Tetrahedron. Vol. 27, pp. 689-700. Pergamon Press 1971. Printed in Great Britain

CALCULATIONS OF GEOMETRIES OF ORGANIC MOLECULES USING THE CNDO/2 METHOD—I

EMPIRICAL CORRELATIONS BETWEEN OBSERVED AND CALCULATED BOND LENGTHS IN SIMPLE ACYCLICS, STRAINED CYCLOALKENES AND SOME POLYCYCLIC MOLECULES*

C. S. CHEUNG,† M. A. COOPER‡ and S. L. MANATT
Space Sciences Division, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91103

(*Received in the USA* 14 *July* 1970; *Received in the UK for publication* 26 *August* 1970)

Abstract—It is shown that calculations of C—C and C—H bond lengths in hydrocarbons using the CNDO/2 semi-empirical MO method exhibit systematic deviations from the observed values. An empirical correlation relating the observed and calculated bond lengths with the number of substituents attached to the bond may be devised. This correlation is capable of providing theoretical bond lengths within 0·008 Å of the experimental for a wide range of simple acyclic molecules. Furthermore, calculations of molecular geometries for some strained cycloalkenes and also some larger molecules, e.g., naphthalene, biphenylene and azulene, are similarly found to be in good agreement with experiment.

WE HAVE recently been investigating the dependence of proton–proton spin–spin coupling constants on subtle aspects of molecular structure.[1–5] In this work, the lack of accurate experimental bond length and bond angle data even for molecules of basic chemical interest has often been a considerable hindrance to the detailed interpretation of our results. This impasse led us to consider using approximate MO theories for the theoretical calculations of molecular geometries, and in particular the CNDO/2 program of Pople *et al.*[6,7] was selected. Although energies calculated by this program are unsatisfactory for quantitative comparison,[8] the basic trends in bond angles and bond lengths seem quite well reproduced. Further, it has been suggested that the semiempirical matrix elements of this program may be reparametrized[9,10] with the aim of improving its reliability in regard to specific applications. Even after reparametrization,[10] however, the agreement between observed and calculated C—C bond lengths over a range of small molecules is no better than ±0·02 Å. This falls considerably short of the presently attainable experimental accuracy (*ca* ±0·005 Å). At first sight then, the CNDO/2 treatment appears insufficiently accurate for many direct quantitative applications, particularly when subtle geometrical effects may be involved. As will be discussed in the present work this restriction can be partially overcome.

We have carefully examined the geometries calculated for simple hydrocarbon

---

* This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

† NASA Resident Research Associate, 1969–71.

‡ NASA Resident Research Associate, 1968–70.

molecules by CNDO/2 (e.g., see Ref 10), and find that in the main the deviations between observed and calculated C—C bond lengths follow consistent patterns. We shall show that it is possible to obtain an empirical correlation between observed and calculated C—C bond lengths which is accurate to an rms error of *ca* ±0·008 Å. We believe that quite apart from the immediate usefulness of this correlation in studies of molecular structure, this result may suggest ways in which CNDO/2 can be modified to give a better account of molecular properties.

Previous theoretical studies of molecular structure which consider all valence electrons have most often been restricted to small molecules of high symmetry, with benzene typically being the largest molecule studied. Indeed, with the agreement between observed and calculated bond lengths exhibited by the initial CNDO/2 results being no better than *ca* ±0·02 Å, calculations on larger molecules may not have seemed justified. Based on the success of our empirical correlation in the case of small unstrained molecules, however, we were encouraged to undertake CNDO/2 calculations on a number of larger molecules of chemical importance. Our prime aim was the calculation of molecular geometries for the benzocycloalkenes. These results were needed as an aid in the interpretation of the NMR data on these molecules.[1] Our calculations for these strained aromatic molecules are given in the following paper,[11] while in the present study we report our results for strained cycloalkenes, *viz* cyclopropene, cyclobutene and cyclopentene, and also for some larger molecules, *viz* naphthalene, azulene and biphenylene. The demonstration of satisfactory agreement between observed and calculated geometries for these test cases was felt to be a necessary step before the calculations on benzocycloalkenes were undertaken.

