Simulated *ab Initio* Molecular Orbital Technique. IV. Cyclohexanes[1a]

James E. Eilers,*[1b] Brian O'Leary, Brian J. Duke,[1c] Arno Liberles,[1d] and Donald R. Whitman[1e]

Contribution from the Theoretical Chemistry Department, University of Oxford, Oxford, England.

Received April 15, 1974

Abstract: The recently reported simulated *ab initio* molecular orbital (SAMO) technique has been used in a study of cyclohexane in both its chair and boat conformations. The adequacy of several different "pattern molecule" systems is examined, and two modifications to the basic SAMO method are introduced. One accounts for steric effects, and the other demonstrates the additive effect of multiple bonding pathways. *Ab initio* calculations on both conformers of cyclohexane are also reported. The SAMO and *ab initio* studies were carried out using the same basis set and geometry so that a direct comparison of results could be made. The total energy, the orbital energies, and the distribution of electrons in the basis orbitals are in excellent agreement. However, the SAMO method proves unsuitable for the energy difference between conformers.

The recently reported simulated *ab initio* molecular orbital (SAMO) technique has proved to be an effective but inexpensive means of producing wave functions of near *ab initio* accuracy for large molecules.[2] This method has been successfully applied to chain hydrocarbons,[2] aromatic rings,[3] simple polymers,[4] certain organic radicals,[5] and one large system of biological interest.[6]

The SAMO calculations on benzene and naphthalene[3] showed extremely good, but not perfect, agreement with complete *ab initio* calculations that employed a wide variety of differing basis sets. This imperfect agreement may simply reflect the use of different basis sets and/or slightly different geometries. However, no detailed comparison of a SAMO and an *ab initio* calculation using identical geometries and identical basis sets has been reported for any ring system. Consequently it has been difficult to determine the source of the small discrepancies between the SAMO and *ab-initio* results for these aromatic rings. The differences in question may well result from a facet of the SAMO method which is, as yet, unexplored. Indeed this facet could be one that will come into play only when SAMO is applied to certain classes of molecules: such as ring systems, unsaturated systems, or conjugated systems. However, the effect may be more specific and be observable only in the case of conjugated rings.

As a first step toward isolating these factors, it is desirable to make a comparative study of the application of the *ab initio* and SAMO methods to saturated ring systems using identical basis sets and geometries. For this purpose cyclohexane in both the chair and boat conformations has been examined. There are several advantages in this choice. First, a saturated ring system enables us clearly to distinguish ring effects from the effects of unsaturation and/or conjugation. Second, there are no strained bond angles; consequently the effects of geometry idealization are minimized and the transferability of Fock matrix elements is maximized. Third, six-membered rings are small enough that *ab initio* calculations are practical. Last, there are subtle geometric differences (other than bond angles and bond lengths) that give rise to an energy difference between the conformers. It would be gratifying were the SAMO method to prove as sensitive to these effects as are *ab initio* techniques.

We report here the results of our comparative study. Since the *ab initio* results represent a slight improvement over those previously available, we report them in some detail, including a component analysis for the energy difference.

Several levels of SAMO calculations are reported. These represent variations in the size of pattern molecules and two modifications to the basic method. The first modification is designed to account for steric effects and the second for multiple bonding pathways. The adequacy of the various sets of pattern molecules is examined in some detail.

Calculations

The geometry of both conformers was standardized. All bond angles were set equal to 109°28′; the C–C distances were set to 1.54 Å and the C–H distances to 1.10 Å. The final basis for both *ab initio* and SAMO calculations consisted of four sp^3 hybrid orbitals and one 1s orbital for each carbon and one 1s orbital for each hydrogen. These hybrid orbitals were obtained as linear combinations of optimized atomic orbitals which were in turn expressed as fixed linear combinations (or contractions) of a 5s 3p set of primitive Gaussian functions. The hydrogen orbital was formed from two 1s Gaussian functions. The optimized exponents and contraction coefficients were obtained from Whitman and Hornback.[7] A detailed discussion of the basis set is given in paper I[2] of this series.

The *ab initio* cyclohexane calculations were performed using the ATMOL system[8] and the ICL 1906A computer. The SAMO calculations were carried out on the ICL 1906A using our own recently developed suite of programs. These programs will be described in detail elsewhere, and they will be made available through the Quantum Chemistry Program Exchange.

SAMO. The SAMO method arises from the Eilers–Whitman observation[2] that matrix elements over the converged Hartree-Fock operator in an LCAO-SCF-MO treatment have closely similar numerical values for similar pairs of basis functions in similar molecules. Thus, for example, the Fock matrix elements for similar interactions in propane and butane are, since they are almost identical, found to be interchangeable. This transferability, however, is generally useful only when a basis set of hybridized atomic orbitals is employed. The explanation for this is that hybrids point along bonds and are thus determined by the internal reference framework of the molecules, whereas unhybridized s and p functions are defined by reference to an often arbitrary external coordinate system.

The SAMO method for closed-shell molecules generates the molecular orbitals and orbital energies for large molecules by a single solution of the eigenvalue problem

$$\mathbf{FC} = \mathbf{SC}\lambda$$

where $\mathbf{S}$ is the matrix of overlap integrals, λ is the diagonal matrix of eigenvalues or orbital energies, $\mathbf{C}$ is the matrix

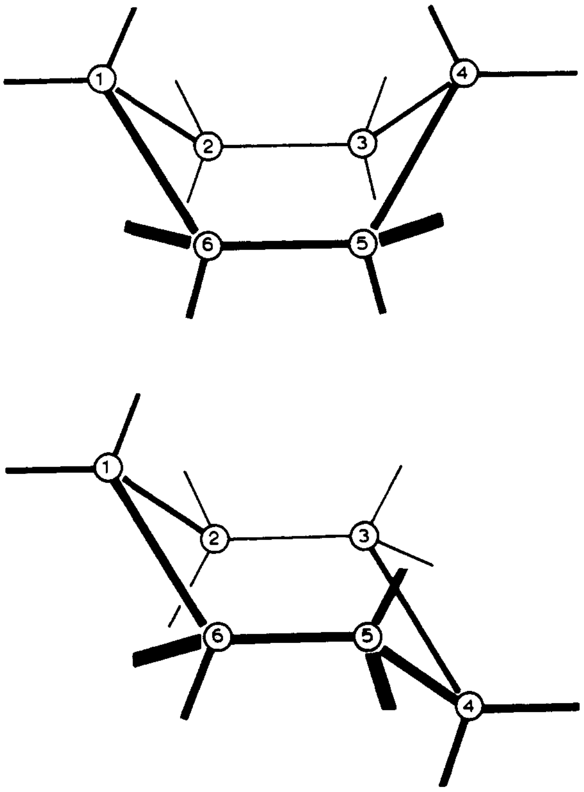

Figure 1. Boat and chair cyclohexanes.

