of the ground state itself is unfavorably high (Figure 3). The charged side chains in the active site lower the energy of the ground state to a much more favorable value, while not completely abolishing the transition-state stabilization gained in the lowpolarization environment of the active site. This phenomenon is not unique to IDH and related decarboxylating dehydrogenases. For example, an antibody which catalyzes the decarboxylation of 3-carboxybenzisoxazoles can be induced by a hapten designed to elicit positively charged carboxylate-binding residues and an apolar environment to stabilize the transition state.<sup>24</sup> When the

(24) Lewis, C.; Kramer, T.; Robinson, S.; Hilvert, D. Science 1991, 253, 1019.

requirements for binding energy and transition-state stabilization are considered together, it becomes clear that both the low-polarizability environment and strong electrostatic field in the active site are critical for enzyme function.

Acknowledgment. J.H.H. thanks Dr. Robert Stroud for encouragement during the early stages of this work. This work was supported by an American Cancer Society postdoctoral fellowship to J.H.H., Grant DMB 88-17438 (National Science Foundation) to S.J.R., and a grant from the Lucille P. Markey Charitable Trust.

Registry No. IDH, 9028-48-2; Mg<sup>2+</sup>, 7439-95-4; decarboxylase, 9027-22-9.

## Molecular Dissimilarity: A Momentum-Space Criterion

David L. Cooper\* and Neil L. Allan<sup>†</sup>

Contribution from the Department of Chemistry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, U.K., and School of Chemistry, University of Bristol, Cantocks Close, Bristol BS8 1TS, U.K. Received September 25, 1991

Abstract: Quantitative measures of the similarity of momentum-space electron densities are investigated using ab initio and semiempirical wave functions. The model series CH<sub>3</sub>OCH<sub>3</sub>, CH<sub>3</sub>SCH<sub>3</sub>, and CH<sub>3</sub>CH<sub>2</sub>CH<sub>3</sub> is examined. A family of molecular dissimilarity indices  $D_{AB}(n)$  appears to be particularly useful. The electron densities of individual C-F bonds and C-H bonds in CH<sub>4</sub>, CH<sub>3</sub>F, CH<sub>2</sub>F<sub>2</sub>, CHF<sub>3</sub>, and CF<sub>4</sub> are compared. Values of  $D_{AB}(-1)$  relative to CH<sub>4</sub> are consistent with the relative reactivities of CH<sub>3</sub>F, CH<sub>2</sub>F<sub>2</sub>, and CHF<sub>3</sub> with the hydroxyl radical.

#### Introduction

The general concept of molecular similarity is becoming increasingly important<sup>1</sup> in many areas of chemistry; it can be especially useful when the molecular processes are very complex, as in pharmacology. A number of authors have proposed quantitative definitions of molecular similarity, with particular emphasis on comparing electrostatic potentials or position-space electron densities,<sup>2-4</sup> including topological analysis of the three-dimensional shapes of charge densities.<sup>5,6</sup> In a previous paper,<sup>7</sup> we suggested a novel alternative based on momentum-space electron distributions, which concentrates on the outer valence-electron density and which is not subject to many of the drawbacks of the more conventional approaches. In this paper, we introduce more discriminating quantities, which we term "momentum-space dissimilarity indices", and we investigate the use of these definitions with ab initio and semiempirical wave functions.

#### Momentum-Space Molecular Similarity

The most straightforward means of obtaining momentum-space wave functions is to take the Fourier transform of the more familiar position-space wave function. The relationship in p-space between the wave function  $\Psi(\mathbf{p})$  and the total electron density  $\rho(\mathbf{p})$  is exactly the same as in **r**-space. In the case of an SCF wave function, for example, the contribution to  $\rho(\mathbf{p})$  of an electron in molecular orbital  $\psi(\mathbf{p})$  is simply  $|\psi(\mathbf{p})|^2$ . The momentum-space electron density falls off very rapidly with increasing  $p = |\mathbf{p}|$ , and thus it is dominated by low values of p, which correspond in position space to the slowly-varying outer valence-electron density.

