

ISSN 0029-3865

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO - CNPq

CENTRO BRASILEIRO DE PESQUISAS FÍSICAS - CBPF

Coordenação de Documentação e Informação Científica - CDI

Área de Publicações

CBPF-NF-063/83

HYDROGEN BONDING AND CHARGE TRANSFER IN SOME  
INTERACTIONS INVOLVING METHYL GLYOXAL AND  
FORMAMIDE

by

M. Segre de Giambiagi, M. Giambiagi and  
C. Alarcon Chamarelli

Rio de Janeiro

CBPF

1983

#### ABSTRACT

The possibility of hydrogen bonding between the nucleic acid bases and methyl glyoxal or formamide is studied through an IEH calculation; the charge transfer and hydrogen bond indices support this hypothesis. The CH...Y indices are similar to those of usual hydrogen bonds. In six of the nine configurations previously proposed by other authors in order to analyze the H<sub>2</sub>S-formamide interactions, formamide is predicted to be an electron acceptor, as expected.

Key-words: H-bonding interactions; Charge transfer; Bond index.

HYDROGEN BONDING AND CHARGE TRANSFER IN SOME INTERACTIONS  
INVOLVING METHYL GLYOXAL AND FORMAMIDE

M. Segre de Giambiagi, M. Giambiagi and C. Alarcon Chamarelli

Centro Brasileiro de Pesquisas Físicas

R. Xavier Sigaud 150, 22290 Rio de Janeiro - RJ, Brasil

1. - Introduction

Reactions involving glyoxal (and its derivatives) and nucleic acids have been extensively studied (<sup>1</sup>). Besides, it has been suggested that these compounds could interact through hydrogen bonds with the nucleic acid bases; the pairs which would appear have been calculated using the CNDO/2 method, focusing on the presumed electron-acceptor properties of glyoxals with respect to proteins (<sup>2</sup>). It has been also mentioned that formamide can hydrogen bond to the nucleic acid bases in a similar manner (<sup>2</sup>).

We wondered if some additional information could be supplied by a bond index giving a quantitative estimate of the electronic cloud along the XH...Y bond (<sup>3</sup>). This bond index, originally devised as a generalization of the Wiberg index to non-orthogonal basis (<sup>4</sup>), has been successfully applied to H-bonding of the nucleic acid bases appearing in DNA (<sup>3</sup>) or tRNA (<sup>5</sup>).

Szent-Györgyi advanced the hypothesis that sulfur-containing compounds could act as donors with respect to proteins (<sup>6</sup>). As a first crude model for this interaction, H<sub>2</sub>S-formamide has been suggested (<sup>7</sup>); of the nine configurations calculated with ab initio methods, only one agrees with the referred supposition. We have attempted an IEH approach (<sup>8</sup>) to the question.

## 2. - Bond index definition

The bond index  $I_{\mu\nu}$  between atoms  $\mu$  and  $\nu$  for the ground state and the electronic density  $q_{\mu}$  are defined as <sup>(4)</sup>

$$I_{\mu\nu} = 4 \sum_{i,j} \sum_{\substack{k \\ \mu}} \sum_{\substack{r \\ \nu}} x_{ik_{\mu}} y_{ir_{\nu}} x_{jr_{\nu}} y_{jk_{\mu}} ; q_{\mu} = \frac{1}{2} \sum_{\nu} I_{\mu\nu} \quad (1)$$

where

$$y_{ir_{\nu}} = \sum_{\substack{k \\ \mu}} S_{k r_{\nu}} x_{ik_{\mu}} \quad (2)$$

$S$  is the overlap matrix and  $x_{ik_{\mu}}$  is the LCAO coefficient of the  $k_{\mu}$ -th atomic orbital belonging to atom  $\mu$  in the  $i$ -th doubly occupied MO.

If the basis system is orthogonal, Mulliken's bond population vanishes. Formula (1) reduces, instead, to the more familiar Wiberg index <sup>(9)</sup>, which was introduced in order to deal with the CNDO/2 approximation. Expression (1) for  $I_{\mu\nu}$  is shown to be the electronic distribution along the  $\mu$ - $\nu$  bond <sup>(4)</sup> and it seems more appropriate than the usual Mulliken population analysis for a comparison with CNDO results.

Recently, a weight factor has been introduced in (1) <sup>(10)</sup>, avoiding thus the equipartition of the bond charge distribution; the theoretical consequences of this are now under investigation. Here, we shall restrict ourselves to the application of definition (1).

3. - H-bonding of methyl glyoxal and formamide with nucleic acids

Table I shows the  $I_{XY}$  values for the  $XH...Y$  bonds of the pairings of Fig. 1 between nucleic acid bases and methyl glyoxal (MG1). Geometries are taken from Refs. 2 and 11.