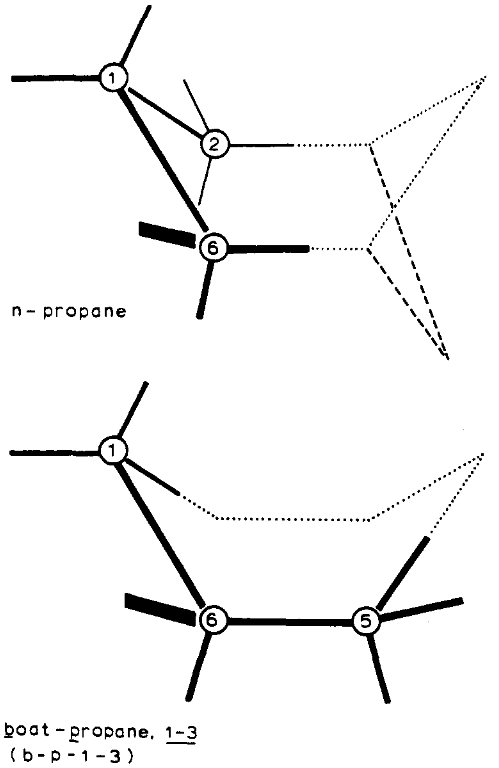

Figure 2. Propane pattern molecules set 1.

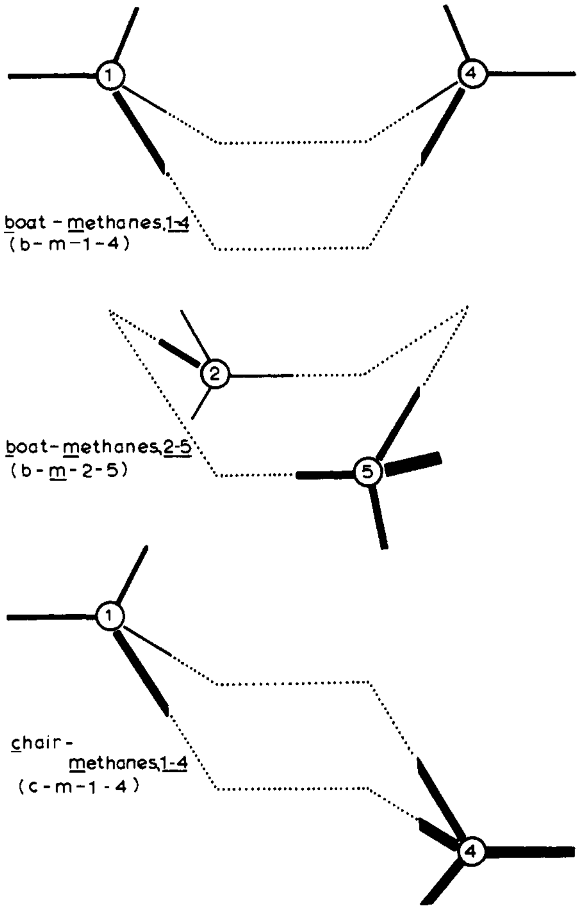

Figure 3. The dimethanes necessary for set II pattern molecules.

whose columns are the eigenvectors or LCAO expansion coefficients, and F is the matrix over the Hartree-Fock operator. The SAMO method obtains the elements of F by transferring values, truncated to four decimal figures, from *ab initio* calculations on similar small molecules known as "pattern molecules." In those cases where interactions between distant orbitals are unavailable from the pattern molecules, the Fock matrix elements are usually small and consequently can be set to zero when constructing the Fock matrix for the large molecule. The total energy is given by

$$E_{\mathrm{T}} = \sum_{\mathrm{occ}} (\epsilon_i^{(0)} + \lambda_i) + V_{\mathrm{NN}}$$

where the $\epsilon_i^{(0)}$ are the expectation values of the one-electron operator, the λ_i are the eigenvalues of the Hartree-Fock operator, V_{NN} is the nuclear repulsion energy, and Σ_{occ} is the summation over occupied molecular orbitals. Except for the F matrix, all terms are evaluated as in *ab initio* methods.

Pattern Molecules. The major parameters to be selected by any user of the SAMO technique are the sizes of the pattern molecules to be employed. It is to be expected that the larger the pattern molecules the more accurate the overall result. This consideration, however, must be balanced against the investment involved in obtaining the requisite *ab initio* calculations on the pattern systems.

In our original investigations on saturated chains[2] it was sufficient to choose pattern molecules containing three carbon centers. However, our investigation of aromatic ring systems[3] showed that pattern molecules containing four carbon centers were required for satisfactory results. With these considerations in mind the present study of cyclohexane has involved the use of three different sets of pattern molecules, each chosen to typify a different cost level.

Figure 1 illustrates the cyclohexanes. For convenience the methylene positions have been numbered. Figures 2, 3, and 4 show *all* the pattern molecules and their relation to the cyclohexanes. Figure 5 shows the numbering scheme for the orbitals of boat cyclohexane.

Set I. Propanes (Figure 2). Propanes represent the smallest feasible pattern molecules for simulating cyclohexane. Since pattern molecules of this size provide no "1,4" interactions, they must fail to include an adequate measure of the steric interaction between methylene groups 1 and 4 (Figure 1). But in the boat conformer these nonbonded in-