In our previous work,<sup>7</sup> we defined a family of momentum-space similarity indices  $S_{AB}(n)$  according to

$$S_{AB}(n) = \frac{100I_{AB}(n)}{\frac{1}{2}(I_{AA}(n) + I_{BB}(n))}$$
(1)

\* University of Liverpool.

```
<sup>†</sup>University of Bristol.
```

in which

$$I_{PQ}(n) = 2 \int p^n \rho_P(\mathbf{p}) \ \rho_Q(\mathbf{p}) \ d\mathbf{p} \qquad (PQ = AA, BB, or AB)$$
(2)

and A and B label the two molecules. The different powers of p in the integrands emphasize different regions of the momentum-space valence-electron density; typical values of n are -1, 0, 1, and 2. The values of  $S_{AB}(n)$ , which lie in the range 0-100, are insensitive to the core-electron densities. Small displacements (in r-space) of heavy atoms have negligible effects on the indices. In addition, the values of  $S_{AB}(n)$  are dependent on the relative orientation of the two molecules in r-space, but not on the distances between them.

The evaluation of  $S_{AB}(n)$  using ab initio momentum-space electron densities is relatively straightforward, and it is not computationally expensive. In our previous work,7 we considered the model systems H<sub>2</sub>O, H<sub>2</sub>S, CH<sub>3</sub>OCH<sub>3</sub>, CH<sub>3</sub>SCH<sub>3</sub>, and CH<sub>3</sub>C- $H_2CH_3$ . It is mainly the cost of carrying out ab initio SCF calculations which limits the use of our procedure for the routine study of large series of large molecules. Consequently, it seems very worthwhile to investigate the reliability of values of  $S_{AB}(n)$ computed using much more readily obtainable semiempirical wave functions. Accordingly, we have reexamined the model systems

(6) Duane-Walker, P.; Artera, G. A.; Mezey, P. G. J. Comput. Chem. 1991, 12, 220

(7) Cooper, D. L.; Allan, N. L. J. Comput.-Aided Mol. Des. 1989, 3, 253-259.

<sup>(1)</sup> Concepts and Applications of Molecular Similarity; Johnson, M. A., Maggiora, G. M., Eds.; Wiley: New York, 1990.

<sup>(2)</sup> Carbó, R.; Leyda, L.; Arnau, M. Int. J. Quantum. Chem. 1980, 17,

<sup>(</sup>a) Hodgkin, E. E.; Richards, W. G. Int. J. Quantum Chem. 1987, 32, 517.
(3) Hodgkin, E. E.; Richards, W. G. Int. J. Quantum Chem., Quantum Biol. Symp. 1987, 14, 105. Richards, W. G.; Hodgkin, E. E. Chem. Br. 1988, 24, 1141 and references therein. Burt, C.; Richards, W. G. J. Comput.-Aided

Mol. Des. 1990, 4, 231.

 <sup>(4)</sup> Ponec, R. Collect. Czech. Chem. Commun. 1987, 52, 555. Ponec, R.
 Z. Phys. Chem. (Leipzig) 1987, 268, 1180.
 (5) Mezey, P. G. J. Math. Chem. 1988, 2, 299.

Table I. Similarity Indices  $S_{AB}(n)$  for the Model Series A = CH<sub>3</sub>OCH<sub>3</sub>, B = CH<sub>3</sub>SCH<sub>3</sub>, C = CH<sub>3</sub>CH<sub>2</sub>CH<sub>3</sub><sup>a</sup>

		1	4			J	3		С			
	-1	0	1	2	-1	0	1	2	-1	0	1	2
					MN	DO Wave F	unctions					
Α	100	100	100	100	97.3	97.0	96.8	96.5	97.6	97.6	97.6	97.4
В	97.3	97.0	96.8	96.5	100	100	100	100	99.9	99.8	99.9	99.7
С	97.6	97.6	97.6	97.4	99.9	99.8	99.9	99.7	100	100	100	100
					Ab Ir	itio Wave	Functions					
Α	100	100	100	100	97.4	97.5	97.5	97.3	98.4	98.5	98.6	98.8
В	97.4	97.5	97.5	97.3	100	100	100	100	99.8	99.8	99.6	99.8
C	98.4	98.5	98.6	98.8	99.8	99.8	99.6	99.8	100	100	100	100

<sup>a</sup> The values calculated with ab initio wave functions are reproduced from ref 7. The columns are headed with the values of n.

