Molecular Electrostatic Potentials:
Comparison of *ab initio* and CNDO Results

C. Giessner-Prettre and A. Pullman

Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique associé au C.N.R.S.
13, rue P. et M. Curie, Paris 5è, France

The possibilities of utilization of CNDO wave functions for computing molecular electrostatic potentials are studied by comparison with *ab initio* results for H_2O and H_2CO.

Les possibilités d'utilisation de fonctions d'onde CNDO pour le calcul des potentiels électrostatiques moléculaires sont étudiées par comparaison avec des résultats *ab initio* pour H_2O et H_2CO.

Die Möglichkeiten der Verwendung von CNDO-Wellenfunktionen zur Berechnung molekularer elektrostatischer Potentiale werden durch Vergleich mit *ab initio* Rechnungen für H_2O und H_2CO untersucht.

Several nonempirical studies using wavefunctions from either STO minimal basis sets [1–3] or GTO basis sets [4, 5] have shown that the value of the electrostatic potential created by the electronic distribution and the nuclear charges of a molecule in the different regions of the space surrounding it gives valuable informations about the sites involved in protonation or in reactions with electrophilic agents. Large-scale non-empirical computations of this sort being unfortunately precluded, we have set out to investigate the possibilities of the CNDO-method in this field: in a preliminary study [6] we have determined within the CNDO/2 semi-empirical framework the isopotential curves for adenine and guanine and have observed that they allow a very neat distinction between the two molecules as to the most probable regions for electrophilic attack, in remarkable conformity with the known experimental facts. On the other hand, the CNDO maps for adenine agree qualitatively with the corresponding *ab initio* maps available [4]. However, some differences appear upon detailed comparison, particularly as concerns the depth of the potential wells and the location of the bottom of these wells.

These differences led us to search for their origin and to try to find an approximation for the calculation of the electrostatic potential from CNDO/2 wave functions which would be in closer agreement with non empirical results. For performing this test we chose water and formaldehyde. For water, non-empirical isopotential maps from a wave function expressed in a minimal STO basis are available for comparison [7]. The choice of formaldehyde was made, aside from the small size of the molecule and the existence in the literature of several *ab initio* minimal STO wave functions which could be used to calculate the non empirical isopotential maps [8, 9], because the structure of protonated formaldehyde is experimentally known [10].

Computational Details

The interaction energy between a molecular distribution and an external unitary positive charge placed at point P is given by [2]:

$$V_P = \sum_A \frac{Z_A}{R_{PA}} - \sum_{\mu\nu} P_{\mu\nu} \int \frac{\chi_\mu \chi_\nu}{r_P} \, d\tau,$$

where the first summation runs over the nuclei of the molecule and the second is carried over all the electrons of the molecule. The integrals appearing in the second summation are the nuclear attraction integrals of a hydrogen nucleus on the distribution $\chi_\mu \chi_\nu$. If the wave functions utilized to calculate the $P_{\mu\nu}$ elements of the density matrix are obtained from CNDO/2, Z_A has to be taken equal to the number of valence electrons.

Approximation I. If one chooses to stay exactly within the CNDO/2 approximations, only the diagonal elements of the density matrix are retained and the nuclear attraction integrals are approximated by minus the repulsion integrals between s orbitals. This was the approximation we used previously [6], in which:

$$V_P = \sum_A \left(\frac{Z_A}{R_{PA}} - P_{AA} \gamma_{AH} \right).$$

This approximation corresponds to Fig. 1.

Approximation II. Another possibility is to keep the same elements of the density matrix, but to abandon the γ-approximation and to use instead the corresponding correct nuclear attraction integrals V_{AH} over s atomic orbitals (so as to retain the spherical symmetry of the interaction). The corresponding maps are those of Fig. 2.

Approximation III. Instead of averaging the nuclear attraction integrals on each atom A, all the nuclear attraction integrals are computed exactly and introduced in the summation with the corresponding $P_{\mu\nu}$ element, provided that χ_μ
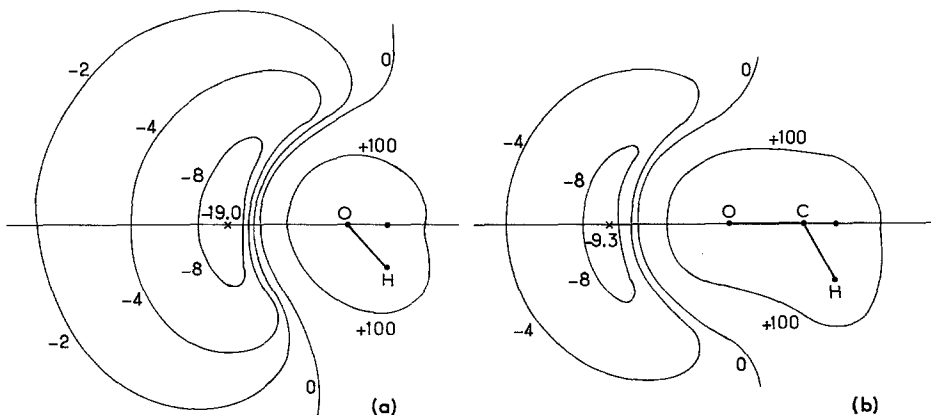