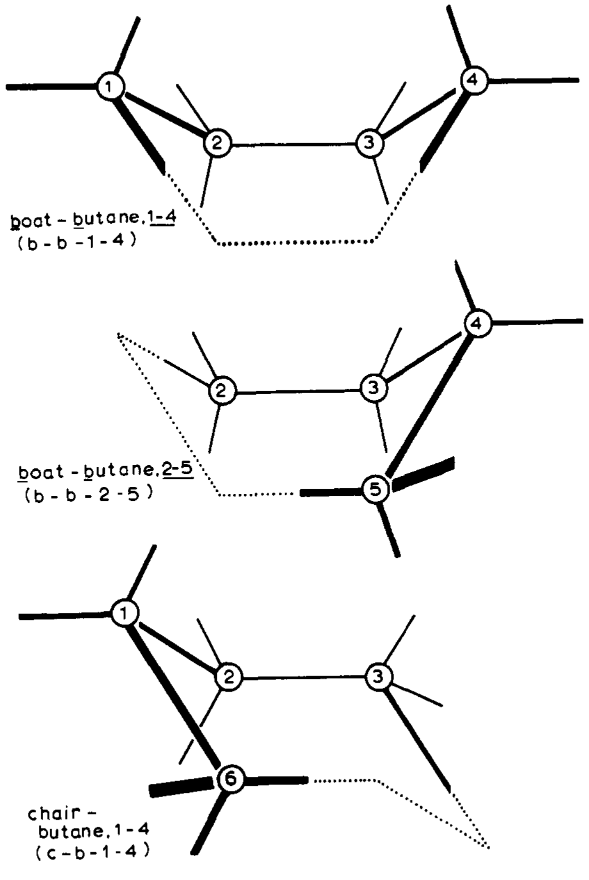

Figure 4. Butane pattern molecules set III.

teractions are likely to be critical, and they may also be necessary to a proper consideration of the chair conformer.

Set II. Propanes (Figure 2) + Methanes (Figure 3). This represents an inexpensive means of including "1,4" interactions. Earlier work[2] has indicated that the magnitude of the Fock matrix element drops off rapidly as the distance between the interacting orbitals increases. This suggests that the relative spatial positioning of the two orbitals is a most important factor in determining the value of a Fock element. If this is the case, *ab initio* calculations on systems consisting of two methanes, spatially located and oriented in exactly the same way as the distant methylene groups of the cyclohexanes (see Figure 3), might provide reliable estimates for those Fock elements that are unavailable from propane. Such a scheme, if successful, could be most useful in later work on long chains that bend back upon themselves, or in a supermolecule approach to reactions.

Set III. Butanes (Figure 4). Since the 1,4 interactions are now part of the pattern molecules, these molecules should provide quite accurate values for all elements necessary to the construction of the Fock matrix for cyclohexane.

Choosing the Fock Elements. The Fock matrix elements, F_{ij}, may be grouped into four classes according to the methylene groups in which the basis orbitals ϕ_i and ϕ_j are located: (a) intrafragment, both ϕ_i and ϕ_j belong to the same methylene group; (b) neighboring, ϕ_i and ϕ_j are within neighboring groups; (c) third center, a third methylene group lies between those containing ϕ_i and ϕ_j; (d) fourth center, ϕ_i and ϕ_j belong to the most distant groups.

Level I uses set I pattern molecules (Figure 2) and presents no real difficulties in choosing the source for the Fock matrix elements. Since the fourth center interactions were unavailable from propane pattern molecules, they were set to zero. Intrafragment interactions were obtained from the central methylene group of the appropriate propanes, since clearly the central methylene group and not the terminal methyls has an environment that most closely approximates that found in the cyclohexanes. Neighboring and third center interactions were taken from the appropriate propanes and, where necessary, by approximating the orbitals of a methylene group by using the corresponding orbitals in a methyl group. This approach has been previously shown to give satisfactory results.[2]

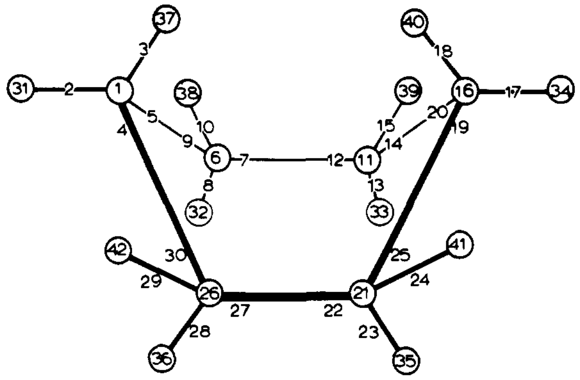

Figure 5. The orbital numbering scheme for boat cyclohexane.

Level II, using set II pattern molecules, is the same as level I except that estimates for all fourth center interactions are now included.

Level III, using butane pattern molecules, provides better estimates for the fourth center interactions because not only are the methyl groups correctly positioned to approximate the methylene groups but also an unbroken bonding pathway exists between the interacting orbitals. The shorter range interactions are also improved, as the need to approximate interactions involving methylene groups by consideration of similarly located methyl groups was much reduced for third center interactions and eliminated entirely for neighboring group interactions. Since the central groups in the butanes have an environment that more nearly approximates that of the methylene groups in the cyclohexane, intrafragment and neighboring group interactions were usually transferred from these central groups. For element $F_{4,14}$ terminal groups cannot be avoided; consequently, an average value from the two pattern molecules was used.

However, the intrafragment Fock matrix elements for methylene groups 1 and 4 of boat cyclohexane (Figure 1) require further consideration. It can be argued that steric interactions between the prows will cause the intrafragment interactions at the prows not only to differ from the other intrafragment interactions, but also to be more appropriately represented by the terminal groups of pattern molecule b-b-1-4 than by the central groups of pattern molecule b-b-2-5 (see Figures 1 and 4). The steric effects will be most noticeable for matrix elements such as $F_{3,3}$, $F_{37,37}$, and $F_{3,37}$ (see Figure 5). Consequently such elements were transferred from the b-b-1-4 pattern molecule in the level III calculations for boat cyclohexane.

The following further modifications also use set III pattern.

Level III(SA) Steric Adjustment. A more detailed consideration of how such steric interactions affect the intrafragment Fock matrix elements is possible; however, operationally, it is somewhat complex. As outlined in paper II of this series,[3] one procedure is to obtain a measure of the steric split experienced by the Fock matrix elements within the terminal groups of pattern molecule b-b-1-4 and to use this split to modify the more appropriate Fock matrix elements obtained from group 2 of pattern molecule b-b-2-5 (see Figures 4 and 5).