**Table II.** Dissimilarity Indices  $D_{AB}(n)$  for the Model Series A = CH<sub>3</sub>OCH<sub>3</sub>, B = CH<sub>3</sub>SCH<sub>3</sub>, C = CH<sub>3</sub>CH<sub>2</sub>CH<sub>3</sub> Calculated Using MNDO Wave Functions<sup>a</sup>

			4			]	В		С				
	-1	0	1	2	-1	0	1	2	-1	0	1	2	
A	0	0	0	0	161	120	109	118	137	93	81	85	
В	161	120	109	118	0	0	0	0	5	4	7	12	
С	137	93	81	85	5	4	7	12	0	0	0	0	

<sup>a</sup> The columns are headed with the values of n.

 $CH_3OCH_3$ ,  $CH_3SCH_3$ , and  $CH_3CH_2CH_3$  using wave functions taken directly from the well-known MOPAC program.<sup>8</sup>

MNDO wave functions were calculated for each of the three systems in the model series, with the  $C_2$ -axis taken to be the z-direction, and with the y-axis in the  $\sigma_v'$  plane. The corresponding momentum-space similarity indices  $S_{AB}(n)$  are reported in Table I together with the values calculated in our previous work using "all-electron" ab initio SCF wave functions (4-31G basis). It is clear that the two sets of numbers exhibit much the same pattern, with the greatest similarity indices between CH<sub>3</sub>SCH<sub>3</sub> and CH<sub>3</sub>CH<sub>2</sub>CH<sub>3</sub>: the replacement of CH<sub>2</sub> by S rather than O is a common device in the pharmaceutical industry.

Of course, the two sets of numbers in Table I are not identical. This is a reflection of the different qualities of semiempirical and ab initio wave functions. In addition, the MNDO-based results utilize "valence-electron" densities rather than "all-electron" densities, but this should be a very small effect, as intimated earlier.

#### Momentum-Space Molecular Dissimilarity

The range of values in Table I is very small, and it is evident that it would be useful at this stage to define a more discriminating quantity. Accordingly, we introduce a family of momentum-space molecular dissimilarity indices  $D_{AB}(n)$  defined by

$$D_{AB}(n) = 100[\frac{1}{2}(I_{AA}(n) + I_{BB}(n)) - I_{AB}(n)]$$
(3)

These quantities are no longer normalized to lie in the range 0-100.  $D_{AB}(n)$  values of 0 indicate two identical momentum-space electron densities. More positive values of  $D_{AB}(n)$  correspond to greater dissimilarity between the two molecules.

Dissimilarity indices for our model series, calculated using MNDO wave functions, are recorded in Table II. Inspection of the values of  $D_{AB}(n)$  leads to the same conclusions as before, but these new numbers are somewhat easier to interpret because of the much larger variation.

#### **Orbital Dissimilarities**

It is easy to envisage a variety of circumstances in which it might be more appropriate to compare individual orbitals rather than total (valence) electron densities for two molecules. For example, it would now be entirely straightforward to calculate dissimilarity indices between HOMOs and/or LUMOs of different molecules. In view of the well-known, mostly qualitative successes of frontier orbital theory,<sup>9</sup> this could be a particularly fruitful avenue for future investigations. In the present paper, we concentrate instead on a somewhat different situation. It is often the case that one would like to compare parts of biologically active systems, such as side chains or particular functional groups. In this instance it is most useful to transform to  $\mathbf{p}$ -space only those parts of the  $\mathbf{r}$ -space wave function which describe the "interesting" regions of the molecule. This presupposes, of course, that a simple, reliable and computationally cheap scheme is available to localize the occupied ab initio or semiempirical molecular orbitals. We now describe such a scheme.

