CBPF-NF-001/81

AVE BOND INDEX IN THE H-BOND OF THE WATSON-CRICK PAIRS

by

Mario Giambiagi, Myriam Giambiagi and
Wilton Barroso Filho

Centro Brasileiro de Pesquisas Físicas - CBPF/CNPq
Av. Wenceslau Braz, 71 - fundos
22290 - Rio de Janeiro - R.J. - Brasil

ABSTRACT

The normal Watson-Crick base pairs are treated as super-molecules. The properties of the electronic distribution along the N-H...Y bonds is studied in an all-valence-electrons calculation, through a bond index formula devised for non-orthogonal basis. Electronic density diagrams of the adenine-uracil base pair are analysed.

1. INTRODUCTION

A typical difficulty encountered in the theoretical study of the nucleic acids hydrogen bond properties is the fact that usually the bases are calculated separately, except for π calculations. To our knowledge, there are neither all-valence-electrons (AVE) nor ab-initio calculations of the base pairs as supermolecules. In a recent paper [1], apparently, they were considered as a single entity; however, this was done only under an indirect form.

Another point is that, as bond orders are non-observable and they do not have an unambiguous physical meaning, perhaps they are not wholly suitable for the interpretation of certain electronic properties; even more in a bond so peculiar as the hydrogen bond.

In this note, we intend to discuss the hydrogen bond considering the basis pairs as supermolecules, and to apply a bond index formula which has been devised for non-orthogonal basis [2]. We also show electronic density contour maps for a base pair; these diagrams are appropriate when considering certain properties of biological systems [3].

Limitiations of the supermolecule approach have been pointed out in semiempirical calculations [4]. Nevertheless, they are beside the point for the magnitudes we are interested in.

The merits and shortcomings for the EHT method, particularly when dealing with biological molecules, are well known [5], and we shall not return on them. After some years of disuse, it has deserved a recent revival [6], and has proven to be useful for the determination of excimer-state characteristics of pyrimidine bases homodimers [7]. In order to overcome some of the EHT limitations we employ the IEHT treatment [8]. We think that this approximation is satisfactory enough for our purposes and, last but not least, it is considerably cheaper than more sophisticated methods.

The application of the mentioned bond index formula [2] to the hydrogen bond of the nucleic acids may enlighten its physical sense; this may turn out particularly helpful in this case where it is difficult to obtain thermodynamic parameters and few reliable experimental data are available [9].

2. BOND INDEX FORMULA

If x_{ik_μ} is the LCAO coefficient of the k_μ-th atomic orbital belonging to atom μ in the i-th MO, S the overlap matrix and

$$y_{ir_\nu} = \sum_{k_\mu} S_{k_\mu r_\nu} x_{ik_\mu} \qquad (1)$$

the bond index $I_{\mu\nu}$ between atoms μ and ν is defined as [2]:

$$I_{\mu\nu} = \sum_{i,j} \sum_{k_\mu, r_\nu} n_i x_{ik_\mu} y_{ir_\nu} n_j x_{jr_\nu} y_{jk_\mu} \qquad (2)$$

where n_i is the occupation number of the i-th level, and the summation in k_μ and r_ν is carried on every atomic orbital belonging to atom μ or atom ν.

The atomic charge q_μ may be written as [2]:

$$q_\mu = \left[N/(N + 2N_d)\right] I_{\mu\mu} + \left[N/(N + 2N_d)\right] \sum_{\nu \neq \mu} I_{\mu\nu} \qquad (3)$$

where N is the number of electrons, and N_d the number of MO whose n_i=2. The first term of (3) may be interpreted as the self-charge of the atom in the molecule, and the second one as the active charge [10] distributed along both effective and formal bonds of atom μ with all the other atoms.

Physically, $I_{\mu\nu}$ resembles Mulliken's population analysis [11], but $I_{\mu\nu}$ has the advantage of not cancelling for $\mu \neq \nu$ if S is the unit matrix; if besides $2N_d$=N, it reduces instead to the more familiar Wiberg bond index [12]. Formula (2), which is rotationally invariant, has been applied successfully to hydrocarbons and some hetero-molecules [2].

3. BOND INDICES RESULTS AND DISCUSSION

We focus here our attention on $I_{\mu\nu}$ when μ and ν are atoms involved in the hydrogen bond of the Watson-Crick base pairs (see Table I).