The data in Table I, used in conjunction with Figures 5

Table I. Steric Adjustments for Boat Cyclohexane

Fock element	*Ab initio*	b-b-1-4	b-b-2-5	Error^a $\times 10^4$	Adjusted value	Error^b $\times 10^4$
$F_{2,2}$	−0.5908	−0.5750	−0.5868^c	+40	−0.5868	+40
$F_{3,3}$	−0.5892	−0.5729^c	−0.5851	+163	−0.5851	+41
Δ	+0.0016	+0.0021	+0.0017		+0.0017	
$F_{31,31}$	−0.5892	−0.5857	−0.5915^c	−23	−0.5915	−23
$F_{37,37}$	−0.6070	−0.6024^c	−0.5944	+46	−0.6082	−12
Δ	−0.0178	−0.0167	−0.0029		−0.0167	
$F_{2,37}$	−0.1515	−0.1530	−0.1579^c	−64	−0.1526	−11
$F_{3,31}$	−0.1597	−0.1603^c	−0.1599	−06	−0.1599	−1
Δ	−0.0082	−0.0073	−0.0020		−0.0073	
$F_{3,37}$	−0.6788	−0.6741^c	−0.6720	+47	−0.6803	−15
$F_{2,31}$	−0.6675	−0.6635	−0.6697^c	−22	−0.6697	−22
Δ	+0.0113	+0.0106	+0.0023		+0.0106	
$F_{2,4}$	−0.0994	−0.1041	−0.1035^c	−41	−0.1004	−10
$F_{3,4}$	−0.1131	−0.1169^c	−0.1100	−38	−0.1132	−01
Δ	−0.0137	−0.0128	−0.0065		−0.0128	

[a] Error = level III value minus the *ab initio* value. [b] Error = the adjusted value minus the *ab initio* value. [c] Elements used in level III procedure.

and 4, will serve to demonstrate this technique. Column two contains values for the boat cyclohexane matrix when obtained from the terminal group of molecule b-b-1-4; while column three contains values for the same intrafragment Fock matrix elements obtained from group 4 of the b-b-2-5 pattern molecule. The values taken from these two sources are significantly different, as is usually the case when considering terminal *vs.* central groups, even when no steric hindrance is present.

Irrespective of which pattern molecule is considered, we see that all elements involving orbital 2 differ from the similar ones involving orbital 3, even though orbitals 2 and 3 are both sp^3 hybrid orbitals pointing toward a hydrogen and are on the same center. Likewise those matrix elements involving orbital 31 differ from those involving orbital 37. This difference (which we call the steric split for such pairs) can be attributed to the different geometric relation of the orbitals in question to the rest of the molecule. Now in the case of the b-b-2-5 pattern molecule the observed split cannot be due to any "prow to prow" steric effect, but rather must reflect the different proximities of orbitals 2 and 3 (Figure 5) to the methylene groups 3 and 5 (Figure 1). Clearly, both this third center effect and a "prow to prow" effect are present in the b-b-1-4 pattern molecule. Therefore, when the "prow to prow" steric effect is important to a pair of matrix elements, the split (Δ) from b-b-1-4 will be much larger than the Δ observed in pattern molecule b-b-2-5. Thus, Table I shows that $F_{2,2}$ and $F_{3,3}$ are only slightly split by the "prow to prow" steric effect while the other elements listed are strongly affected.

When as for $F_{2,2}$ and $F_{3,3}$ there is no steric split, we use the values from b-b-2-5. When the "prow to prow" steric effect is significant, we must modify the b-b-2-5 values to take the "prow to prow" steric effect into account. We allow for this as follows. (a) Since $F_{31,31}$ is unlikely to be strongly influenced by the steric effect, it is given its values from b-b-2-5 of 0.5915. The value for $F_{37,37}$ is then obtained by imposing the split observed in pattern molecule b-b-1-4 onto the $F_{31,31}$ value from b-b-2-5 [*i.e.*, $F_{31,31} + \Delta = -0.5915 + (-0.0167) = -0.6082$]. Note that this approach only changes the value for one matrix element in the split pair. (b) On the other hand, where both elements of the pair are apparently influenced by steric effects, it is best to superimpose the split from b-b-1-4 upon the mean value observed in b-b-2-5. An example of this is the $F_{2,4}$ and $F_{3,4}$ pair, which are assigned values as follows.

$$F_{2,4} = \tfrac{1}{2}(-0.1035 - 0.1100) - \tfrac{1}{2}(-0.1169 + 0.1041) = -0.1004$$

$$F_{3,4} = \tfrac{1}{2}(-0.1035 - 0.1100) + \tfrac{1}{2}(-0.1169 + 0.1041) = -0.1132$$

Column six in Table I lists the values given by this procedure, while columns five and seven show the improvement achieved by the level III(SA) procedure. Such modifications are unnecessary for the chair conformer.

Bonding Pathways and Long Range Fock Elements Level III(PA). Earlier work[2,3] has shown that the magnitude of F_{ij} drops off rapidly as ϕ_i and ϕ_j get farther apart. However, this decay is not uniform; indeed, some long range Fock elements have small positive values. Moreover, the behavior of nmr coupling constants suggests that the nature and length of the bonding pathway(s) between ϕ_i and ϕ_j may have as much to do with the magnitude of the interaction as does the character and relative spatial positions of the orbitals.

We studied the question of the bonding pathway's influence on the fourth center Fock matrix elements by examining Fock matrix data from three sources: (1) properly positioned methanes (Figure 3); (2) butanes; and (3) from the complete *ab initio* SCF calculations on the cyclohexanes. As is seen in Table II, the values from both the methanes and the butanes are in reasonable, but not exact, agreement with each other and with the *ab initio* ones. It is reasonable to assume that the difference between the values obtained from methanes and the values from butanes is due to electrostatic forces transmitted along the bonds, or more precisely due to the presence of nuclei in the region between the methylene groups under consideration. Likewise the difference between the butane values and the true self-consistent field cyclohexane values can be attributed to the presence of a second bonding pathway in the cyclohexane.

This is easily checked for elements such as $F_{3,18}$ or $F_{2,34}$ where the two pathways are, by symmetry, identical. The above arguments lead to the following equation relating these Fock matrix elements

$$F(\text{cyclohexane}) = F(\text{methanes}) + 2[F(\text{butane}) - F(\text{methane})]$$

Elements such as $F_{4,19}$, where the two pathways are unequal, require

$$F(\text{cyclohexane}) = F(\text{methanes}) + [F_{\text{short}}(\text{butane}) + F_{\text{long}}(\text{butane}) - 2F(\text{methanes})]$$

Here, we add a separate correction for each pathway.