In general, localization schemes rely on the fact that SCF wave functions are invariant to arbitrary unitary transformations of the doubly-occupied MOs among themselves. A wide variety of localization schemes, of differing expense, efficiency, and objectivity, have been used to obtain localized molecular orbitals (LMOs) by means of orbital rotations. In the course of the present work, we have implemented for MOPAC wave functions the population localization criterion described by Pipek and Mezey.<sup>10</sup> This general procedure is also extensively used as a precursor to ab initio spin-coupled valence bond calculations<sup>11</sup> and has proved itself to be very reliable in a wide variety of applications.<sup>12</sup>

The basic idea of the population localization scheme is to maximize the quantity Z defined by

$$Z = \sum_{A} \sum_{i} [P_A(i)]^2$$
(4)

in which  $P_A(i)$  is the contribution made by each electron in MO  $\psi_i$  to the Mulliken population on center A. This was achieved in the present work by means of successive  $2 \times 2$  orbital rotations using a strategy analogous to that described in ref 11 for ab initio wave functions. In common with our similarity/dissimilarity and plotting programs, our population localization program for MO-PAC wave functions has been written so as to operate directly on the standard "graphical" files produced by the MOPAC program using the GRAPH directive.

We decided to employ a very stringent test of LMO dissimilarity indices. As we have discussed previously,<sup>13</sup> C-F bonds at primary sites in fluoroalkanes differ remarkably little from one system to another. A particularly severe task, therefore, would be to try to use values of  $D_{AB}(n)$  to discriminate between the C-F bonds

<sup>(8)</sup> Stewart, J. J. P. J. Comput.-Aided Mol. Des. 1990, 4, 1 and references therein.

<sup>(9)</sup> See, for example: Fleming, I. Frontier orbitals and organic chemical reactions; Wiley: Chichester, 1978.

<sup>(10)</sup> Pipek, J.; Mezey, P. G. J. Chem. Phys. 1989, 90, 4916.

<sup>(11)</sup> Cooper, D. L.; Gerratt, J.; Raimondi, M. J. Mol. Struct. (THEO-CHEM) 1991, 229, 155-162.

<sup>(12)</sup> Cooper, D. L.; Gerratt, J.; Raimondi, M. Chem. Rev. 1991, 91, 929-964. Sironi, M.; Raimondi, M.; Cooper, D. L.; Gerratt, J. J. Phys. Chem. 1991, 95, 10617-10623.

 <sup>(13)</sup> Cooper, D. L.; Allan, N. L.; Powell, R. L. J. Fluorine Chem. 1990, 46, 317–337. Cooper, D. L.; Wright, S. C.; Allan, N. L.; Winterton, N. J. Fluorine Chem. 1990, 47, 489–507. Cooper, D. L.; Allan, N. L.; Powell, R. L. J. Fluorine Chem. 1990, 49, 421–432.

Table III. Dissimilarity Indices  $D_{AB}(n)$  for C-F Bonds in Fluoromethanes, Calculated Using AM1 Wave Functions<sup>a</sup>

	·	CI	H₃F		CH <sub>2</sub> F <sub>2</sub>				CHF3				CF <sub>4</sub>				
	-1	0	1	2	-1	0	1	2	-1	0	1	2	-1	0	1	2	
CH₃F	0	0	0	0	0.2	0.1	0.2	0.2	0.4	0.3	0.3	0.4	1.5	0.8	0.7	0.7	
CH <sub>2</sub> F <sub>2</sub> CHF <sub>3</sub>	0.2 0.4	0.1 0.3	0.2 0.3	0.2 0.4	0 0.2	0 0.2	0 0.2	0 0.1	0.2 0	0.2 0	0.2 0	0.1 0	1.2 0.8	0.6 0.5	0.5 0.4	0.5 0.4	
CF₄	1.5	0.8	0.7	0.7	1.2	0.6	0.5	0.5	0.8	0.5	0.4	0.4	0	0	0	0	

<sup>a</sup> The columns are headed with the values of n.