The possibility of these pairings was suggested taking into account that it is geometrically realizable. In a CNDO/2 calculation of the energies involved, it has been estimated that MG1 may compete with the bases for binding through hydrogen; this in turn could be related to a "brakelike" action on cell proliferation of malignant cells and would lead to an increase of the relative concentration of the amino-acids which are coded by A and U (<sup>2</sup>). The presence of  $CH...N$  or  $CH...O$  bonds has been raised as an objection against this model, with the argument that they are rather improbable (<sup>12</sup>); the interaction energies are moreover contested.

We desire hence to add to this controversy the information which bond indices supply. We take as references the  $I_{XY}$  values which may be correlated with experimental results (<sup>3,5</sup>); the Watson-Crick pairs have

$$\begin{aligned} \text{for } NH...O \quad I_{NO} &= 0.026 - 0.039 ; I_{HO} = 0.033 - 0.051 \\ \text{for } NH...N \quad I_{NN} &\cong 0.053 ; I_{HN} \cong 0.076 \end{aligned} \quad (3)$$

It is seen (Table I) that the  $I_{CN}$  values are higher than these and than all those calculated for the pairs appearing

in yeast tRNA<sup>Phe</sup>, except for one NH...N bond of a G-G pair. The  $I_{CO}$  values are lower, but similar to some  $I_{NO}$ 's. Of the four  $I_{NO}$  values, three are as big as the highest encountered in other pairs (5).

It has been argued that H-bonds with participation of carbon atoms result from an electrostatic attraction between the C and O(N) atoms (2); Table II shows that, in fact, their electronic densities are opposite in sign, while in usual XH...Y bonds both X and Y are negatively charged. There has been much discussion about whether this kind of hydrogen bonds are to be considered on the same grounds as other hydrogen bonds (13), or should be entitled "pseudo-hydrogen bonds". Exhaustive crystallographic information confirms the existence of CH...O, CH...N and even CH...Cl bonds (14). Bond indices values for the associations between methyl glyoxal and the nucleic acid bases, together with results not shown here, permit us to conclude that these bonds have the same status. Let us recall that the origin and nature of the hydrogen bond are still very far from clear (15,16).

If atom  $\mu$  furnishes  $N_{\mu}$  valence electrons to the molecule, its net charge will be  $Q_{\mu} = N_{\mu} - q_{\mu}$ . The total net charge transferred within a pairing will then be  $Q = \sum_{\mu} Q_{\mu}$ . In the pairings with MG1 we have

$$Q(U) = -0.050; Q(A) = 0.010; Q(C) = 0.007; Q(G) = -0.030 \quad (4)$$

so that MG1 behaves as an electron acceptor with respect to

A and C, and as donor with respect to U and G. Let us remark that U and G are acceptors also in the Watson-Crick pairs, where  $Q(U) = -0.051$  and  $Q(G) = -0.072$ .

As an example of the methylation influence on the charges of the pairs, we show in Fig. 2 the variations experienced by electronic densities when passing from G1-A to MG1-A. All the atoms not shown have the same  $q_{\mu}$ . The differences are not relevant and the  $I_{XY}$  values are the same.

Geometrically, MG1 could allow triple complexes analogous to those appearing in yeast tRNA<sup>Phe</sup> (17), for instance the U-MG1-A shown in Fig. 3. The four corresponding hydrogen bond indices are

$$I_{CO} = 0.018; I_{NO}(U-MG1) = 0.076; I_{CN} = 0.067; I_{NO}(A-MG1) = 0.051 \quad (5)$$

The somewhat low  $I_{CO}$  value is compensated by the big ones of the other three. Hydrogen bonds involving  $CH_3$  have been hinted at (13). This triple may shadow an opening for the interactions between MG1 and nucleic acids.

When DNA and RNA are dissolved in formamide (F), they appear to lose their formal secondary structure (18). The denaturation effect of F has been related to its ability to form hydrogen bonds with nucleic acids (19). Having been suggested that these H-bonds are stronger than those with glyoxal (2), we have calculated the bond indices of the pairs shown in Fig. 1.



Two recent analyses of experimental data for liquid formamide yield different conclusions regarding the ease of formamide to build hydrogen bonds. In a low frequency Raman study, chain structures are expected to be the dominant species <sup>(20)</sup>. According to electron, neutron and X-ray measures, instead, there should be four H-bonds per molecule <sup>(21)</sup>.

We see that  $I_{NN}$  in Table III is higher than the values of (3) for A-F and less than them for C-F;  $I_{NN}$  of A-F is as high as the  $I_{CN}$ 's of Table I. The enhanced possibility of interaction mentioned earlier is more evident for NH...O bonds since  $I_{NO}$ 's in Table III are appreciably higher than those of (3) and compete clearly with those of MGl. The bonding of F with A and U seems favoured with respect to the bonding with G and C.