In the first place, let us underline that the significance of the values in the tables is given by the fact that every other intermolecular $I_{\mu\nu}$ has at least two orders of magnitude less than those reported. This effect cannot be ascribed to the associated distances. For example, in the G-C base pair, C_2 of guanine is 4.33 $\mathring{A}$ away both from C_8 of the imidazole ring and from the cytosine carbon of the carbonyl group. Now, in the first case $I_{\mu\nu}$=0.0147, and in the second one $I_{\mu\nu}$=0.0000. As in the supermolecule the formula does not discriminate between inter-or-intramolecular $I_{\mu\nu}$, these values give a clear idea of the physical meaning of formula (2), being in agreement with chemical intuition.

The constancy of the NH...N value ($\sim$0.054) is striking; it holds also for the A^*-U^* pair and for other pairs not shown in this paper [13].

The O...HN index is nearly half the former one in A-T and A-U. Its value is enhanced in the presence of three H-bonds, as in G-C. It is known that the experimental enthalpies of hydrogen bonding are much higher in the G-C dimer (10-11.5 kcal/mol) than in the A-U dimer (6.2 kcal/mol) [14], and this is usually credited to the existence of three H-bonds in the first case and two in the second one. It may then be also due to the reinforcement of the O...HN bonds. Tautomerization appears to have too such an effect, and this certainly deserves further study [13]. Anyway,

from the point of view of the electronic distribution along the bond, the NH...N bond seems the strongest of the H-bonds of the normal Watson-Crick base pairs. This inference, which now has a more definite physical meaning, has been mentioned once in connection with π bond orders for the base pairs [15].

The I_{NY} (in NH...Y) values are wholly σ, the I^{π}_{NY} being zero up to four decimal places. Therefore, one should handle with some caution the widely accepted assumption of π delocalization between the basis forming a pair. A dynamical model with this underlying hypothesis has been recently proposed, e.g., for the description of DNA as a superconductant [16]. Our results predict instead that conjugation along the bridge is of σ nature.

The above considerations about the NH...Y bond index arise actually from the values of $I_{\mu\nu}$ for H...Y. For example, $I_{\mu\nu}$ for N...H (in N...HN) is ∿0.077. The O...H value in O...HN is less, about a half in the A-U and A-T pairs. That is nitrogen and hydrogen share about 0.08 of an electron in H-bond, while oxygen and hydrogen share appreciably less. Experimental NMR spectra of 9-ethyladenine, 1-cyclohexylthymine and 1-cyclohexyluracil (associated and not associated) indicate a stronger association in the position corresponding to the NH...N bond than in the position corresponding to the O...HN bond. In agreement with our results, also, the proton chemical shifts for A-U and A-T derivatives are equal [17] in the hydrogen-bonded positions.

One could wonder where the electronic distribution forming the bridge comes from. In order to answer this question the I_{NH} value of Table I should be compared with those of the separate bases in Table II. It is thus seen that the fraction of electron loosed by the N-H bond when forming a hydrogen bond goes *virtually all* to the bridge. The usual picture of an XH...Y hydrogen bond is made in terms of proton-donating character of X-H groups or electron-accepting tendency from Y. Now, our results lead to a model where the N-H group transfers an electronic fraction to H...Y (see Section 4). Both descriptions agree with a consequent weakening of the X-H bond.

We report in Tables III and IV the charge partition (in self and active charge) only for the hydrogen atoms involved in these bonds. Table III shows the values for the pairs, and Table IV those for the separate bases. It is seen that their behaviour is very nearly constant, not depending of the non-adjacent principal atom being N or O. We may say that the total charges are constant within 2%, the self charges within 4% and the active charges within 1%. Let us remark that the hydrogen "keeps" less charge than it "gives up" to the molecule.

In Table III we have reported also, only for the sake of comparison, the π bond orders P_{NN} and P_{NO} of the hydrogen bonds, together with those of ref.[15], but perhaps they do not have much significance.

4. ELECTRON-DENSITY DIAGRAMS

Drawing an electron density diagram helps us to illustrate what has been discussed in the preceding section. In Fig. 2 we show the hydrogen bonding region of the A-U pair. As we have seen that the π conjugation accross the H bond may be safely neglected, we have drawn the diagram on the molecular plane. In experimental three-dimensional Fourier electron density functions close to the molecular plane determined by X-ray diffraction methods of ref. [18] hydrogens, and hence the hydrogen bonds, are not clearly resolved. A further difference Fourier synthesis finally reveals the position of all hydrogen atoms [19]; but here the contours which could describe hydrogen bonds are not shown.