Table II. Second Bonding Pathway Adjustments for Prow–Prow Interactions in Boat Cyclohexane

Fock matrix element	From methanes	From butanes	From *ab initio*	Adjusted[a] for second pathway
1–16	−0.0002	−0.0002	−0.0002	−0.0002
2–16	+0.0591	+0.0594	+0.0594	+0.0597
2–17	−0.0132	−0.0194	−0.0257	−0.0256
2–18	+0.0586	+0.0603	+0.0619	+0.0620
2–19	+0.0586	+0.0716	+0.0732	+0.0733
2–34	+0.0154	+0.0159	+0.0165	+0.0164
2–40	+0.0680	+0.0698	+0.0718	+0.0716
3–16	−0.0240	−0.0241	−0.0241	−0.0242
3–17	+0.0587	+0.0603	+0.0619	+0.0620
3–18	−0.0889	−0.0912	−0.0935	−0.0935
3–19	−0.0481	−0.0504	−0.0503	−0.0503
3–34	−0.0079	−0.0080	−0.0081	−0.0081
3–40	−0.1173	−0.1188	−0.1202	−0.1203
4–16	−0.0240	−0.0243	−0.0242	−0.0245
4–17	+0.0586	+0.0716	+0.0732	+0.0733
4–18	−0.0481	−0.0504	−0.0503	−0.0503
4–19	−0.0889	−0.1103	−0.1125	−0.1126
4–20	−0.0481	−0.0504	−0.0524	−0.0529
4–34	−0.0079	−0.0080	−0.0082	−0.0081
4–40	−0.0504	−0.0536	−0.0538	−0.0540
31–16	−0.0001	−0.0001	−0.0001	−0.0001
31–17	−0.0154	−0.0159	−0.0165	−0.0164
31–18	−0.0079	−0.0080	−0.0082	−0.0081
31–19	−0.0079	−0.0080	−0.0082	−0.0081
31–34	+0.0024	+0.0011	+0.0006	+0.0008
31–40	−0.0080	−0.0079	−0.0079	−0.0078
37–40	−0.1496	−0.1512	−0.1528	−0.1528

[a] Butane values plus an adjustment determined from the difference observed in level II and level III elements.

We find that while the sign and size of the Fock elements for such long range interactions are primarily determined by the type and spatial positioning of the interacting orbitals, the presence of additional intervening nuclei not present in the pattern molecule will make a noticeable adjustment. Moreover, the size of this adjustment depends upon the length of the bonding pathways, and the effects are precisely additive when more than one pathway is present.

Table II shows this effect and adjusted values for the "prow to prow" interactions in boat cyclohexane. Most of these fourth center interactions are small and these adjustments to them are still smaller; thus we cannot expect this modification to have much effect on the SAMO calculation.

Results

Ab Initio. Few theoretical calculations have been reported on the cyclohexanes; the available results are summarized in Table III along with the experimental values for the ionization potential and the chair–boat energy difference. All the calculations seem to predict adequately the energy difference and ionization potentials.

SAMO. The energies obtained from the SAMO calculations are compared to the *ab initio* SCF results in Table IV. All levels (shown in Table IV in order of increasing sophistication) do well in regard to total energy, the best values being accurate to 1 part in 10^4. This accuracy is not reflected in the chair–boat energy difference because: (a) it is obtained as a small difference between large numbers; (b) at each level the chair calculation is closer to the *ab initio* than is the boat calculation; (c) at the third level the deviations of the chair and boat calculations from the *ab initio* ones are of different sign.

In general, the use of more accurate Fock elements, progressing toward level III(PA), succeeds in more accurately mimicking the *ab initio* results.

SAMO-level I, where we ignore fourth center interactions, predicts the boat conformer to be more stable by 219 kcal/mol. This clearly unsatisfactory result is surprising, since one might expect that ignoring fourth center electronic terms and including long range nuclear–nuclear repulsions would, while destabilizing both the boat and the chair, have a greater effect on the boat. The chair should also be favored through consideration of staggered and eclipsed C–H bonds.

SAMO-level II which includes approximate fourth center interaction terms from methanes lowers the barrier to 44 kcal/mol, but still in the wrong direction.

All SAMO-level III calculations which use butane pattern molecules correctly predict the chair to be the more stable conformer. SAMO-III gives a difference of 70 kcal/mol (ten times the experimental value). A comparison of the SAMO-III and the *ab initio* Fock matrices showed larger deviations in the intrafragment Fock elements than in the smaller fourth center elements. Improving these intrafragment Fock matrix elements with the steric adjustment, SAMO-III(SA), lowers the difference to 39.5 kcal/mol. SAMO-III(PA), which adjusts fourth center interactions for pathway effects, does not improve the calculation with respect to energy.

Tables V and VI compare the *ab initio* and SAMO eigenvalue spectra. The general features and ordering are well reproduced. Level II shows a few moderately disparate values (±5%) and three cases where an energy level is "out of order" as defined by the *ab initio work*. The level III approach (using butanes) removes some of these discrepancies and halves the root mean square deviation; however, the $7a_1$ orbital is out of order.

Level III(SA) shows a modest overall improvement and a significant improvement of orbitals $3b_1$, $7a_1$, and $10a_1$ for the boat conformer. Level III(PA) gives slightly improved values for most of the occupied orbitals. However, the $8a_1$ orbital in boat is more deviant and the $1a_{1u}$ orbital in chair now becomes out of order.

The total orbital populations from the *ab initio* and SAMO calculations are presented in Tables VII and VIII. The agreement is good for SAMO-II, SAMO-III, SAMO-III(SA), and SAMO-III(PA), but only fair for SAMO-I. Surprisingly SAMO-II gives better agreement than the level III results.

Component Analysis of the Energy Difference. Table IX

Table III. *Ab Initio* Results

	Energy, hartrees		Energy difference (chair–boat), kcal/mol	Ionization potential, eV	
	Chair	Boat		Chair	Boat
This work	−233.149238	−233.134296	9.41	12.36	12.18
Experimental[a]			6.59[a]	10.4[e]	
Hoyland[b]	−232.910588	−232.899136	7.19		
Preuss[c]	−229.6239			9.34	
CNDO[d]			5.02	13.8	
NDDO[d]			4.83	11.8	

[a] Reference 9. [b] Reference 10. [c] Reference 11. [d] Reference 12. [e] Reference 13.