Table IV. Dissimilarity Indices  $D_{AB}(n)$  for C-H Bonds in Fluoromethanes, Calculated Using AM1 Wave Functions<sup>a</sup>

		CH <sub>3</sub> F			CH <sub>2</sub> F <sub>2</sub>				CHF3							
	-1	0	1	2	-1	0	1	2	-1	0	1	2	-1	0	1	2
CH₄	0	0	0	0	1.1	0.7	0.5	0.4	3.9	1.8	1.1	0.9	13.1	4.5	2.3	1.6
CH	1.1	0.7	0.5	0.4	0	0	0	0	1.9	0.9	0.7	0.5	9.2	3.3	1.8	1.3
CH <sub>2</sub> F,	3.9	1.8	1.1	0.9	1.9	0.9	0.7	0.5	0	0	0	0	4.5	1.7	1.1	0.8
CHF <sub>3</sub>	13.1	4.5	2.3	1.6	9.2	3.3	1.8	1.3	4.5	1.7	1.1	0.8	0	0	0	0

<sup>a</sup> The columns are headed with the values of n.

in the series of fluoromethanes  $CH_3F$ ,  $CH_2F_2$ ,  $CHF_3$ , and  $CF_4$ . MNDO tends to produce large errors for highly fluorinated compounds, probably because of overestimating the FF repulsions,<sup>14</sup> and so we have used AM1 instead (see also ref 13). All the molecules were oriented with one C-F bond pointing along the z-axis.

The AM1 method suggests very similar Mulliken charges on fluorine for each of  $CH_3F$ ,  $CH_2F_2$ , and  $CHF_3$ . The F atoms in  $CF_4$  are predicted to be slightly less negatively charged. As is to be expected, the AM1 Mulliken charges on carbon increase markedly with successive fluorination. The C-F bond lengths in all four molecules are very similar: AM1 predicts a very small decrease (by 0.02 Å overall) on successive fluorination from  $CH_3F$ to  $CF_4$ .

The population localization scheme was used to obtain LMOs for the series of fluoromethanes, and orbital dissimilarities were calculated for the C-F bonds pointing along the z-axis. The values of  $D_{AB}(n)$  generated in this way are listed in Table III. All of the values are very small, as we would expect for such remarkably similar C-F bonds. Nonetheless, it is clear that the C-F bonds in "adjacent" molecules in the series CH<sub>3</sub>F, CH<sub>2</sub>F<sub>2</sub>, and CHF<sub>3</sub> are the least dissimilar. The values of  $D_{AB}(n)$  differ very little with changes in n = -1, 0, 1, and 2, and so the electron densitiesmust be very similar to one another over extended regions of momentum space. Values of  $D_{AB}(n)$  for the  $CF_4/CH_xF_{4-x}$  comparisons are slightly larger, although still very small, and increase with increasing x, as one might anticipate. For comparisons involving CF<sub>4</sub>, the largest values of  $D_{AB}(n)$  occur for n = -1, so that the densities differ most at low momentum. Evidently,  $D_{AB}(n)$ is sufficiently discriminating to portray the very small differences between C-F bonds in fluoromethanes.

AM1 wave functions for the series  $CH_4$ ,  $CH_3F$ ,  $CH_2F_2$ , and  $CHF_3$  were calculated, with each molecule oriented with one C-H bond pointing along the z-axis, and LMOs were generated using the population localization procedure. According to AM1, and indeed experimental measurements, there is scarcely any variation in the C-H bond length with increasing fluorination. Nonetheless, there are significant differences in the chemical reactivity of these C-H bonds as manifested, for example, by the rate constants and activation energies for hydrogen-atom abstraction by various free radicals. Of particular importance is the reaction with hydroxyl radicals: this is the dominant loss process in the troposphere for hydrofluorocarbons (CFCs). In this context, various models predict very much longer residence times for CHF<sub>3</sub> than  $CH_2F_2$  or  $CH_3F.^{15,16}$  With this in mind, we would expect the LMO dis-