The σ conjugation curves are in agreement with the $I_{\mu\nu}$ values of Table I. Thus the contour of 0.025 e/au^3 conjugates in the lower bond (I_{NN}=0.054), while it does not in the upper one. For this one, the conjugating contour line is 0.015 (I_{ON}=0.027).

Around the oxygen atom taking part in the bond, the closed curve of 1.0 clearly depicts a "horseshoe" distribution which issues from a combination of the two sp^2 electronic orbitals corresponding to electron pairs. The perturbation brought about by the hydrogen bond appears to be very small, in accordance with the model arising from the $I_{O...H}$ and I_{NH} values discussed in the precedind section. The hybridization of this oxygen is $sp^{2.03}$, almost rigorously trigonal, while the other oxygen has more p character ($sp^{2.24}$).

In a classical paper [20], it has been stated that in the G-C base pair the HOMO and the LEMO originate from different single bases. The consequence of this is the prediction of a charge transfer complex in the transition between both molecules. We shall discuss elsewhere [13] our detailed results, but it seems worthwhile to spend some words about the HOMO and LEMO of the A-U base pair. Inspection of the molecular orbitals shows unambiguously that the LEMO comes from adenine, has π character, and is practically unaltered by the dimer formation.

As to the HOMO, the molecular orbitals (all of them σ) indicate that it arises from uracil, undergoing a certain modification, with a slight contribution from adenine. Fig. 3 depicts this clearly. It is seen that adenine has little to do with the dimer's HOMO, and also the modification suffered by uracil's HOMO. The electronic density in U and A-U tends to concentrate around uracil's oxygen taking part of the hydrogen bond. All the orbitals of uracil's HOMO change their orientation upon entering the dimer. In the case of the $2p_x$ orbital of the upper oxygen, this favours the formation of the hydrogen bond, where appears the only contribution coming from adenine, through the amino nitrogen. Curiously, therefore, the A-U base pair HOMO suggests the O...HN bond, and not the HN...N one.

5. CONCLUSIONS

1. The conjugation along the hydrogen bond between the bases forming a pair is of σ nature.

2. The fraction of electron loosed by the NH bond in the separate bases goes virtually all to the bridge when forming a pair.

3. The hydrogen bond in the nucleic acids is formed by a fraction of 0.08 of an electron in H...N while for H...O it goes from 0.03 up to 0.05 of an electron.

4. The bond index $I_{\mu\nu}$ seems a suitable quantity for describing hydrogen bonds properties.

The authors wish to express their warmest acknowledgement to Dr. Jorge Torres for the long discussions around these arguments.

REFERENCES

[1] M.Geller, A.Jaworski and A.Pohorille, Int. J.Quantum Chem. 15 (1979) 369.

[2] M.Giambiagi, M.S. de Giambiagi, D.R.Grempel and C.D. Heymann, J. Chim. Phys. 72 (1975) 15.

[3] D.M.Esquivel, M.S. de Giambiagi and M.Giambiagi, Chem. Phys. Lett. 64 (1979) 131; D.B.Boyd, J. Am. Chem. Soc. 94 (1972) 64.

[4] G.R.Pack, G.H.Loew, S.Yamabe and K.Morokuma, Int. J. Quantum Chem., Quantum Biol. Symp. 5 (1978) 417.

[5] See for example: A.Pullman, Int. J. Quantum Chem. 2s (1968) 187; E.Fraga, S.Fraga, Biomoléculas (Ed. Alhambra, Madrid, 1976).

[6] A.Gavezzoti and M.Simonetta, Chem. Phys. Lett. 48 (1977) 434; M.Barzaghi, C.Oliva and M.Simonetta, J.Phys. Chem. 84 (1980) 1959; L.Eisenstein, D.R.Franceschetti and K.L.Yip, Theoret. Chim. Acta 49 (1978) 349.

[7] V.I.Danilov, V.I.Pechenaya and M.R.Sharafutdinov, Chem. Phys. Lett. 59 (1978) 545.

[8] QCPE program Nº 256, to which we have added the subroutine for bond indexes.

[9] J.N.Spencer et al., J. Phys. Chem. 83 (1979) 2615.

[10] C.A.Trindle, J. Am. Chem. Soc. 91 (1969) 219.

[11] R.S.Mulliken, J. Chem. Phys. 23 (1955) 1833.

[12] K.B.Wiberg, Tetrahedron 24 (1968) 1083.

[13] M.S. de Giambiagi, M.Giambiagi and D.M.Esquivel, to be published.

[14] Y.Kyogoku, R.C.Lord and A.Rich, Biochim. Biophys. Acta 179 (1969) 10.