CALCULATION PROCEDURE

Calculations were performed on a UNIVAC 1108 computer using a FORTRAN deck of CNDO/2 obtained from Dr. G. A. Segal. Our procedure for energy minimization is similar to that discussed by other authors,[10] namely: (1) selection of a set of independent bond lengths and angles which uniquely determine the system; (2) evaluation of the total energy at two points around this origin in turn for each variable; and (3) extrapolation to a new origin followed by repetition of steps (2) and (3). This iterative procedure was repeated until the total energy was constant to *ca* 5×10^{-5} au. In general this means that bond lengths are calculated to within ±0·002 Å and bond angles to ±0·2°.

The first step of this energy minimization procedure,[10] *viz* selection of a set of independent bond length and bond angle parameters, merits some further discussion with respect to cyclic molecules. In the case of the benzocycloalkenes (to be discussed in the accompanying paper), and also naphthalene and biphenylene, suitable parameters characterizing the carbon skeleton of the benzene moiety were a, b, c, θ and ϕ (see molecular fragment Fig 1). These parameters are particularly convenient for use with CNDO/2, since the latter program contains an option in the present version for specifying molecular geometry in terms of bond angles and bond lengths, rather than in Cartesian coordinates.

It is evident that any increase in, for example, bond length a, while b, c, θ and ϕ were kept constant would produce a concomitant increase in d. In the event that energy minimization lay in the direction of increasing a while decreasing d, or *vice versa*, the energy minimization process might well be slow, and therefore wasteful of

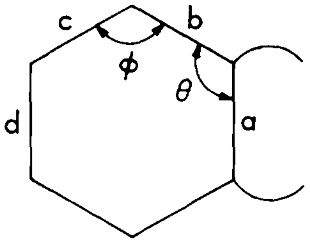


Fig 1.

computational time. A convenient method of accelerating the minimization process which we have used makes use of the much lower energies of bond angle bending processes as opposed to bond length changes. Thus, the angle ϕ instead of being kept totally independent of θ was taken to be,

$$\phi = 360 - 2\theta - w - (\Delta a + \Delta b - \Delta c)/0{\cdot}04 \quad (1)$$

where the angles are in degrees and the terms Δa, Δb, Δc represent small changes (in Å) in bond lengths a, b and c, respectively, relative to the values at the origin, and w is a new variable. The independent parameters are now taken as a, b, c, θ and w. The latter parameter can be used to change ϕ, and hence d independently of the other parameters. The last factor in Eq (1) allows for the trigonometrically derived relationship that a 0·1° increase in ϕ together with either a 0·004 Å decrease in a or b, or else a 0·004 Å increase in c will leave bond length d essentially unchanged. The above procedure gave results essentially identical to the more straightforward method, while greatly saving in computational time.

A second economy in time was effected by assuming standard values for certain of the parameters. For the calculations on the aromatic molecules, C—H bond distances in the aromatic ring were taken equal to those obtained in the CNDO/2 calculation on benzene (1·117 Å). In addition, these C—H bonds were first assumed to bisect the exterior C—C—C angles at the relevant C atoms. Trial calculations showed that neither of these assumptions was severe and that compared to results of a full minimization, errors in C—H distance were less than 0·001 Å and errors in the C—C—H angles were less than a degree.

Finally, the assumptions made as to molecular symmetry need to be discussed. Usually these were the highest consistent with chemical properties, e.g. D_{6h} for benzene and D_{2h} for naphthalene and biphenylene. Relative to this point it is interesting to note the suggestion[12] that in the larger acenes the molecular symmetry may be reduced from D_{2h} to C_{2v} by a distortion of the type shown in Fig 2. This has not been considered in the present work. C_{2v} symmetry was taken for the cycloalkenes and benzocycloalkenes. In the cases of, for example, cyclopentene and indane, it is more likely that the actual symmetry is C_s. However, the calculations are greatly simplified

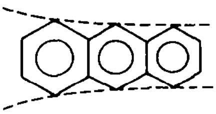


Fig 2.

by using C_{2v} symmetry. Trial calculations on cyclopentene were carried out which showed that CNDO/2 calculates a slightly lower energy for a planar conformation, although this conflicts with experiment.[13] However, the calculated bond lengths were equal for both the planar geometry and one which is nonplanar (with the C_3–C_4–C_5 plane making an angle of 20° with the C_5—C_1–C_2—C_3 plane).

Tables 2 through 5 present the results of our calculations, while the CNDO/2 data in Table 1 is from Ref 10.[*]

DISCUSSION

Careful examination of observed and CNDO/2 calculated carbon–carbon bond lengths for a series of simple hydrocarbons as summarized in Table 1 reveals a number of consistent features. For example, in the case of double bonds, the differences between observed[14, 15] and calculated bond lengths, as shown in Fig 3, are approximately linearly related to the number of hydrogens attached to the C═C system.