Table IV. Comparison of Energies (SAMO *vs.* SCF)

Conformer	Type of calculation	Energy, au	Deviation from SCF	ΔE between conformers
Chair	*Ab initio* SCF	−233.14928		
	Level I	−233.40119	−0.25191	
	Level II	−233.11917	+0.03011	
	Level III	−233.16847	−0.01919	
	Level III(PA)	−233.17056	−0.02128	
Boat	*Ab initio* SCF	−233.13430		+0.01498
	Level I	−233.75513	−0.62083	−0.35394
	Level II	−233.19167	+0.05738	−0.07250
	Level III	−233.05646	+0.07784	+0.11201
	Level III(SA)	−233.10578	+0.01752	+0.06269
	Level III(PA)	−233.10417	+0.03013	+0.06639

Table V. Eigenvalue Spectra Chair Cyclohexane

		Deviation from SCF × 10^4		
Molecular orbital	*Ab initio* SCF, au	SAMO-III (PA)	SAMO-III	SAMO-II
$1e_u$	−11.3570	−46	−46	−44
$2e_u$	−11.3570	−46	−46	−44
$1e_g$	−11.3570	−46	−46	−44
$2e_g$	−11.3570	−46	−46	−44
$1a_{1g}$	−11.3569	−46	−46	−46
$1a_{2u}$	−11.3569	-48^a	-48^a	−44
$2a_{1g}$	−1.1647	+21	+42	+50
$3e_u$	−1.0443	−13	−23	−63
$4e_u$	−1.0443	−13	−23	−63
$3e_g$	−0.8607	−12	−04	−64
$4e_g$	−0.8607	−12	−04	−64
$2a_{2u}$	−0.7848	+02	+04	−12
$3a_{1g}$	−0.6766	+72	+99	+282
$3a_{2u}$	−0.6296	+11	+01	$+13^a$
$5e_g$	−0.6266	−02	−11	−33
$6e_g$	−0.6266	−02	−11	−33
$5e_u$	−0.5546	−08	−51	−212
$6e_u$	−0.5546	−08	−51	−212
$1a_{1u}$	−0.5522	-47^a	−56	0
$7e_u$	−0.5105	−46	+21	+19
$8e_u$	−0.5101	−46	+21	+19
$4a_{1g}$	−0.4811	+19	+04	−10
$7e_g$	−0.4543	−24	+09	+19
$8e_g$	−0.4543	−24	+09	+19
$5a_{1g}$	+0.4090	+186	+176	$+374^a$
$4a_{2u}$	+0.4305	+18	−17	−08
$9e_u$	+0.4493	+71	−19	+43
$10e_u$	+0.4493	+71	−19	+43
$9e_g$	+0.4778	+190	−60	−115
$10e_g$	+0.4778	+190	−60	−115

[a] Predicted out of proper order.

Table VI. Eigenvalue Spectra of Boat Cyclohexane

		Deviation from SCF × 10^4			
Molecular orbital	*Ab initio* SCF, au	SAMO-III(PA)	SAMO-III(SA)	SAMO-III	SAMO-II
$1a_2$	−11.3556	−45	−45	−45	−58
$1b_2$	−11.3556	−45	−45	−45	−58
$1b_1$	−11.3556	−63	−63	−63	−59
$1a_1$	−11.3556	−61	−61	−61	−57
$2b_1$	−11.3553	−49	−49	−49	−29
$2a_1$	−11.3551	−49	−49	−49	−29
$3a_1$	−1.1689	+29	+42	+30	+80
$2b_2$	−1.0444	−12	−21	−21	−54
$3b_1$	−1.0381	−02	−08	−35	−83
$4a_1$	−0.8799	−09	+05	−05	−40
$2a_2$	−0.8600	−03	+03	+03	−78
$4b_1$	−0.7495	−07	+05	−01	−40
$5a_1$	−0.6909	+25	+47	+56	+19
$6a_1$	−0.6798	−86	+118	+119	+302
$3b_2$	−0.6282	+01	−13	−13	−32
$5b_1$	−0.5770	+16	−05	+16	−66
$4b_2$	−0.5554	−19	−57	−55	−192
$7a_1$	−0.5457	-31^a	$+45^a$	$+90^a$	−09
$5b_2$	−0.5441	−09	−63	−63	+11
$6b_1$	−0.5409	+16	−10	+06	-247^a
$3a_2$	−0.5163	−02	+16	+15	−51
$8a_1$	−0.4662	+77	−34	−38	−15
$7b_1$	−0.4546	−09	−24	+19	−24
$4a_2$	−0.4513	−07	+33	+33	+07
$9a_1$	+0.3956	+55	+36	+34	+69
$10a_1$	+0.4042	+50	−45	+132	+118
$11a_1$	+0.4341	+09	+41	+46	+49
$6b_2$	+0.4377	+100	+41	+14	+20
$8b_1$	+0.4671	+01	−34	−06	−147
$7b_2$	+0.4753	+17	−51	−57	−52

[a] Predicted out of order.

Table VII. Orbital Populations for Chair Cyclohexane

	Ab initio SCF	SAMO-III (PA)	SAMO-III	SAMO-II	SAMO-I
1s (C)	2.0036	2.0037	2.0037	2.0036	2.0036
C–C	0.9949	0.9952	0.9952	0.9955	0.9984
C–H (ax)	1.1749	1.1787	1.1787	1.1754	1.1761
C–H (eq)	1.1796	1.1756	1.1760	1.1737	1.1867
H (ax)	0.8284	0.8250	0.8249	0.8290	0.8290
H (eq)	0.8235	0.8266	0.8264	0.8272	0.8078

examines the components of the barrier in scaled[14-16] and unscaled form for both the *ab initio* results and the best SAMO result.

Scaling changes the sign of only the kinetic energy component; the barrier is repulsive dominant by all four interpretations. The SAMO results badly underestimate the nuclear attraction component. This is partially offset by compensating deviations in kinetic energy and electron-electron repulsion components.

Discussion

As has been the case with previous SAMO applications, the total energy, the orbital energies, and the distribution of electrons in the basis orbitals are in excellent agreement with the *ab initio* results. The agreement here is comparable to that found for unsaturated chains and somewhat better than for benzene and naphthalene.

In general the use of more accurate Fock matrix elements has led to better calculations (see Table IV). The use of four-carbon pattern molecules (SAMO-III) gives significantly better results than does the SAMO-II approach, a calculation which also includes estimates for all interactions. We would emphasize that the improvement in the matrix elements is not confined to the fourth center interactions; the others are, by being in an environment that more nearly reflects that found in cyclohexane, also improved. The use of the steric adjustment, proposed earlier[3] on largely intuitive ground, has been shown to give more accurate individual Fock matrix elements and to improve the energy calculations. Employing the pathways adjustment produces better Fock elements for fourth center interactions, but the effect on the overall calculation is mixed: slightly better eigenvalues and orbital populations but a marginally poorer energy.

We have shown that the magnitude of a Fock element depends upon (a) the types of orbitals involved, (b) their spatial location and orientation relative to one another, and (c) the length of the bonding pathway(s) between them. However, this does not preclude the use of schemes where long range interactions are estimated from two properly positioned but nonbonded small molecules when the bonding pathway in the large molecule contains more than five or six bonds or when considering two unbound molecular moieties in a "near approach." This type of approach could prove quite valuable for studies of coiled molecules or the study of transition states for enzyme reactions.