[15] A.Imamura, H.Fujita and C.Nagata, Bull. Chem. Soc. Japan 40 (1967) 522.

[16] D.A.Paine and W.L.Pensinger, Int. J. Quantum Chem. 15 (1979) 333.

[17] I.Morishima, T.Inubushi, T.Yonezawa and Y.Kyogoku, J. Am. Chem. Soc. 99 (1977) 4299.

[18] H.M.Sobell, J. Mol. Biol. 18 (1966) 1.

[19] F.Mazza, H.M.Sobell and G.Kartha, J. Mol. Biol. 43 (1969) 407.

[20] R.Rein and J.Ladik, J. Chem. Phys. 40 (1969) 2466.

FIGURE CAPTIONS

Figure 1 — Watson-Crick pairs.

Figure 2 — Contour density diagram of the H-bond region in A-U, in the molecular plane. Units are e/au^3.

Figure 3 — a) HOMO of the A-U bases pair; b) HOMO's of U and A, both in the molecular plane. The contours are respectively, from outside, in e/au^3: 0.001, 0.005, 0.01 and 0.1.

TABLE I

$I_{\mu\nu}$ for μ,ν atoms involved in the hydrogen bonds. In the case of the G-C base pair, where there are two bonds of the same type, the first value corresponds to the O...HN bond above, and the second value to the bond below (see Fig. 1).

	A-U	A-T	G-C
I(ON) in O...HN	0.027	0.025	0.039 0.032
I(NN) in NH...N	0.054	0.054	0.053
I(O...H) in O...HN	0.029	0.032	0.051 0.042
I(H...N) in NH...N	0.076	0.077	0.077
I(NH) in O...HN	0.942	0.942	0.923 0.932
I(NH) in NH...N	0.899	0.895	0.889

TABLE II

$I_{\mu\nu}$ for the group N-H involved in the hydrogen bonds, for the bases taken separately, and for the pairs of Table I. In the case of the G-C base pair, as in Table I, G_1 corresponds to the bond above, and G_2 to the bond below.

	I(N-H) (sep.bases)	I(N-H) (pair)	I(sep.bases) - I(pair)	I(H...Y) (pair)
A	0.977	0.942	0.035	0.029
U	0.974	0.899	0.075	0.076
T	0.974	0.895	0.079	0.077
G_1	0.968	0.889	0.079	0.077
G_2	0.974	0.932	0.042	0.042
C	0.980	0.923	0.057	0.051

TABLE III

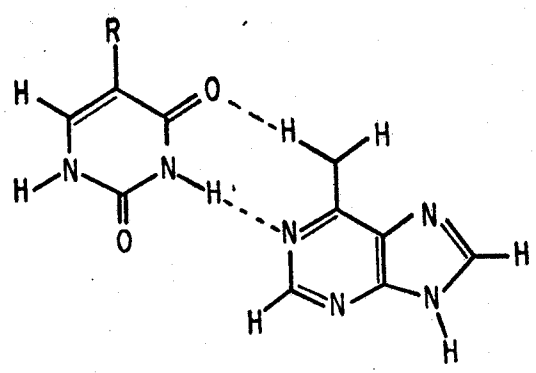
Partition of charge in the hydrogen atoms involved in the H bond NH...Y, and π bond order between N and Y in the pairs of table I. In the last column, between brackets, values from ref. [15].

Pair	Bond	Total charge	Self-charge	Active charge	$P^{\pi}_{\mu\nu}$
A-T	O...HN	0.875	0.383	0.492	0.0002(0.0001)
	NH...N	0.871	0.379	0.491	0.0039(0.0025)
A-U	O...HN	0.873	0.382	0.492	0.0004
	NH...N	0.871	0.380	0.496	0.0039
G-C	NH...O	0.876	0.384	0.492	0.0000(0.0001)
	N...HN	0.863	0.372	0.490	0.0000(0.0023)
	O...HN	0.871	0.380	0.491	0.0000(0.0003)

TABLE IV

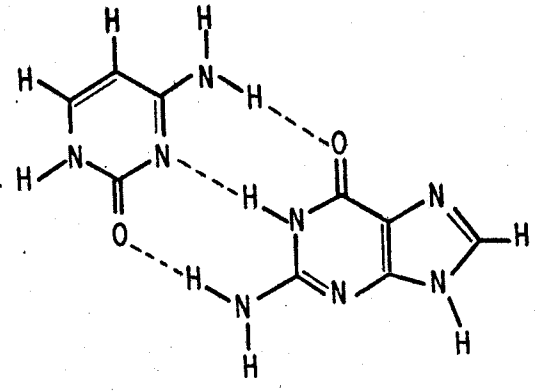
Partition of charge in the hydrogen atoms involved in the H bond, in the separate bases. G_1 and G_2 have the same significance as in Table III.