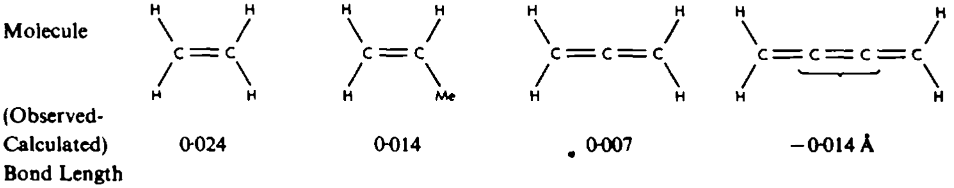


FIG 3.

The difference between observed and calculated C═C bond lengths is therefore decreased by *ca* 0·009 Å for every H-atom that is removed from the C═C skeleton. A similar additive substituent effect was found for single bonds, except that here the difference between observed and calculated was decreased by *ca* 0·015 Å for each H-atom removed from the —C—C— skeleton. For triple bonds, the change was observed to be *ca* 0·006 Å. Recognizing these types of consistent errors, it is possible to correlate more satisfactorily the observed and CNDO/2 calculated carbon–carbon bond lengths (R_o and R_c, respectively) by one simple expression of the form,

$$R_o = R_c + 0{\cdot}001263\, n_H/(1{\cdot}563 - R_c) \tag{2}$$

where n_H is the number of hydrogens attached to the carbon atoms of the carbon–carbon bond. For the molecules of Table 1, Eq (2) reproduces the observed bond lengths to a root mean square (rms) accuracy of ±0·008 Å. This figure is little more than the experimental error, and compares with rms accuracies of ±0·03 Å for unmodified CNDO/2 results, and ±0·02 Å for the reparametrized CNDO/2[10] results on the same molecules as considered by us. The only remaining instances in which the deviations are greater than twice this value are for the central bonds in butadiene (−0·029 Å) and butadiyne (−0·018 Å). These discrepancies remained even after the reparametrization[10] to improve the reliability of calculated bond lengths, and seem to indicate a definite shortcoming in CNDO/2, which is unrelated to substitution patterns.

---

[*] Our calculations for a selection of molecules treated in Ref 10 were in excellent agreement, hence the latter results have been used directly in Table 1.

From equally careful scrutiny of the C—H bond distance results, it is possible to derive a similar relation between the CNDO/2 and experimental bond lengths. For the five typical C—H bonds given in Table 1, the observed and CNDO/2 calculated C—H bond lengths ($\mathbf{R}'_o$ and $\mathbf{R}'_c$, respectively) are related by

$$\mathbf{R}'_o = \mathbf{R}'_c - 0{\cdot}008\,(5 - n_H) \tag{3}$$

where n_H is the total number of protons attached to the C atom. For the five examples cited, Eq (3) reproduces the bond lengths to an rms error of ±0·003 Å as compared to an error of ±0·025 Å for CNDO/2 itself, and a value of ±0·018 Å after reparametrization.[10]

Based on the satisfactory simple correlations discussed above, we infer that CNDO/2 contains features which give rise to errors in carbon–carbon and C—H bond distances which are dependent on two factors, *viz*, the type of bond, e.g., single, double or triple, and also upon the degree of substitution. It may be that the pattern of errors arises from the assignment of relatively wrong electron attracting powers to carbon and hydrogen in the program.[10] On a qualitative basis, we might then expect that for a particular type of C—C bond, the amount of charge placed on the C atoms would be dependent on the number of attached hydrogens. In the energy minimization process, it is reasonable that the C—C distance found is sensitive to the amounts of charge on the C-atoms, and therefore to the extent of H-substitution. Furthermore, for a given substituent pattern, it might be expected that the relatively small error wrought in the C—C distance by the slightly erroneous charges placed on carbon would depend on the C—C distance. Because the energy required to distort a bond length is greater for multiple bonds than for single bonds, it is logical to expect that the errors introduced into C—C distances by substitution of equal numbers of hydrogens would be in the order C—C > C═C > C≡C; this is as observed.