Fig. 1a and b. Electrostatic molecular potential energy maps for a) water molecule, b) formaldehyde, using Approx. I. Energies are in kcal/mole. The upper part of the figure corresponds to the y or z plane, the lower part to the molecular plane
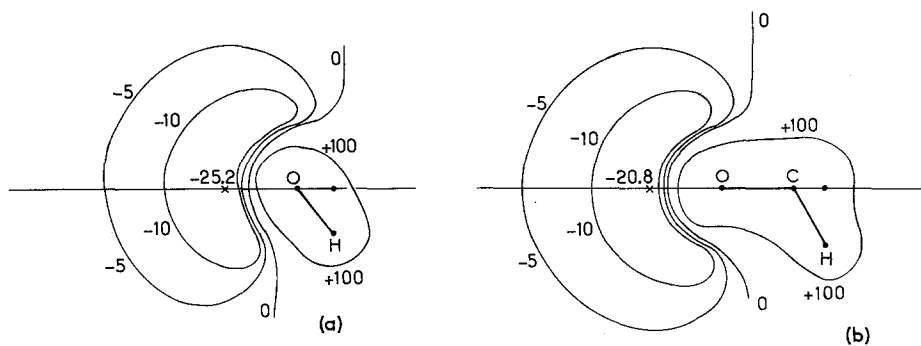
Fig. 2. Same as Fig. 1. Approx. II

and χ_v are located on the same atom. The only nuclear attraction terms left out are the three-center terms. This approximation corresponds to the maps of Fig. 3.

Approximation IV. The first three approximations utilize the usual CNDO/2 density matrix made with the coefficients admittedly considered as coefficients over Slater orbitals in spite of the ZDO approximation. Now, the CNDO eigen-vectors may be conceived instead [11] as expressed in terms of orthogonalized orbitals λ derived from Slater orbitals χ by the Löwdin's transformation [12]. Thus, one may retransform the CNDO coefficients into coefficients over Slater orbitals by the matrix product:

$$C^\chi = S^{-1/2} C^\lambda ,$$

where S is the overlap matrix over the Slater orbitals, and use these "deortho-gonalized" coefficients for calculating a new density matrix. This approach was utilized for studying charges and dipole moments [13, 14], and very recently, density distributions [20].

If one uses this deorthogonalized density matrix for the computation of the potentials, all the nuclear attraction integrals must be retained since the two-center distributions $\chi_\mu \chi_v$ are no longer zero. This approximation of the potential V_P differs from an exact computation only by the freezing of the 1s electrons into the nuclei and by the fact that the density matrix is obtained through a CNDO deorthogonalization process instead of a complete *ab initio* SCF procedure. This approximation corresponds to Fig. 4.
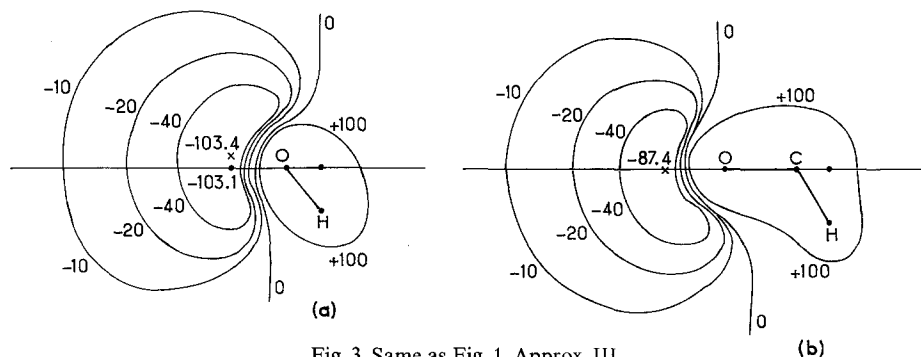
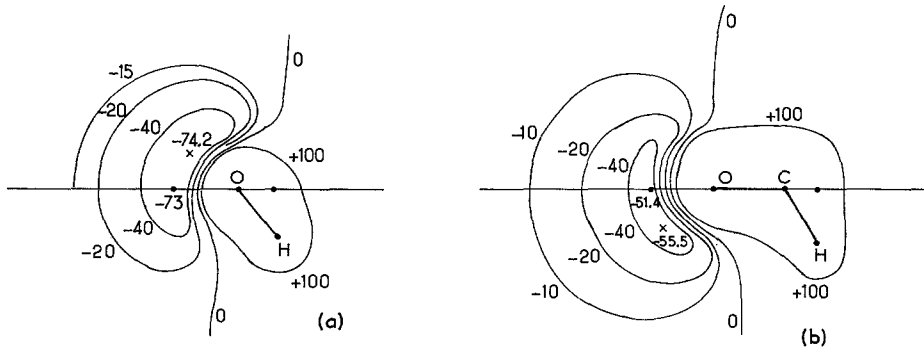