**Table V.** Values of  $D_{AB}(-1)$  for C-H Bonds (Relative to CH<sub>4</sub>) and Kinetic Data for Reaction of Fluoromethanes with Hydroxyl Radicals<sup>*a*</sup>

	$D_{AB}(-1)$	$E_{\rm A}/R$ (K)	$k (10^{-14} \text{ cm}^3 \text{ s}^{-1})$
CH₄	0	$1700 \pm 200$	0.77
CH <sub>1</sub> F	1.1	$1700 \pm 300$	1.8
CH <sub>2</sub> F <sub>2</sub>	3.9	$1650 \pm 200$	1.0
CHF3	13.1	$2650 \pm 500$	0.021

<sup>a</sup>Taken from ref 17.  $E_A$  is the activation energy and k is the rate constant at 298 K.

similarity indices to indicate that the C-H bonds in this series of molecules are all very similar to one another, but we would expect the values of  $D_{AB}(n)$  to show larger values and a greater variation with *n* than was the case for the C-F bonds. This assumes, of course, that there is some correlation between properties of the isolated molecules and properties of the corresponding transition states.

Dissimilarity indices for the C-H bonds in fluoromethanes, which are collected in Table IV, show the expected pattern. The p-space electron densities corresponding to the various C-H bonds in the fluoromethanes appear to differ most at low momentum. The largest values of  $D_{AB}(n)$  occur for the comparisons involving  $CHF_3$ . It is, perhaps, useful to put the magnitudes of these numbers in context:  $D_{AB}(-1) = 5$  for the comparison of CH<sub>3</sub>SCH<sub>3</sub> with  $CH_3CH_2CH_3$  (see Table II) whereas  $D_{AB}(-1) = 13$  for the C-H bonds in CHF<sub>3</sub> and CH<sub>4</sub>. In Table V, we list values of  $D_{AB}(-1)$  relative to CH<sub>4</sub>, and the activation energies and rate constants for hydrogen-atom abstraction by OH.<sup>17</sup> It is clear that the dissimilarity indices can be used to put the rate constants for CH<sub>3</sub>F, CH<sub>2</sub>F<sub>2</sub>, and CHF<sub>3</sub> in order, although no linear correlations are apparent. The LMO dissimilarities certainly highlight the much lower reactivity of CHF<sub>3</sub>. It would be very useful to extend this work to other potential CFC replacements.

#### Conclusions

We have extended our previous study of the similarity of ab initio momentum-space electron densities by considering wave functions obtained using much cheaper semiempirical methods. In addition, we have introduced a new and more discriminating index, namely, the molecular dissimilarity  $D_{AB}(n)$ . The suitability of this approach was demonstrated using the model series CH<sub>3</sub>-OCH<sub>3</sub>, CH<sub>3</sub>SCH<sub>3</sub>, and CH<sub>3</sub>CH<sub>2</sub>CH<sub>3</sub>. The same general pattern emerges for similarity indices computed from either ab initio SCF (4-31G) or semiempirical (MNDO) wave functions.

We have outlined how a suitable localization scheme makes it possible to compare the electron densities of individual bonds. This approach was tested by considering the C-F bonds and the C-H bonds in the series of molecules  $CH_4$ ,  $CH_3F$ ,  $CH_2F_2$ ,  $CHF_3$ ,

 <sup>(14)</sup> Dewar, M. J. S.; Rzepa, H. S. J. Am. Chem. Soc. 1978, 100, 58.
 (15) Derwent, R. G.; Volz-Thomas, A.; Prather, M. J. UNEP/WMO Scientific Assessment of Stratospheric Ozone; Appendix, AFEAS Report; 1989; Chapter V.

<sup>(16)</sup> Cooper, D. L.; Allan, N. L.; McCulloch, A. Atmos. Environ. 1990, 24A, 2417-2419, 2703. Cooper, D. L.; Cunningham, T. P.; Allan, N. L.; McCulloch, A. Atmos. Environ. 1992, 26A, 1331-1334.