It may be that Eqs (2) and (3) and their rationale as discussed here will serve to indicate ways in which CNDO/2 could be improved to give a better account of bond lengths. For the present, we shall use Eq (2) in a purely empirical manner for the more straightforward comparison of observed and calculated C—C bond lengths. It should be noted at the outset that the consistent errors in bond length made by CNDO/2 will necessarily lead to somewhat inaccurate values of bond angles, particularly in cyclic molecules.

Cycloalkenes

Our results for cyclopropene, cyclobutene, and cyclopentene are given in Tables 2 and 3. The CNDO/2 treatment correctly predicts that the C═C bond in cyclopropene is shorter than that in either cyclobutene or cyclopentene, with the latter two being almost equal. The C—C bonds adjacent to the double bond are calculated to have nearly the same length in all the molecules, as found experimentally,[13, 16, 17] while the unique C—C bond in cyclobutene is correctly calculated to be very long. Apart from these general considerations, it is seen that Eq (2) gives bond lengths correct to *ca* ±0·01 Å except for the C═C bond in cyclopropene, which is in error by −0·025 Å. The strain in this molecule is severe; so the poor value from CNDO/2 is not altogether surprising. For these three molecules, Eq (3) reproduces the C—H bond lengths to an rms error of ±0·006 Å, with almost all this error coming from the methylene hydrogens in cyclopropene. It is furthermore interesting to note the C—C—H angle dependences

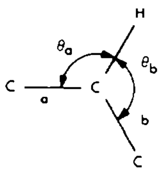


FIG 4.

found by CNDO/2 for the environment of Fig 4. Where bond *a* is longer than bond *b*, CNDO/2 calculates that the C—H bond (or H—C—H plane for the methylene groups) lies in such a position as to make $\theta_b > \theta_a$. This seems to be in accord with the limited experimental data at hand, e.g. that for the methine hydrogens in cyclopropene and cyclobutene. In the case of molecules where bond angles are close to 120°, the total energy is only slightly sensitive to the C—C—H bond angles, provided that θ_a and θ_b are equal to within one or two degrees.

Polycyclic conjugated molecules

For molecules the size of naphthalene or larger, the energy minimization procedure becomes very costly in computer time because of the large number of orbitals, and also the large number of independent bond angle and bond length parameters involved. Therefore, calculations were performed only in the cases of naphthalene, biphenylene and azulene. The results are given in Tables 4 and 5. With the exception of the 5,5′ bond in biphenylene and the 9,10 bond in azulene, the observed bond lengths and those from Eq (2) agree to better than $\pm 0{\cdot}006$ Å. The CNDO/2 treatment appears to fail badly in the case of the long bonds between sp^2 carbons in biphenylene and azulene. As mentioned above, similar results were found for the long bonds between unsaturated C atoms in butadiene and butadiyne (Table 1).

It must be noted that Hückel π-electron theory and also its refinements[18, 19] are capable of calculating C—C bond lengths in conjugated hydrocarbons to an accuracy comparable with the present work. However, these former methods proceed by first calculating π-bond orders and then using an empirical correlation between π-bond orders and carbon–carbon distances, and so are not applicable to saturated systems. Furthermore, these π-electron methods generally do not pay attention to bond angle changes. In the present CNDO/2 work, all valence electrons have been included, and the iterative energy minimization procedure has been carried out over all the bond length and bond angle parameters to yield an optimum theoretical geometry.

CONCLUSION

The present CNDO/2 calculations on cycloalkenes and three polycyclic conjugated molecules confirm that Eq (2) is capable of correlating observed and calculated C—C bond lengths to an accuracy of *ca* $\pm 0{\cdot}008$ Å.* This good agreement encouraged us to undertake the calculations on benzocycloalkenes as described in the following paper.[11]

---

* After this work was completed, a paper appeared giving bond length results from Dewar's MINDO method.[20] It seems from the data presented there for small molecules that MINDO can successfully reproduce C—C bond lengths, including even those for long bonds between sp^2 carbons.