Fig. 3. Same as Fig. 1. Approx. III

Fig. 4. Same as Fig. 1. Approx. IV

Reference ab initio Computations. The reference curves for H_2O (Fig. 5a) are those of Bonaccorsi *et al.* [7], corresponding to the molecule-optimized Slater basis set of Aung, Pitzer and Chan [15].

For H_2CO we have computed the isoenergy values using two different wave functions: one in a Slater minimal basis set (best atom ζ's with $\zeta_H = 1.2$) [8], the other in a molecule-optimized Slater basis (isotropic) [9]. The corresponding maps are those of Figs. 5b1, and 5b2 respectively.

For these potential computations (and those of Approx. IV), we have used an adaptation of a part of an integral program written by Stevens [16].
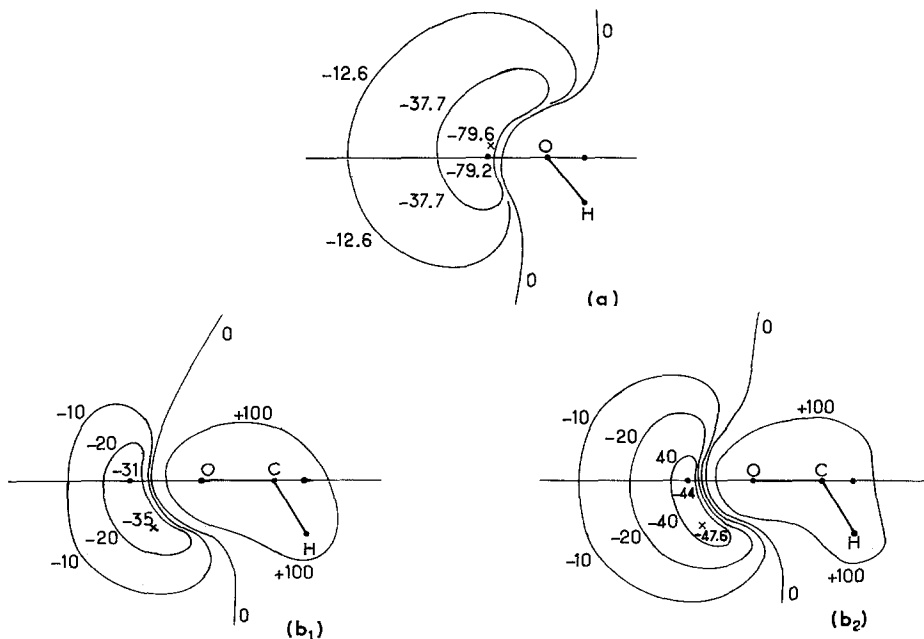

Fig. 5a and b. Non-empirical electrostatic potential energy maps for: a H_2O of Ref. [7], b1 H_2CO (Slater exponents, $\zeta_H = 1.2$), b2 H_2CO (optimized isotropic basis)

Results and Discussion

The maps of Approx. I (full CNDO) show a similar behaviour for the two molecules: the potential wells are rather shallow and located on the molecular axis. These results when compared to the non-empirical ones of Bonaccorsi *et al.* for water and ours for formaldehyde show that the computed values are too small (in absolute value) and that the bottoms of the wells are located too far from the oxygen atoms when approximating the nuclear integrals by the γ-values. This appears as a general deficiency of Approx. I, which has partly been discussed before [6]. As expected, replacing the γ's by the average V_{AH}-values (Approx. II) increases the numerical potential values and brings the minima closer to the molecule (Fig. 2). But in the two approximations, the minima are located on the molecular symmetry axis for both molecules whereas the *ab initio* calculations find the bottoms of the potential wells away from the molecular axis, somewhat above (and below) the molecular plane in the case of water, and in the molecular plane but clearly away from the C=O axis in the case of formaldehyde (Fig. 5). For this molecule the non-empirical isopotential maps are in agreement with experimental findings on the structure of protonated formaldehyde: high resolution proton magnetic resonance indicates unambiguously [10] an unsymmetrical structure for the protonated species[1].