At all levels of SAMO the calculation on the chair con-

Table VIII. Orbital Populations for Boat Cyclohexane

Orbital (Figure 5)	Type	Methylene group (Figure 1)	SCF	SAMO-III(PA)	SAMO-III(SA)	SAMO-III	SAMO-II	SAMO-I
1	1s	1	2.0037	2.0037	2.0037	2.0037	2.0036	2.0033
6	1s	2	2.0036	2.0036	2.0036	2.0036	2.0036	2.0036
5	C–C	1	0.9936	0.9937	0.9941	0.9963	0.9972	1.0187
9	C–C	2	0.9954	0.9965	0.9961	0.9943	0.9941	1.0025
7	C–C	2	0.9958	0.9966	0.9966	0.9965	0.9973	0.9931
3	C–H (ax)	1	1.1707	1.1608	1.1628	1.1542	1.1705	1.1330
2	C–H (eq)	1	1.1808	1.1702	1.1734	1.1677	1.1818	1.2200
8	C–H (ax)	2	1.1779	1.1802	1.1802	1.1802	1.1771	1.1776
10	C–H (eq)	2	1.1800	1.1762	1.1765	1.1765	1.1800	1.1888
37	H (ax)	1	0.8314	0.8406	0.8389	0.8472	0.8338	0.8445
31	H (eq)	1	0.8261	0.8308	0.8237	0.8240	0.8261	0.7626
32	H (ax)	2	0.8216	0.8237	0.8271	0.8317	0.8162	0.8265
38	H (eq)	2	0.8236	0.8265	0.8263	0.8263	0.8237	0.8074

Table IX. The Cyclohexane Chair–Boat Energy Difference by Components

	Ab initio		SAMO-III(SA)	
	Unscaled	Scaled	Unscaled	Scaled
$\Delta E_{\mathrm{Total}}$	+0.01498	+0.01501	+0.06268	+0.06266
Δh_0	−2.76565	−2.56765	−2.687730	−2.41078
ΔV_{ee}	+1.39068	+1.27656	+1.3604601	+1.2017673
ΔV_{NN}	+1.38995	+1.30610	+1.38995	+1.27168
ΔT	+0.08184	−0.11116	+0.06225	−0.19397
ΔV_{Ne}	−2.84749	−2.45649	−2.74998	−2.2168
ΔV_{At}	−2.76565	−2.56765	−2.68773	−2.41078
ΔV_{rep}	+2.78066	+2.58266	+2.750409	+2.47344

former is nearer to the *ab initio* result than is the analogous calculation on the boat conformer (Table IV). This is not surprising since it is in the boat form that steric effects are most pronounced. Although applying the steric adjustment to the boat conformer results in a reduction of its energy deviation (from +0.078 to +0.029), this deviation is still larger than that of the chair conformer (−0.019). Fundamentally, the chair form would seem to be more amenable to the SAMO technique. There are two pattern molecules for boat cyclohexane but only one for the chair; also all methylene groups in the chair form are equivalent. Thus, although for both conformers one must occasionally choose between two different, but apparently equally appropriate, sources for certain Fock elements, the need to make this choice arises more frequently in boat cyclohexane. Further, when choosing third and fourth center Fock elements, one must occasionally approximate a hybrid orbital pointing to a carbon atom by one which, in the "pattern molecule," points to a hydrogen atom. Again, this approximation arises more frequently in the boat form.

The calculation on the energy difference for the cyclohexane conformers suffers, as do all such calculations, from being the result of subtracting two large numbers. However, in *ab initio* calculations, it is at least known that the energy calculated for each conformer is higher than the "true value." Unfortunately, the variation theorem does not apply to our scheme; and, as we have again seen in this work, the molecular energies calculated by SAMO lie in a narrow region about the *ab initio* result. This randomness can, in extreme cases, lead to energy differences that are incorrect in both sign and size. We therefore suggest caution in using SAMO for energy differences or conformational analysis when (a) the energy difference is very small or (b) it is reasonable to expect one form to be more amenable to a SAMO treatment than the other.

We have pointed out that in this work we obtained an agreement between SAMO and *ab initio* results comparable to that found for unsaturated chains and better than that found for benzene and naphthalene. Although it is worth remembering that in our work with aromatic rings we used a basis set that differed from those of the published *ab initio* calculations, we feel this to be merely a contributory factor, rather than the fundamental reason for the poorer agreement of the SAMO and *ab initio* calculations on benzene and naphthalene. When SAMO is applied to benzene-like systems we observe π orbital compression and/or an abnormally low energy for the first virtual orbital (π symmetry). Although these appear to be related rather than independent phenomena, it is difficult to pinpoint their cause. Neither effect is observed in the treatment of polyene,[4] a conjugated system of infinite length, or in cyclohexane rings. Since we only observe this behavior when conjugated unsaturation is present in a ring, we feel that it probably arises from SAMO's treatment of fourth center interactions in benzene rings. Recognizing that in the benzene ring there are two bonding pathways through which the orbitals may "interact," we feel that a modification similar to that used to correct for the second bonding pathway in the cyclohexanes should have been applied to all fourth center interactions in benzene, naphthalene, anthracene, and phenanthrene.

Conclusion

For both chair and boat cyclohexane the total energy, the orbital energies, and the distribution of electrons in the basis orbitals are in excellent agreement with the *ab initio* results. The use of a steric adjustment for boat cyclohexane has led to more accurate individual Fock matrix elements and to an improved energy. The introduction of the bonding pathways modification, although it has had little marked effect here, may well prove valuable in future work. We stress caution in using SAMO for energy differences or conformational analysis (a) when the energy difference is very small or (b) when it is reasonable to expect one conformer to be more amenable to a SAMO treatment than the other.

References and Notes

(1) (a) This work was initiated at Case Western Reserve University; (b) State University College at Brockport, Brockport, N.Y.; (c) The University of Lancaster (England); (d) Fairleigh Dickenson University; Teaneck, N.J.; (e) Case Western Reserve University, Cleveland, Ohio.

(2) J. E. Eilers and D. R. Whitman, *J. Amer. Chem. Soc.*, **95**, 2067 (1973).

(3) J. E. Eilers, B. O'Leary, A. Liberles, and D. R. Whitman, *J. Amer. Chem. Soc.*, submitted for publication.

(4) B. J. Duke and B. O'Leary, *Chem. Phys. Lett.*, **20**, 459 (1973).