TABLE 1. EXPERIMENTAL AND CALCULATED BOND LENGTHS (Å) FOR SMALL HYDROCARBONS

Bond	Compound	Experimental[a]	CNDO/2[b]	Eq (2)
C≡C	Acetylene	1·205	1·197	1·204
	Propyne	1·206	1·207	1·211
	Butenyne	1·215[c]	1·208	1·212
	Butadiyne	1·205[d]	1·208	1·212
C=C	Ethylene	1·335	1·311	1·331
	Propene	1·336[e]	1·322	1·338
	Butadiene	1·342[d]	1·324	1·340
	Butenyne	1·344[c]	1·322	1·338
	Allene	1·312	1·305	1·315
	Butatriene 1–2	1·309	1·305	1·315
	Butatriene 2–3	1·284	1·298	1·298
$C{=}C_{arom}$	Benzene	1·397	1·384	1·398
C—C	Ethane	1·536	1·459	1·532
	Propane	1·526	1·468	1·534
	Cyclopropane	1·524	1·474	1·531
	Propene	1·501[e]	1·455	1·502
	Propyne	1·459	1·430	1·458
	Butadiene	1·463[d]	1·466	1·492
	Butenyne	1·434[c]	1·421	1·430
	Butadiyne	1·384	1·402	1·402
				Eq(3)
C—H	Methane	1·106	1·114	1·106
	Ethane	1·107	1·120	1·104
	Ethylene	1·084	1·112	1·088
	Acetylene	1·059	1·093	1·061
	Benzene	1·084	1·117	1·085

[a] Experimental data from Ref 14 except where stated.

[b] CNDO/2 results from Ref 10.

[c] T. Fukuyama, K. Kuchitsu and Y. Morino, *Bull. Chem. Soc. Japan* **42**, 379 (1969).

[d] K. Kuchitsu, T. Fukuyama and Y. Morino, *J. Mol. Struct.* **1**, 463 (1968).

[e] Ref 15.

Table 2. Experimental (in parentheses) and CNDO/2 calculated geometries for lower cycloalkenes.[a,b,c,d] (For references see footnotes to table 3)

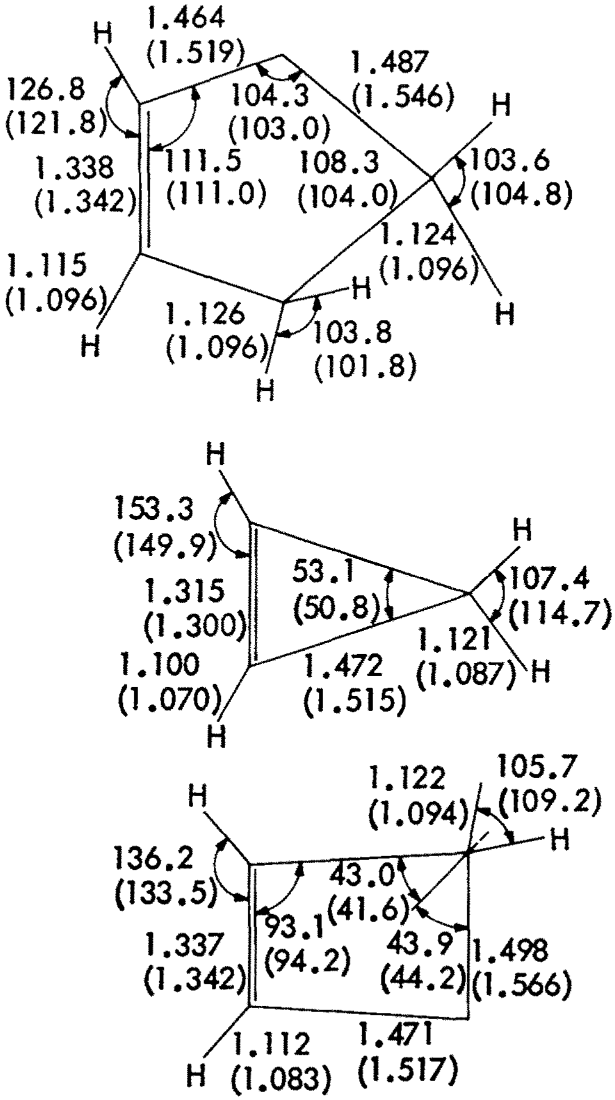
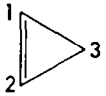
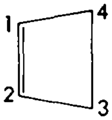
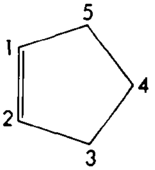


TABLE 3. C—C BOND LENGTHS FOR CYCLOALKENES: OBSERVED, CNDO/2, AND VALUES FROM EQ (2) COMPARED

Compound	Bond	Experimental[a]	CNDO/2[a]	Eq (2)
	1,2	1·300[b]	1·315	1·325
	1,3	1·515	1·472	1·514
	1,2	1·342[c]	1·337	1·348
	1,4	1·517	1·471	1·512
	3,4	1·566	1·498	1·576
	1,2	1·342[d]	1·338	1·349
	1,5	1·518	1·464	1·502
	3,4	1·546	1·487	1·553

Footnotes to Tables 2 *and* 3

[a] Bond lengths in Å; bond angles in degrees. All C—H and C—C—H angles were included in the energy minimization.