As concerns the non-empirical curves it is interesting to note that the location of the potential well is very similar whatever basis set is utilized. From a qualitative point of view the two sets of curves are very similar, but they differ in the extension of the attractive zone and in the numerical value of the minima. The greater extension occurs for the optimized basis and is probably related to the fact that the optimized exponents for the oxygen atom are smaller than the corresponding Slater exponents.

Approx. III leads to numerical values of the potentials which are much larger than in the preceeding computations, and even larger than the *ab initio* values. For water the largest potential is now slightly off the molecular plane. But for formaldehyde the molecular axis is still found to be the most attractive region for a positve charge.

If we turn now to Approx. IV, we observe a better quantitative agreement with non-empirical calculations. The numerical agreement between the two types of calculations shows that the accuracy of the representation of the core electrons (the $1s$ electrons from oxygen and carbon atoms in the present cases) plays a negligible role on the value of the electrostatic potential at the distances considered. On the other hand, the introduction of the two-center distributions seems required for obtaining the directionality of the attraction energy when the bottoms of the potential wells are very small depressions in a rather deep arc-shaped valley, as it is the case in the neighbourhood of oxygen atoms.

As far as the depth of potential wells is concerned, Approx. IV appears in better agreement with the values obtained from the wave function using an optimized basis set than those from a wave function using regular Slater exponents.

---

[1] The experimental evidence on the geometry of H_3O^+ is less clearcut: crystal data indicate pyramidal structures but the free ion is claimed to be planar [17–19] with very easy distortion.

Conclusion

These calculations show that one way to obtain *quantitatively* good isopotential maps from a CNDO/2 wavefunction is to transform the eigenvectors from the orthogonal basis set into a regular Slater basis and to introduce all the integrals. Although the calculation of nuclear integrals is very rapid for one point of space, the large number of points required for a map renders the computer time necessary for such calculations rapidly prohibitive when the size of the molecule increases. However, for large molecules this approximation might be useful for the exploration of a relatively small region of space where a preliminary study has shown that the potential well has the largest probability to be found. Such an exploratory study may be performed very rapidly using Approx. II. (The use of Approx. I, which is less good, and slower, may be preferred only when consistency with other CNDO results is desired).

References

1. Bonaccorsi, R., Petrongolo, E., Scrocco, E., Tomasi, J.: Quantum aspects of heterocyclic compounds in chemistry and biochemistry, Jerusalem Symposia, Vol. II, 181 (1970).
2. — Scrocco, E., Tomasi, J.: J. chem. Physics **52**, 5270 (1970).
3. — — — Theoret. chim. Acta (Berl.) **20**, 331 (1971).
4. — Pullman, A., Scrocco, E., Tomasi, J.: Theoret. chim. Acta (Berl.), in press.
5. — — — — Chem. Physics Letters **12**, 622 (1972).
6. Giessner-Prettre, C., Pullman, A.: C. R. Acad. Sci. **272** c, 750 (1971).
7. Bonaccorsi, R., et al.: Private communication and Ref. [3].
8. Aung, S., Pitzer, R. M., Chan, S. I.: J. chem. Physics **45**, 3457 (1966).
9. Switkes, E., Stevens, R. M., Lipscomb, W. N.: J. chem. Physics **51**, 5229 (1969).
10. Olah, G. A., O'Brien, D. H., Calin, M.: J. Amer. chem. Soc. **89**, 3582 (1967).
11. Pople, J. A., Segal, G. A.: J. chem. Physics **43**, S 136 (1965).
12. Löwdin, P. O.: J. chem. Physics **18**, 365 (1950).
13. Giessner-Prettre, C., Pullman, A.: Theoret. chim. Acta (Berl.) **11**, 159 (1968).
14. Shillady, D. D., Billingsley, F. P., Bloor, J. E.: Theoret. chim. Acta (Berl.) **21**, 1 (1961).
15. Aung, S., Pitzer, R. M., Chan, S. I.: J. chem. Physics **49**, 2071 (1961).
16. Stevens, R. M.: J. chem. Physics **52**, 1397 (1970).
17. Grahn, R.: Arkiv fysik **19**, 147 (1961).
18. de Paz, M., Ehrenson, S., Friedman, L.: **52**, 3362 (1970).
19. Moskowitz, J. M., Harrison, M. C.: J. chem. Physics **43**, 3550 (1965).
20. MacIver, J. W., Jr., Coppens, P., Novak, D.: Chem. Physics Letters **11**, 82 (1971).

Prof. Dr. A. Pullman
Institut de Biologie Physico-Chimique
13, rue Pierre et Marie Curie
Paris 5è, France