(5) B. J. Duke, J. E. Eilers, and B. O'Leary, *J. Chem. Soc., Faraday Trans. 2*, **70**, 386 (1974).

(6) B. O'Leary, B. J. Duke, J. E. Eilers, and E. W. Abrahamson, *Nature, (London)*, **246**, 166 (1973).

(7) D. R. Whitman and C. J. Hornback, *J. Chem. Phys.*, **51**, 398 (1966); C. J. Hornback, Ph.D. Thesis, Case Institute of Technology, 1967.

(8) V. R. Saunders, ATMOL2 Program, Atlas Computer Laboratory, Chilton, Berks, England.

(9) N. L. Allinger, J. A. Hirsch, M. A. Mille, I. J. Tyminsky, and F. A. Van Catledge, *J. Amer. Chem. Soc.*, **90**, 1199 (1968).

(10) J. R. Hoyland, *J. Chem. Phys.*, **50**, 2775 (1969).

(11) H. Preuss and R. Janoschek, *J. Mol. Struct.*, **3**, 423 (1969).
(12) R. B. Davidson, W. L. Jorgensen, and L. C. Allen, *J. Amer. Chem. Soc.*, **92**, 749 (1970).
(13) F. H. Field and J. L. Franklin, "Electron Impact Phenomenon," Academic Press, New York, N.Y., 1957, Table 10A.
(14) J. O. Hirschfelder and J. F. Kincaid, *Phys. Rev.*, **52**, 658 (1937).
(15) P. O. Löwdin, *J. Mol. Spectrosc.*, **3**, 46 (1959).
(16) L. C. Allen, *Chem. Phys. Lett.*, **2**, 597 (1968); W. H. Fink and L. C. Allen, *J. Chem. Phys.*, **46**, 2261 (1967); R. B. Davidson and L. C. Allen, *J. Chem. Phys.*, **54**, 2828 (1971).

Localized Orbital Studies of Hydrogen Bonding. II. Dimers Containing H_2O, NH_3, HF, H_2CO, and HCN

Mark S. Gordon,*[1] Dennis E. Tallman, Cathy Monroe, Michael Steinbach, and Jill Armbrust

Contribution from the Department of Chemistry, North Dakota State University, Fargo, North Dakota 58102.

Received June 5, 1974

Abstract: INDO localized molecular orbitals (LMO's) are utilized for investigating the nature of intermolecular hydrogen bonding in the fully geometry optimized dimers $(HF)_2$, $(H_2O)_2$, $(NH_3)_2$, FH-OH_2, HOH-FH, FH-NH_3, H_2NH-FH, H_2O-HNH_2, HOH-NH_3, HCN-HF, and H_2CO-HF. The results suggest that a reasonable measure of relative hydrogen bond strengths should be the intrabond, two-center, one-electron interference energy connecting the acceptor atom and donated proton. This approach views the net stabilization energy of a hydrogen bonded dimer as arising from a large energy decrease due to formation of the hydrogen bond, modified by smaller energy increases due to internal decreases in monomer bond energies upon formation of the dimer. Hydrogen bond stabilization appears to be closely related to the extent of charge transfer within the hydrogen bonded complex. The calculated transfer of charge can largely be explained in terms of electron density shifts within the acceptor lone pairs, while the decrease in electron density on the proton is discussed in terms of the donor XH bond. The approach presented should be particularly useful for analyzing intramolecular hydrogen bonding systems where the hydrogen bond energy is not simply obtainable from monomer–dimer energy differences.

I. Introduction

There has been considerable literature recently involving the application of quantum mechanics to the study of hydrogen bonding. Most of these investigations have been on the molecular orbital level (*ab initio*[2-15] or semiempirical[2,16,17]) and primarily concerned with elucidating the nature of *inter*molecular hydrogen bonding. Our understanding of the hydrogen bond has been vastly improved by this effort; however, the computational results on systems capable of forming *intra*molecular hydrogen bonds are few in comparison. Pople and coworkers carried out minimal basis *ab initio* calculations on a series of 1,2-disubstituted ethanes;[18] however, their effort was mainly concerned with the possibility of attractive 1,3 interactions in particular molecular conformations and the stability of such conformations relative to those in which hydrogen bonding is unlikely. While these authors concluded that hydrogen bonding stabilizes a configuration, little attempt was made at formulating a quantitative measure of the strength of these hydrogen bonds. Johansson, Kollman, and Rothenberg[19] studied intramolecular hydrogen bonding in ethylene glycol using an STO 3G basis set.

In fact, aside from sheer molecular size, the major difficulty in calculations on intramolecularly hydrogen bonded systems is to find an energy quantity which can be associated with a hydrogen bond energy, at least on a relative scale. For an intermolecular system, the problem is simple since the hydrogen bond energy is just the difference in energy between the dimer and separated monomers. Unfortunately, such a simple approach is not applicable when there is only one molecule, the complex, to consider.

Attempts have been made[16,19] to estimate the energy associated with an intramolecular hydrogen bond using the energy difference between the complex and some "nonhydrogen-bonding" conformation of the same molecule. While there is some merit in this approach, it suffers from the assumption that all of the energy difference between the two conformations may be attributed to the formation of the hydrogen bond.

In the present work an attempt is made, on the basis of calculations on small hydrogen-bonded dimers, to define an energy quantity which correlates with the calculated hydrogen bond energies on the one hand, and which is accessible in a calculation on an intramolecular complex on the other. The intermolecular systems are used as a starting point for the very reason stated above: the energy differences in these cases *are* available for comparison.

To carry out the analysis use is made of localized molecular orbitals[20] (LMO's) obtained from INDO[21] canonical orbitals. Although a semiempirical method is used in the present work to simplify the calculations, the basic approach is extendable to nonempirical methods. The use of localized orbitals aids the calculation and analysis by providing a theoretical description of the bonds in the monomers and of the way in which these bonds change on formation of a hydrogen-bonded complex. Further, since hydrogen bonding is often thought of as interaction between a donor hydrogen and an acceptor lone pair, analysis of the LMO's provides a convenient means of investigating such characteristics as hybridization and availability of lone pairs, hydrogen bond strength *vs.* dipole moment enhancement relationships, and changes in lone pair electron density distributions on hydrogen bond formation.

All of the complexes studied in the present work have been subjected to previous MO calculations, either *ab initio*[2-15] or semiempirical[2,16,17] or both; however, two aspects of the present study seem to be unique. First, little use has been made previously of the application of energy localization to the study of hydrogen bonding, while, as stated above, we feel an analysis using LMO's should be particu-