[b] Cyclopropene: Ref 16.

[c] Cyclobutene: Ref 17. Dotted line in Table 2 shows plane of CH_2 group.

[d] Cyclopentene: Ref 13. Planes of CH_2 groups assumed to bisect exterior C—C—C angles of planar carbon system. Note that experimental and calculated bond angles cannot be well compared because of the planar geometry used in CNDO/2. Also, note that ref. 13 assumed all C—H bond lengths equal.

TABLE 4. EXPERIMENTAL (IN PARENTHESES) AND CNDO/2 CALCULATED GEOMETRIES FOR POLYCYCLIC MOLECULES: BOND LENGTHS IN Å AND BOND ANGLES IN DEGREES. (FOR REFERENCES SEE FOOTNOTES TABLE 5)

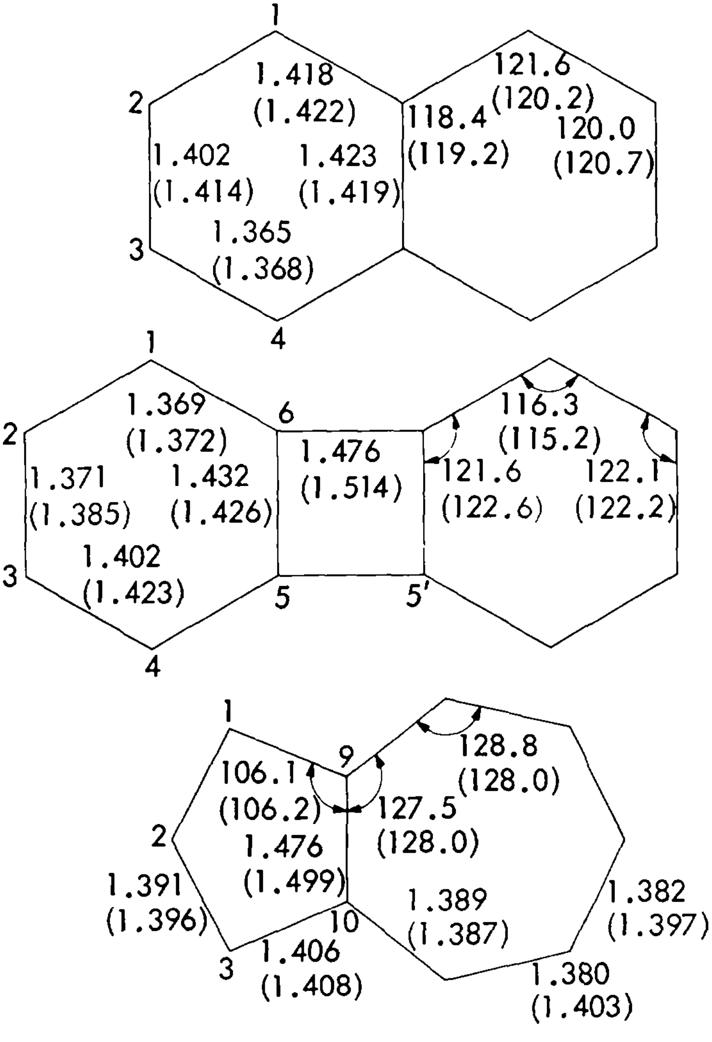
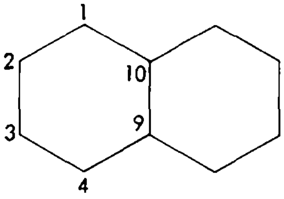
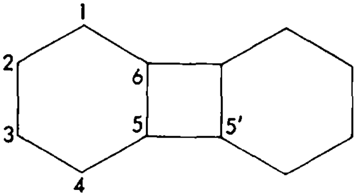
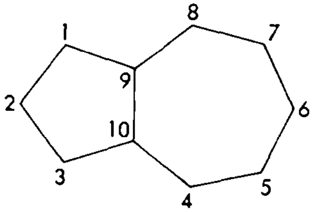