The total net charges of the F-partner in the pairings of Table III are

$$Q(U) = 0.006; Q(A) = 0.014; Q(C) = -0.030; Q(G) = 0.000 \quad (6)$$

That is, U and A would be electron donors and C an acceptor with respect to F, while G would not form a charge-transfer complex with F. It has been noted that, when interacting with nucleic acids, F could behave either as an electron acceptor or donor <sup>(19)</sup>.

In an odd comparative calculation of EH and CNDO/2 atomic charges and dipole moments for about 15 molecules (involving C, O, N and H), it was found that multiplying EH values

by 0.3 gives approximate CNDO/2 values (<sup>22</sup>). We explored if a similar relationship could be found between bond indices of definition (1) and Wiberg's calculated with the usual CNDO/2 formalism, although hydrogen bonding would perhaps require a reparameterization (<sup>23</sup>). The pairs of Tables I and III (together with other H-bond associations not reported here) give the following approximate values for  $I_{XY}(\text{IEH})/I_{XY}(\text{CNDO}/2)$ :

$$\text{NN: } 0.17 \quad ; \quad \text{NO: } 0.21 \quad ; \quad \text{CN: } 0.25 \quad ; \quad \text{CO: } 0.27 \quad (7)$$

so that we may say, on an equal footing with Ref. 22, that H-bond IEH indices lead to CNDO/2 ones when multiplied by 0.2 - 0.3.

In short, the bond index results are in agreement with an H-bond model for the interactions between nucleic acids and MgI or F.

#### 4. - H<sub>2</sub>S-formamide interaction

The system H<sub>2</sub>S-formamide has been proposed as a model simulating the interaction between polypeptide groups and sulfur-containing substances (<sup>7</sup>). Ab initio calculations with different basis sets were performed for nine relative geometrical positions (Fig. 4); against expectation, in eight of them H<sub>2</sub>S turned out to act as an electron acceptor (<sup>7</sup>). We calculated hence the same nine configurations within the simple IEH framework; as in Ref. 7, we do not introduce d orbitals for S. Table IV

shows that in six of the nine configurations our results ascribe a donor character to  $H_2S$ . In the three other cases, the charge transfer is of the same order of magnitude as the ab initio one.

In configurations a), b), c), e), f) and g) the distance from sulfur to formamide is high, while in d), h) and i) it is much lower. In d) and i) we meet bond indices such as

$$\begin{array}{ll} \text{for d)} & I_{SO} = 0.684 & I_{SC} = 0.892 \\ \text{for i)} & I_{SO} = 0.779 & I_{SC} = 0.946 \end{array} \quad (8)$$

that must be related to the strikingly high charge transfer which these configurations display in Table IV; a charge transfer may be considered small if lower than 0.1<sup>(24)</sup>.

Case f) shows a NH...S bond with  $I_{NS} = 0.057$ . The deviation of this bond from linearity is 15°. Liquid formamide shows an average NH...O angle of 10°<sup>(21)</sup> and, in general, it is accepted that a departure from linearity up to 30° is not a relevant factor<sup>(25)</sup>. Charges are respectively -0.183 for N and -0.035 for S, and  $I_{HS} = 0.083$ . The data are consistent with all previous results, so that this bond has the characteristics of a conventional hydrogen bond.

Configuration g) shows, geometrically, a linear hydrogen bond. In it,  $I_{SO} = 0.040$  and  $I_{SH}$ , which is 0.985 in  $H_2S$ , becomes 0.959. It is known that, in fact, the XH bond weakens when entering a hydrogen bond<sup>(25)</sup>. However, we have

$I_{HO} = 0.037$ , while all the other cases which we considered up to now (<sup>3,5</sup>) obey  $I_{HY} > I_{XY}$  (see Tables I and III).

The ab initio calculations confront serious convergence difficulties, which were overcome by taking the diagonal blocks of the starting charge-bond order matrix of the complex from the final SCF results of the single constituent molecules and taking zeros in the off-diagonal blocks (<sup>7</sup>). We met convergence problems only for case h); as the distance between H<sub>2</sub>S and formamide is arbitrary, we displaced H<sub>2</sub>S up to an SO distance of 2.0 Å and thus reached convergence. Let us analyze what happens on both sides of the convergence limit. It has been noted that nonconvergence in SCF calculations is mainly due to the mixing among MO's belonging to the various subsets of the molecular basis, along the iterative path (<sup>26</sup>). In Fig. 5 we show that this is the case for the IEH calculation of configuration h). The HOMO and LEMO cross each other for the nonconvergent geometry; this denotes mixing among  $\sigma$  and  $\pi$  orbitals. A displacement in the relative positions suffices here to restore the ordering, as is seen in Fig. 5. We have also carried out a calculation with an SO distance of 1.5 Å, which converges. The  $I_{SO}$  value is completely upset, however; for 2.0 Å,  $I_{SO} = 0.019$  with a HOMO of  $\sigma$  character, while the 1.5 Å distance gives  $I_{SO} = 0.814$  and a  $\pi$  HOMO. Configuration h), therefore, converges both for a  $\sigma$  and a  $\pi$  HOMO, provided they maintain this character along the iterative path.

