *Ibid.* **1980**, *7*, 107-126. (e) "Chemical and Physical Aspects of Catalytic Oxidation"; Portefaix, J. L.; Figueras, F., Eds.; Editions du C.N.R.S.: Lyon, 1980.

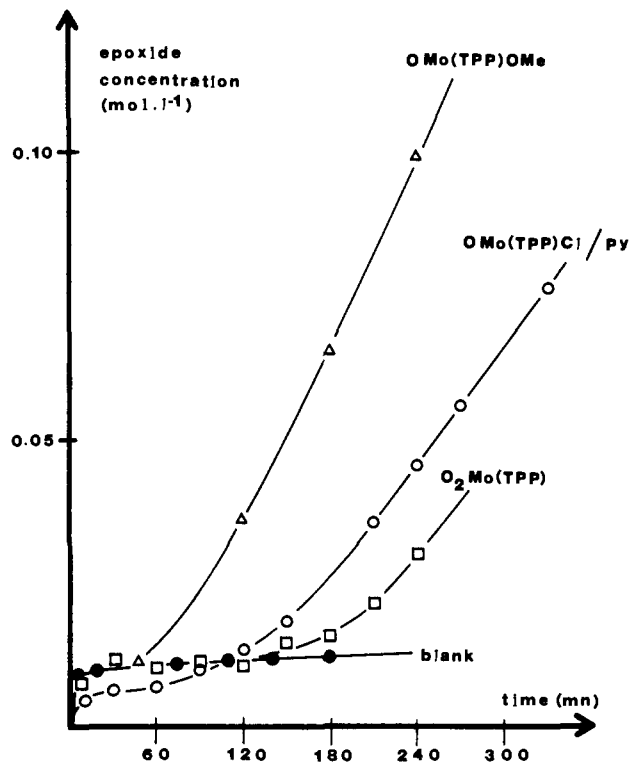


Figure 1. Cyclohexene epoxidation with *t*-BuOOH catalyzed by several molybdenum porphyrin complexes.

Our current interest in the reactivity of molybdenum porphyrins³⁻⁷ led us to study their catalytic properties. Owing to the steric hindrance of the macrocyclic ligand, the coordination, at the same time, of the olefin and the hydroperoxide to the metal center would be unlikely. Here we report on the selective epoxidation of olefins with *tert*-butyl hydroperoxide catalyzed by (5,10,15,20-tetra-phenylporphyrinato)molybdenum(V) complexes [OMo(TPP)X].^{8,9}

In a typical experiment, OMo(TPP)OMe⁶ (ca. 4 mg; 5×10^{-6} mol) was dissolved in dry oxygen-free benzene (13 mL). Then cyclohexene (2.5 mL; 24.6 mmol) and *n*-decane (used a GC internal standard) were added. The accurate catalyst concentration was measured by UV-visible spectrophotometry, and the mixture was heated to 80 ± 2 °C under argon. After equilibration, *t*-BuOOH (0.5 mL; 5 mmol) was added, and aliquots were periodically taken for GC analysis and recording of the UV-visible spectra.¹⁰

A high selectivity to cyclohexene oxide was obtained at total conversion of the hydroperoxide (Table I). Figure 1 shows the rates of appearance of cyclohexene oxide according to the nature

(2) (a) Sheldon, R. A.; Van Doorn, J. A. *J. Catal.* **1973**, *31*, 427-437. (b) Chong, A. O.; Sharpless, K. B. *J. Org. Chem.* **1977**, *42*, 1587-1590. (c) Mimoun, H. *J. Mol. Catal.* **1980**, *7*, 1-29. (d) Mimoun, H.; Charpentier, R.; Mitschler, A.; Fischer, J.; Weiss, R. *J. Am. Chem. Soc.* **1980**, *102*, 1047-1054. (3) Ledon, H.; Bonnet, M.; Lallemand, J.-Y. *J. Chem. Soc., Chem. Commun.* **1979**, 702-704.

(4) Ledon, H.; Bonnet, M. *J. Mol. Catal.* **1980**, *7*, 309-313. (5) Mentzen, B. F.; Bonnet, M. C.; Ledon, H. *J. Inorg. Chem.* **1980**, *19*, 2066-2069.

(6) Ledon, H. *C. R. Hebd. Seances Acad. Sci., Ser. C* **1978**, *287*, 59-61. Ledon, H. J.; Bonnet, M. C.; Brigandat, Y.; Varescon, F. *Inorg. Chem.* **1980**, *19*, 3488-3491.

(7) Ledon, H.; Mentzen, B. *Inorg. Chim. Acta* **1978**, *31*, L393-L394.

(8) Oxygenation of cyclohexene in a complex mixture containing both iron and molybdenum (or vanadium) octaethylporphyrin was claimed to produce cyclohexene oxide and 2-cyclohexen-1-ol in a 1:1 ratio, but very few experimental details were given. See: Baccouche, M.; Ernst, J.; Fuhrhop, J.-H.; Arzoumanian, H. *J. Chem. Soc., Chem. Commun.* **1977**, 821-822.

(9) Part of this work has been presented at the 2nd International Symposium on Homogeneous Catalysis, Sept 1-3, 1980, Düsseldorf, Federal Republic of Germany: Ledon, H.; Durbut, P.; Bonnet, M. "Abstracts of Papers", p 186.

(10) Concentrations of OMo(TPP)X complexes were monitored by measuring the absorbance at 454 nm after conversion into OMo(TPP)OMe by addition of methanol.