TABLE 5. C—C BOND LENGTHS (Å) FOR POLYCYCLIC MOLECULES:[a] OBSERVED VALUES AND THOSE CALCULATED BY CNDO/2 AND EQ (2)

Compound	Bond	Experimental	CNDO/2	Eq (2)
	1,2	1·368[b]	1·365	1·378
	2,3	1·414	1·402	1·418
	1,10	1·422	1·418	1·427
	9,10	1·419	1·423	1·423
	1,2	1·423[c]	1·402	1·418
	2,3	1·385	1·371	1·384
	1,6	1·372	1·369	1·376
	5,6	1·426	1·432	1·432
	5,5′	1·514	1·476	1·476
	1,2	1·396[d]	1·391	1·405
	1,9	1·408	1·406	1·413
	4,5	1·403	1·380	1·394
	4,10	1·387	1·389	1·396
	5,6	1·397	1·382	1·396
	9,10	1·499	1·476	1·476

[a] C—H bonds were assigned the length 1·117 Å, as in benzene, and were taken to bisect the exterior C—C—C angles (see text).

[b] D. W. J. Cruickshank and R. A. Sparks, *Proc. Roy. Soc.* **A258**, 270 (1960).

[c] J. K. Fawcett and J. Trotter, *Acta. Cryst.* **20**, 87 (1966).

[d] Average of X-ray crystallographic results [A. W. Hanson, *Acta. Cryst.* **19**, 19 (1965)] and electron diffraction results [O. Bastiensen and J. L. Dennison, *Acta. Chem. Scand.* **20**, 1319 (1966)].

Acknowledgements— We should like to thank Professor G. A. Segal (University of Southern California) for supplying the CNDO/2 deck, and both him and Dr. R. L. Poynter (JPL) for helpful discussion.

REFERENCES

[1] M. A. Cooper and S. L. Manatt, *J. Am. Chem. Soc.* **92**, 1605 (1970)
[2] M. A. Cooper and S. L. Manatt, *Ibid.* **91**, 6325 (1969)
[3] M. A. Cooper, D. D. Elleman, C. D. Pearce and S. L. Manatt, *J. Chem. Phys.* **53**, 2343 (1970)
[4] M. A. Cooper and S. L. Manatt, *J. Am. Chem. Soc.* **92**, 4646 (1970)
[5] M. A. Cooper and S. L. Manatt, *Org. Magn. Resonance* **2**, in press (1970)
[6] J. A. Pople, D. P. Santry and G. A. Segal, *J. Chem. Phys.* **43**, *S* 129 (1965)
[7] J. A. Pople and G. A. Segal, *Ibid.* **43**, *S*, 136 (1965); *Ibid.*, **44**, 3280 (1966)
[8] H. H. Jaffe, *Accounts Chem. Res.* **2**, 136 (1969)
[9] K. B. Wiberg, *J. Am. Chem. Soc.* **90**, 59 (1968)
[10] H. Fischer and H. Kollmar, *Theoret. Chim. Acta.* **13**, 213 (1969)
[11] C. C. Cheung, M. A. Cooper and S. L. Manatt, *Tetrahedron* **27**, 701 (1971)
[12] G. Binsch, E. Heilbronner and J. N. Murrell, *Mol. Phys.* **1**, 305 (1966)
[13] G. W. Rathjens, *J. Chem. Phys.* **36**, 2401 (1962); M. I. Davis and T. W. Muecke, *Ibid.* **74**, 1104 (1970)
[14] *Interatomic Distances.* Chemical Society Special Publication No. 18, London (1965)

[15] D. R. Lide and D. E. Mann, *J. Chem. Phys.* **27**, 868 (1967)

[16] R. H. Kasai, R. J. Meyers, D. F. Eggers and K. B. Wiberg, *Ibid.* **30**, 512 (1959)

[17] B. Bak, J. J. Led, L. Nygaard, J. Rastrup-Andersen and G. O. Sorensen, *J. Mol. Structure* **3**, 69 (1969)

[18] e.g., L. Salem, *Molecular Orbital Theory of Conjugated Systems*, pp. 134–143. Benjamin, New York, N.Y. (1966)

[19] C. A. Coulson and A. Golebiewski, *Proc. Phys. Soc.* **78**, 1310 (1961)

[20] M. J. S. Dewar and E. Haselbach, *J. Am. Chem. Soc.* **92**, 590 (1970)