	Total charge	Self-charge	Active Charge
U	0.873	0.381	0.492
T	0.869	0.377	0.492
C	0.879	0.385	0.494
A	0.876	0.383	0.493
G_1	0.859	0.369	0.490
G_2	0.867	0.376	0.491



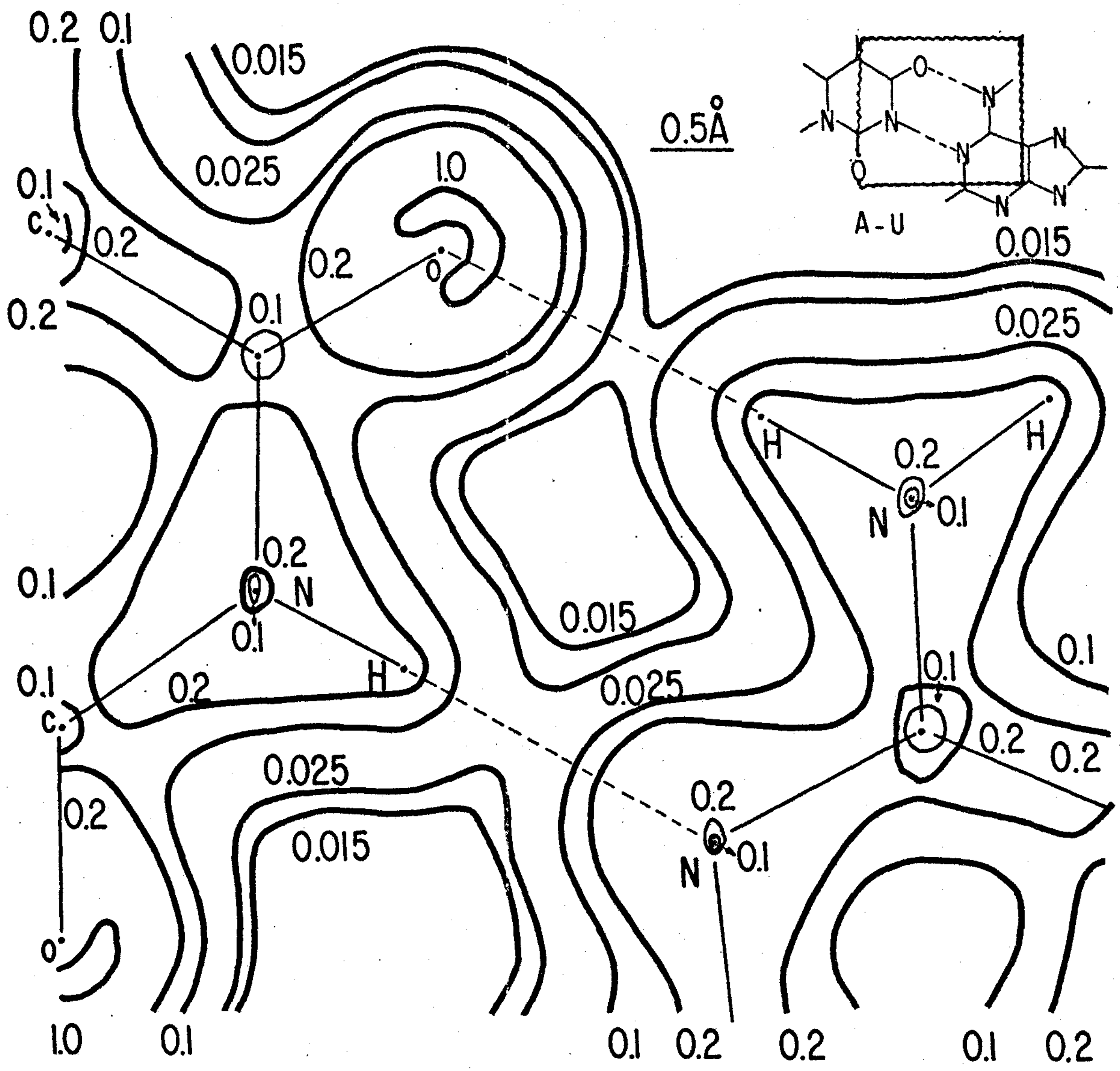
A — U (R = H)