5. - Conclusions

- Bond indices results support the hypothesis that nucleic acids can interact through hydrogen bonds both with methyl glyoxal and formamide.

- Hydrogen bonds of the type CH...Y show bond index values similar to usual hydrogen bonds.

- The H<sub>2</sub>S group behaves as an electron donor with respect to formamide in six of the nine relative positions which in an ab initio calculation lead to the opposite result.

Bibliography

- (<sup>1</sup>) see for example K. Nakaya, O Takenaka, H. Horinishi and K. Shibata: *Biochim. Biophys. Acta*, 161, 23 (1968); R. Shapiro, B. I. Cohen, S.-J. Shiuey and H. Maurer: *Biochemistry*, 8, 238 (1969); N. E. Broude and E. I. Budowsky: *Biochim. Biophys. Acta*, 254, 380 (1971).
- (<sup>2</sup>) S. F. Abdulnur: *Int. J. Quantum Chem., Quantum Biol. Symp.*, 3, 59 (1976).
- (<sup>3</sup>) M. Giambiagi, M. S. de Giambiagi and W. B. Filho: *Chem. Phys. Lett.*, 78, 541 (1981).
- (<sup>4</sup>) M. Giambiagi, M. S. de Giambiagi, D. R. Gempel and C. D. Heymann: *J. Chim. Phys.*, 72, 15 (1975).
- (<sup>5</sup>) M. S. de Giambiagi, M. Giambiagi and D. M. S. Esquivel: *Z. Naturforsch.*, 38c, 000 (1983).
- (<sup>6</sup>) A. Szent-Györgyi: *Electronic Biology and Cancer* (Dekker, 1976), chap. 10.
- (<sup>7</sup>) P. Otto, S. Suhai and J. Ladik: *Int. J. Quantum Chem., Quantum Biol. Symp.*, 4, 451 (1977).
- (<sup>8</sup>) P. Dibout: *QCPE* 10, 256 (1973).
- (<sup>9</sup>) K. Wiberg: *Tetrahedron*, 24, 1083 (1968).
- (<sup>10</sup>) F. E. Jorge, M. Giambiagi and M. S. de Giambiagi: *Theoret. Chim. Acta*, in press.
- (<sup>11</sup>) L. Pauling: *The Nature of the Chemical Bond* (Cornell, 1960), pp. 306-308 and 502.

- (<sup>12</sup>) D. Demoulin, A.-M. Armbruster and B. Pullman: Theoret. Chim. Acta, 48, 143 (1978).
- (<sup>13</sup>) R. Gay and G. Vanderkooi: J. Chem. Phys., 75, 2281 (1981).
- (<sup>14</sup>) R. Taylor and O. Kennard: J. Am. Chem. Soc., 104, 5063 (1982).
- (<sup>15</sup>) M. D. Joesten and L. J. Schaad: Hydrogen Bonding (Dekker, 1974), p. 111.
- (<sup>16</sup>) R. M. Minyaev and V. I. Pavlov: J. Mol. Struct., THEOCHEM, 92, 205 (1983).
- (<sup>17</sup>) J. D. Robertus, J. E. Ladner, J. T. Finch, D. Rhodes, R. S. Brown, B. F. Clark and A. Klug: Nature, 250, 546 (1974).
- (<sup>18</sup>) P. O. P. T'so, G. K. Helmkamp and C Sander: Biochim. Biophys. Acta, 55, 584 (1962).
- (<sup>19</sup>) G. K. Helmkamp and P. O. P. T'so: J. Am. Chem. Soc., 83, 138 (1961).
- (<sup>20</sup>) O. Faurskov Nielsen, P.-A. Lund and E. Praestgaard: J. Chem. Phys., 77, 3878 (1982).
- (<sup>21</sup>) E. Kálmán, I Serke, G. Pálinkás, M. D. Zeidler, F. J. Wiesman, H. Bertagnolli and P. Chieux: Z. Naturforsch., 38a, 231 (1983).
- (<sup>22</sup>) P. R. Andrews: Theoret. Chim. Acta, 14, 261 (1969).
- (<sup>23</sup>) S. Noor Mohammad and A. J. Hopfinger: Int. J. Quantum Chem., 22, 1189 (1982).

(<sup>24</sup>) P. Otto and J. Ladik: Int. J. Quantum Chem., 18, 1143 (1980).

(<sup>25</sup>) L. Paoloni: Legami a idrogeno, in "Enciclopedia internazionale di chimica" (Pem, 1972), pp. 310-389.