<sup>(17)</sup> Hampson, R. F.; Kurylo, M. J.; Sander, S. P. UNEP/WMO Scientific Assessment of Stratospheric Ozone; Appendix, AFEAS Report; 1989; Chapter III.

and CF<sub>4</sub>. The values of  $D_{AB}(-1)$  for the C-H bonds highlight the much lower reactivity of CHF<sub>3</sub> toward the hydroxyl radical.

In the cases considered here, the need to rotate one molecule relative to another did not arise. However, it should be stressed that this is very straightforward, and the ability to do this has already been implemented in our programs. We have not, as yet, considered distortion of the molecule(s) by means of, e.g., rotations about individual bonds.

The ability to use semiempirical wave functions and also to concentrate on particular regions of a molecular system means that it is now possible to compute values of  $D_{AB}(n)$  for large series

of large molecules. It is also possible to compare different regions of the same molecule. Comparisons of HOMOs and/or LUMOs could also be worthwhile. No single measure of similarity can be expected to be a universal panacea, but we believe that the momentum-space indices described here are likely to prove useful in structure-activity relations, particularly those involving binding energies to biologically active systems.

Registry No. CH<sub>3</sub>OCH<sub>3</sub>, 115-10-6; CH<sub>3</sub>SCH<sub>3</sub>, 75-18-3; CH<sub>3</sub>CH<sub>2</sub>CH<sub>3</sub>, 74-98-6; CH<sub>3</sub>F, 593-53-3; CH<sub>2</sub>F<sub>2</sub>, 75-10-5; CHF<sub>3</sub>, 75-46-7; CF<sub>4</sub>, 75-73-0; CH4. 74-82-8.

# Stereochemically Active or Inactive Lone Pair Electrons in Some Six-Coordinate, Group 15 Halides

### Ralph A. Wheeler\* and P. N. V. Pavan Kumar<sup>†</sup>

Contribution from the Department of Chemistry and Biochemistry, 620 Parrington Oval, Room 208, University of Oklahoma, Norman, Oklahoma 73019. Received October 21, 1991

Abstract: The valence shell electron pair repulsion (VSEPR) model has enjoyed remarkable success explaining the shapes of molecules consisting of main group atoms. Exceptional compounds are known, however, which display the "inert pair effect": they contain a lone pair of electrons that appear spherically symmetrical and are thus stereochemically inactive. We consider  $Bil_{6}^{3-}$ ,  $BiCl_{6}^{3-}$ ,  $Bil_{3}$ , and  $Sbl_{3}$  as prototypical molecules and solids containing six-coordinate, group 15 atoms with stereochemically active or inactive lone pair electrons on the central atom. Extended Huckel molecular orbital and crystal orbital calculations show an underlying tendency for a second-order Jahn-Teller distortion from octahedral to trigonal coordination of the central atom. The distortion mixes the central atom p<sub>z</sub> character of the unoccupied orbitals with the s character of the HOMO and thus hybridizes the lone pair electrons toward the longer M-X bonds (M = Bi or Sb; X = I or Cl). For the MX<sub>3</sub> compounds, the hybrid lone pair points toward the smaller X-M-X angles, and the basic assumption of the VSEPR model is confounded for extended solids. Whether or not the distortion is structurally realized for the  $BI_6^{3-}$  and  $BiCl_6^{3-}$  molecules depends upon the balance between HOMO-LUMO mixing and (1) the mixing between the HOMO and a lower energy 1a, orbital with ligand lone pair and Bi p, character and (2) repulsions among ligand lone pairs. The HOMO-1a, mixing is strongest and ligand lone pair-lone pair repulsions are greatest for large, electropositive ligands. HOMO-LUMO mixing predominates for the compounds that display trigonal distortions ( $BiCl_6^{3-}$  and  $SbI_3$ ), but not for those that remain octahedral ( $BiI_6^{3-}$  and Bil<sub>1</sub>).