A — T (R = CH_3)



G — C

FIG.1



0.2
0.1
0.015
0.5Å
A - U
0.025
1.0
0.1
c.
0.2
0.2
O
0.015
0.2
0.1
0.025
H
0.2
H
N
0.1
N
0.2
0.1
0.1
N
0.1
0.015
0.1
H
0.1
0.2
0.025
0.1
c.
0.2
0.2
0.025
0.2
0.2
0.015
N
0.1
O
1.0
0.1
0.1
0.2
0.2
0.1
0.2

FIG.2

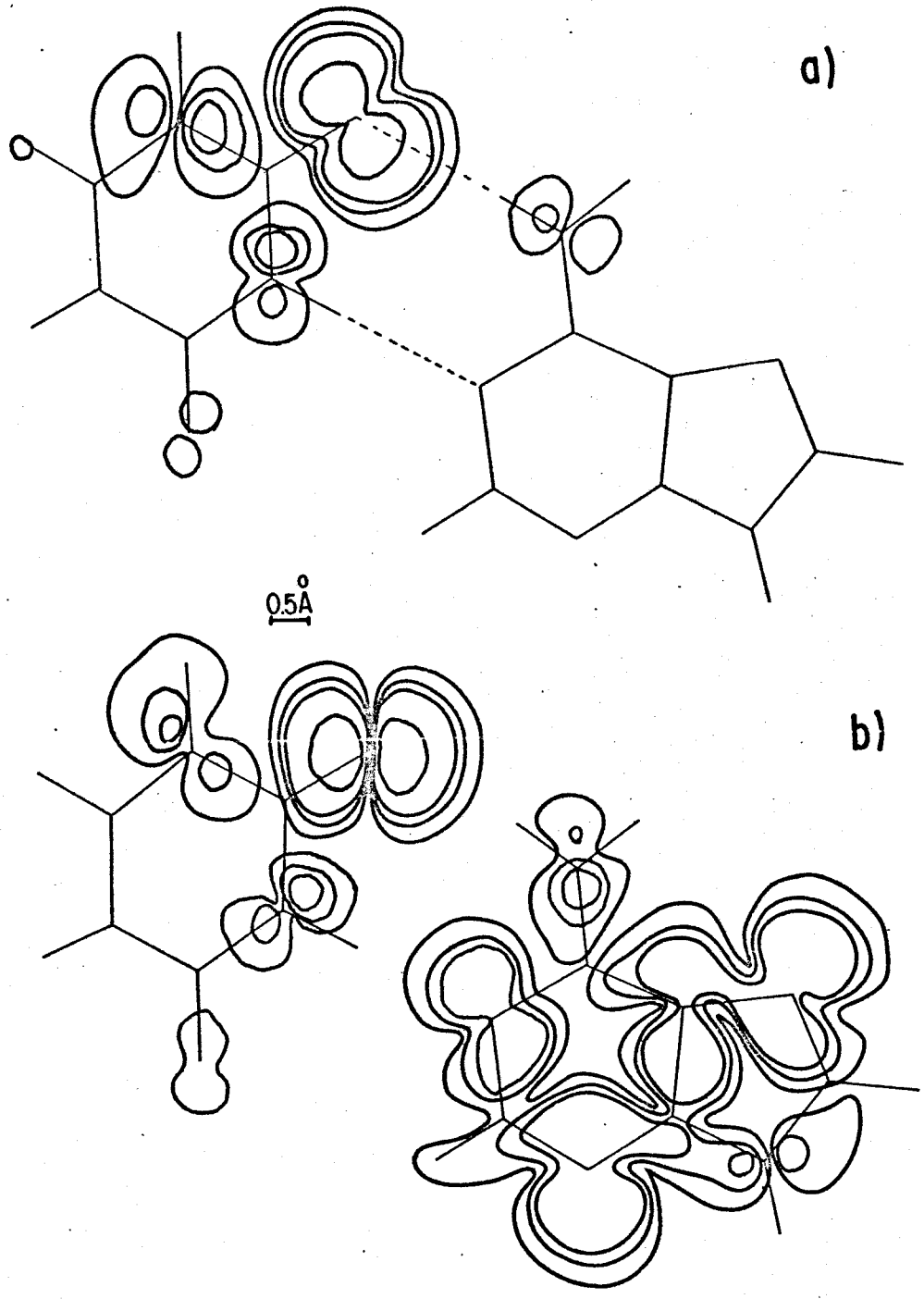


FIG. 3