(<sup>26</sup>) R. Carbó, J. A. Hernández and F. Sanz: Chem. Phys. Lett., 47, 581 (1977).



Figure captions

Figure 1 - Pairings of the nucleic acid bases with methyl glyoxal or formamide (F).

Figure 2 - Variation of some electronic densities in the pair glyoxal-adenine (G1-A) under methylation of glyoxal. The numbers indicate  $\Delta q_{\mu} [(MGl-A) - (G1-A)]$ .

Figure 3 - Possible U-MGl-A triple.

Figure 4 - The configurations proposed in Ref. 7 for the H<sub>2</sub>S-formamide interaction.

Figure 5 - Behaviour of HOMO (●) and LEMO (○) along the iterative path for two different SO distances in configuration h) of Fig. 4. The full line represents a  $\pi$  MO and the broken line a  $\sigma$  MO.

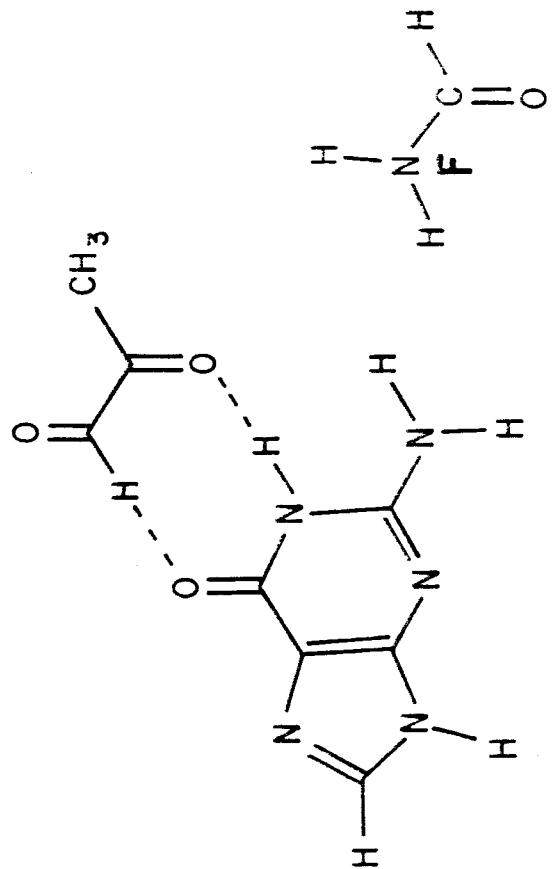
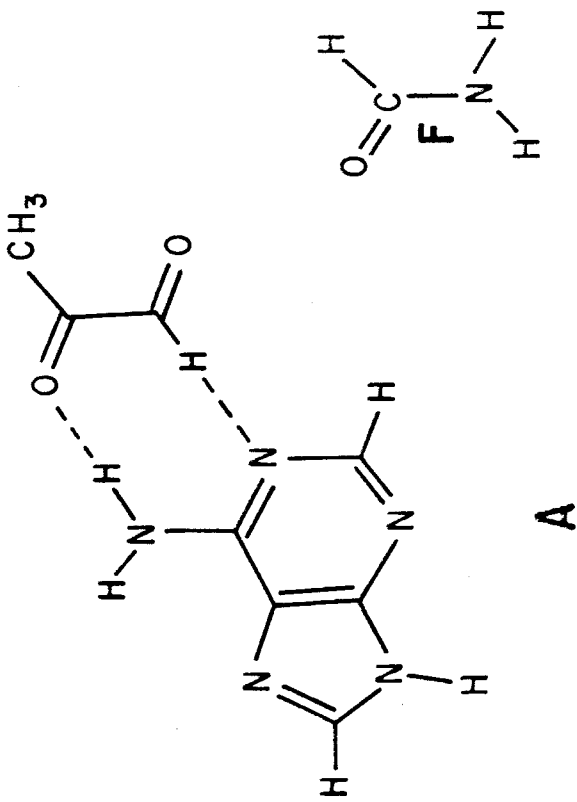
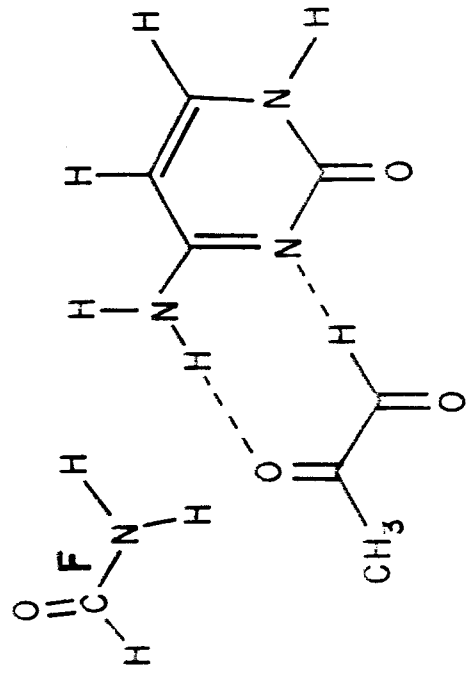
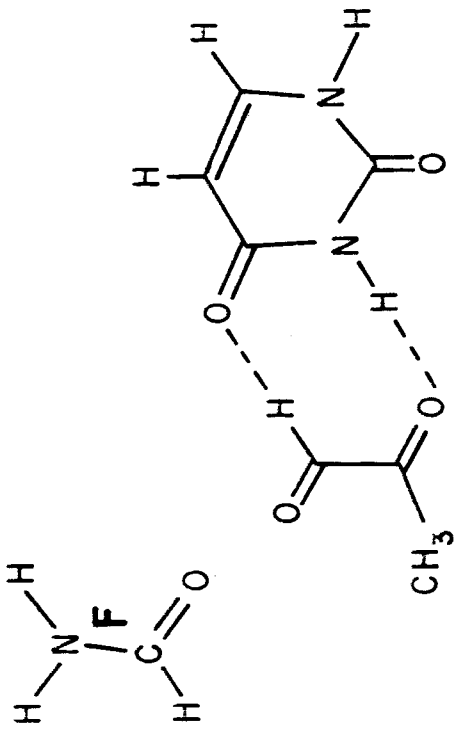