#### Introduction

Compounds containing six-coordinate main group elements are remarkable for the variety of coordination geometries observed. Several striking examples are the isoelectronic molecules Bil63  $(1)^{2h}$  and BiCl<sub>6</sub><sup>3-</sup>  $(2)^{2a-g}$  as well as the extended solids BiI<sub>3</sub> (3)and SbI<sub>3</sub>.<sup>3a</sup> Each pair of compounds (BiI<sub>6</sub><sup>3-</sup> and BiCl<sub>6</sub><sup>3-</sup>; BiI<sub>3</sub>



<sup>†</sup>Current address: Department of Chemistry, University of Texas at Arlington, Arlington, TX 76019-0065.

and SbI<sub>3</sub>) is extraordinary because both structures of the pair cannot be explained by simple models such as valence shell electron pair repulsion (VSEPR).<sup>4</sup> In particular, lone pair electrons on the central Bi atoms of  $BiI_6^{3-}$  and  $BiI_3$  are not evident in the structures. They are stereochemically inactive and give rise to what has been called the "inert pair" effect.<sup>4</sup> We will develop these ideas more fully later, after discussing the structures shown in 1-3 and after introducing a number of compounds with similar structures. Then we will present a simple, unified model based

(3) BiI<sub>3</sub>: (a) Trotter, J.; Zobel, T. Z. Kristallogr. 1966, 123, 67-72. (b) Davidovich, R. L.; Buslaev, Y. A. Koord. Khim. 1988, 14, 1011-1036. (c) Pearson, W. B. The Crystal Chemistry and Physics of Metals and Alloys; Wiley: New York, 1972; p 445. (d) Greenwood, N. N.; Earnshaw, A. Chemistry of Elements; Pergamon Press: Oxford, 1984; p 652.

 (4) VSEPR: (a) Gillespie, R. J.; Nyholm, R. S. Quart. Rev. Chem. Soc.
 1957, 11, 339. (b) Searcy, A. W. J. Chem. Phys. 1958, 28, 1237–1242. (c) Gillespie, R. J. J. Am. Chem. Soc. 1960, 82, 5978–5983. (d) Gillespie, R. J. J. Chem. Educ. 1963, 40, 295–301; Ibid. 1970, 47, 18–23. (e) Gillespie, Computer Science (Science) R. J. Molecular Geometry; Van Nostrand and Reinhold: London, 1972. (f) Huheey, J. E. Inorganic Chemistry, 3rd ed.; Harper and Row: New York, 1983; p 208. (g) Jolly, W. L. Modern Inorganic Chemistry, 2nd ed.; McGraw-Hill: New York, 1991.

0002-7863/92/1514-4776\$03.00/0 © 1992 American Chemical Society

<sup>(1)</sup> General reference for  $MX_6^{n-}$  and  $MX_n$  (M = group 13-18 central atom, X = halogens): Wells, A. F. Structural Inorganic Chemistry, 5th ed.;

atom, X = halogens): Wells, A. F. Structural Inorganic Chemistry, 5th ed.; Clarendon Press: Oxford, 1984. (2) BiCl<sub>6</sub><sup>3-</sup>: (a) Morss, L. R.; Robinson, W. R. Acta Crystallogr. 1972, 28 B, 653-654. (b) Lazarini, F. Acta Crystallogr. 1987, 43C, 637-638. (c) Battaglia, L. P.; Corradi, A. B.; Nardelli, M.; Tani, M. E. V. J. Chem. Soc., Dalton Trans. 1978, 583-587. (d) Atoji, M.; Watanabe, T. J. Chem. Phys. 1952, 20, 1045-1046. (e) du Bois, A.; Abriel, W. Z. Naturforsch. 1988, 43B, 1003-1009. (f) du Bois, A.; Abriel, W. Z. Kristallogr. 1988, 182, 36-38. (g) Benachenhou, F.; Mairesse, G.; Nowogrocki, G.; Thomas, D. J. Solid State Chem. 1986, 65, 13-26. (h) Bil<sub>6</sub><sup>3-</sup>: Lazarini, F. Acta Crystallogr. 1977, 33B, 1057-1059. 1957-1959