FIG. 1

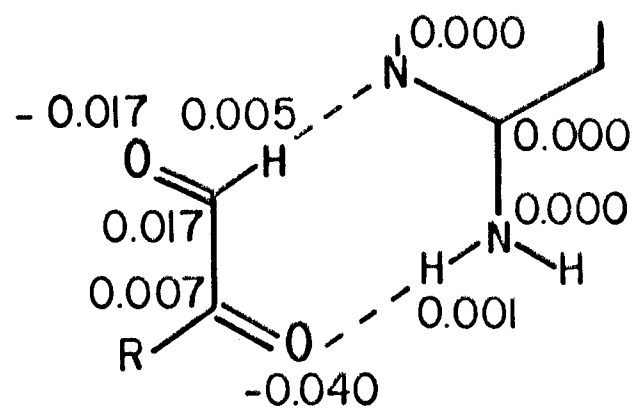


FIG. 2

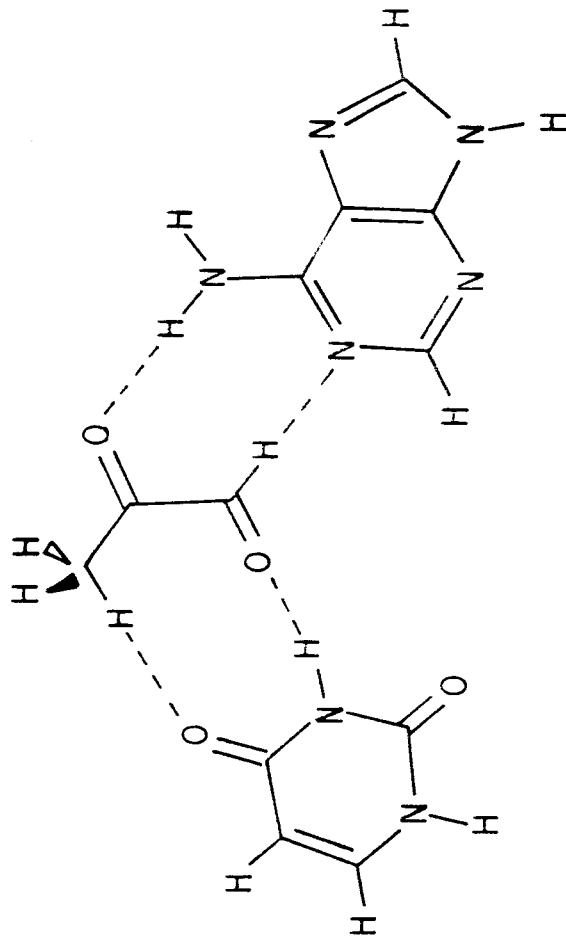


FIG. 3

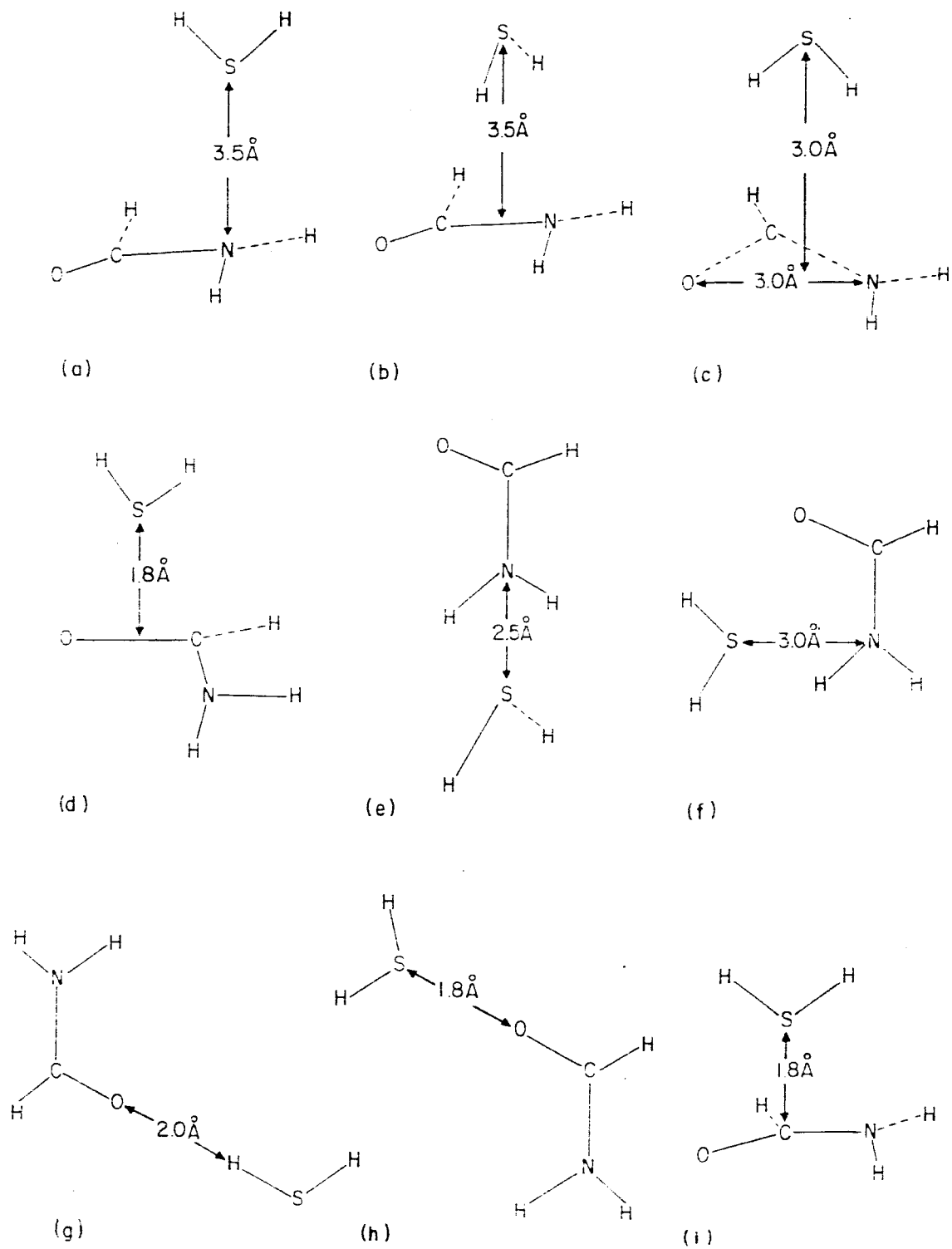


FIG 4

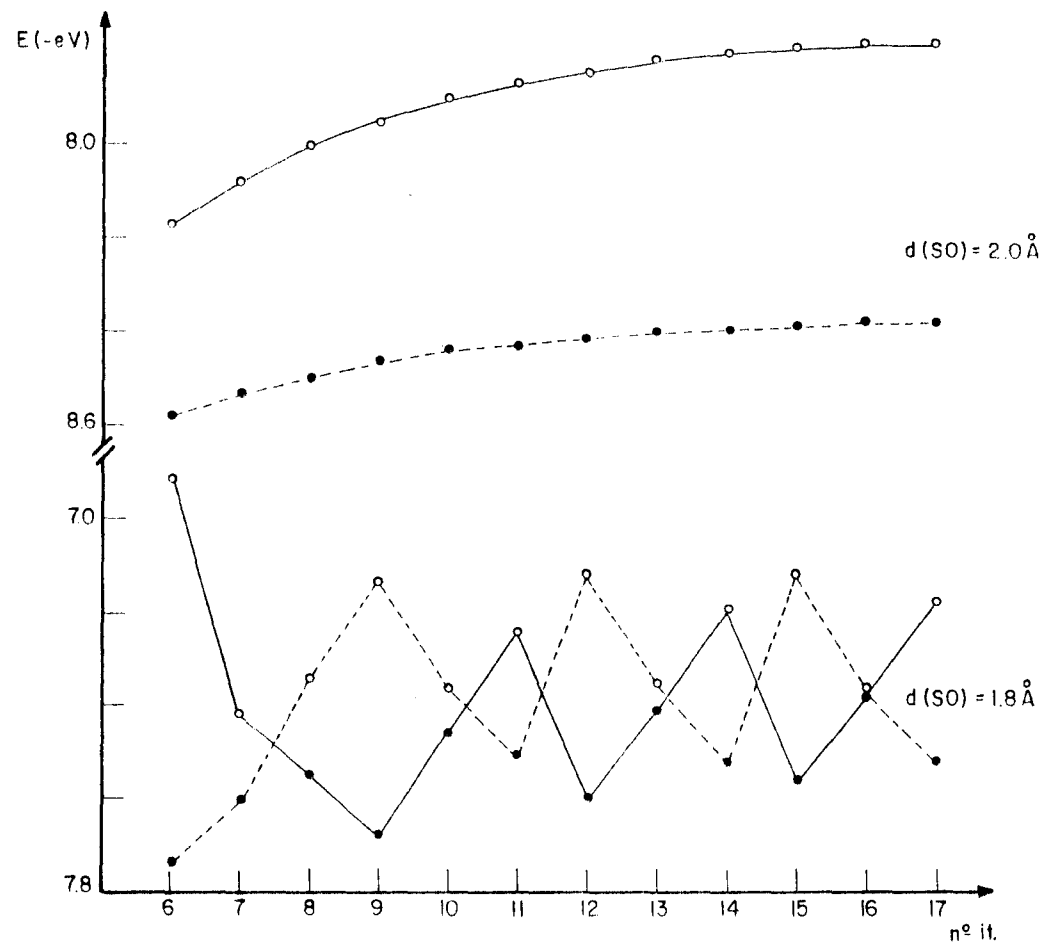


FIG. 5

Table I. - Bond indices I in XH...Y bonds of the pairs between methyl glyoxal (MG1) and the nucleic acid bases (see Fig. 1).

| Pair  | X | Y | $I_{XY}$ | $I_{HY}$ | $I_{XH}$ |
|-------|---|---|----------|----------|----------|
| MG1-U | C | O | 0.026    | 0.029    | 0.931    |
|       | N | O | 0.058    | 0.081    | 0.892    |
| MG1-A | C | N | 0.067    | 0.084    | 0.886    |
|       | N | O | 0.051    | 0.070    | 0.905    |
| MG1-C | C | N | 0.065    | 0.081    | 0.896    |
|       | N | O | 0.052    | 0.071    | 0.907    |
| MG1-G | C | O | 0.025    | 0.027    | 0.935    |
|       | N | O | 0.038    | 0.053    | 0.913    |

Table II. - Electronic densities of atoms X and Y of the XH...Y bonds appearing in the pairs between methyl glyoxal (MG1) and the nucleic acid bases (see fig. 1).

|       | MG1-U  | MG1-A  | MG1-C  | MG1-G  |
|-------|--------|--------|--------|--------|
| X = C |        | 0.088  | 0.093  |        |
| Y = N |        | -0.276 | -0.293 |        |
| X = C | 0.112  |        |        | 0.113  |
| Y = O | -0.266 |        |        | -0.301 |
| X = N | -0.150 | -0.175 | -0.180 | -0.106 |
| Y = O | -0.218 | -0.231 | -0.225 | -0.219 |



Table III. - Bond indices I in XH...Y bonds of the pairs between formamide (F) and the nucleic acid bases (see Fig. 1).

| Pair | X | Y | $I_{XY}$ | $I_{HY}$ | $I_{XH}$ |
|------|---|---|----------|----------|----------|
| U-F  | N | O | 0.067    | 0.095    | 0.880    |
|      | N | O | 0.061    | 0.086    | 0.886    |
| A-F  | N | O | 0.054    | 0.073    | 0.900    |
|      | N | N | 0.063    | 0.092    | 0.883    |
| C-F  | N | O | 0.059    | 0.076    | 0.900    |
|      | N | N | 0.032    | 0.047    | 0.904    |
| G-F  | N | O | 0.044    | 0.063    | 0.894    |
|      | N | O | 0.040    | 0.055    | 0.903    |

Table IV. - Charge transfer in the nine geometrical configurations of Ref. 7 proposed for the interaction H<sub>2</sub>S-formamide. Charges are referred to H<sub>2</sub>S. See Fig. 4.

| Configuration | IEH    | <u>Ab initio</u> (7) |
|---------------|--------|----------------------|
| a)            | 0.002  | -0.297               |
| b)            | -0.004 | -0.006               |
| c)            | -0.038 | -0.078               |
| d)            | 0.478  | -0.154               |
| e)            | 0.140  | -0.118               |
| f)            | 0.082  | -0.102               |
| g)            | -0.040 | -0.051               |
| h)            | 0.005  | -0.131               |
| i)            | 0.583  | 0.045                |