Theoret. chim. Acta (Berl.) 31, 75—82 (1973)

# *Ab initio* Calculations on Large Molecules Using Molecular Fragments. Characterization of the Zwitterion of Glycine*

Lester L. Shipman** and Ralph E. Christoffersen***
Department of Chemistry, University of Kansas, Lawrence, Kansas 66044

Received April 19, 1973

The *ab initio* molecular fragment approach is applied to a characterization study of the ground state of the zwitterion of glycine. Included among the properties studied are the $\phi - \psi$ conformational energy surface, the electronic structure, and the magnitude and direction of the dipole moment. The results of the present study are compared to the results of other theoretical and experimental studies.

*Key words:* Zwitterion of glycine – Conformational energy surface – *Ab initio* molecular fragment approach.

## 1. Introduction

Glycine, a biologically important molecule having the distinction of being the smallest amino acid, exists as a zwitterion (see Fig. 1) in aqueous solution in the range of biological pH's [1]. In previous theoretical studies, the $\phi - \psi$ conformational energy surface of the zwitterion of glycine has been studied by Ponnuswamy and Sasisekharan [2] using an empirical potential energy function; Imamura *et al.* [3a] and Oegerle *et al.* [3b] have studied the electronic structure and preferred geometry of the zwitterion and neutral forms of glycine in CNDO studies; and the zwitterion of glycine has been characterized in the X-ray crystallographic geometry [4, 5] of α-glycine in an *ab initio* study by Ryan and Whitten [6].

In the present study, the *ab initio* SCF molecular fragment approach [7–19] is applied to a characterization study of the ground state of the zwitterion of glycine. These studies serve partly to provide comparisons with experimental and other theoretical studies, and partly to elucidate the ability of the molecular fragment approach to describe systems with large charge polarization. The entire $\phi - \psi$ conformational energy surface is studied, and a $\phi - \psi$ conformational energy contour map is presented. The molecular orbital structure, as well as the magnitude and direction of the dipole moment vector are discussed for the conformation of lowest energy.

---

* This work was supported in part by the National Science Foundation, the University of Kansas, and the Upjohn Company, Kalamazoo, Michigan 49001.

** NSF Trainee (1969–1972). Current address — Department of Chemistry, Cornell University, Ithaca, New York 14850.

*** Alfred P. Sloan Research Fellow (1971–1973). Author to whom correspondence should be directed.

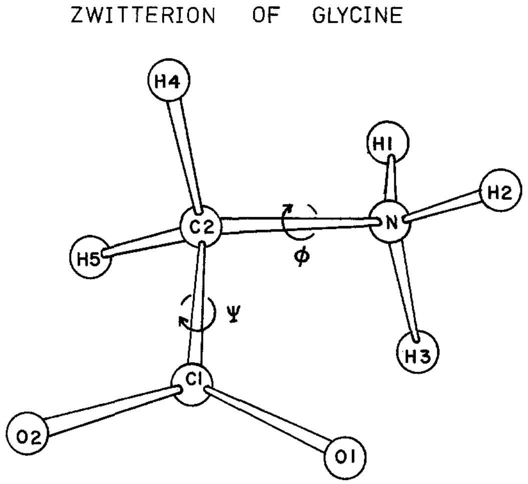

Fig. 1. Nuclear numbering system and definition of $\phi$ and $\psi$ for the zwitterion of glycine

The nuclear geometry used in the present study is taken from the recently determined neutron diffraction geometry of $\alpha$-glycine by Jönnsson and Kvick [20]. This geometry is modified slightly according to several constraints, so that many fewer calculations are needed to construct a grid for the entire conformational energy surface in $\phi(\sphericalangle \mathrm{H3{-}N{-}C2{-}C1})$ and $\psi(\sphericalangle \mathrm{N{-}C2{-}C1{-}O1})$ [21]. The convention that $(\phi, \psi) = (0, 0)$ [22] is the fully eclipsed conformation is used, and the constraints applied are the following:

1) A $\phi$-rotation of 120° results in a final structure that is identical to the initial structure.

2) A $\psi$-rotation of 180° results in a final structure that is identical to the initial structure.

3) Reflection through the C1–C2–N plane takes H4 identically into H5, and vice versa.

The imposition of these constraints results in only small deviations from the experimental structure (see Table 1). The largest bond distance deviation is 0.015 Å and the RMS deviation is 0.007 Å, while the largest bond angle deviation is 1.2°, and the RMS deviation is 0.6°. With these three constraints imposed, the entire conformational energy surface may be studied with $\phi$ running from 0 to 120° and $\psi$ running from 0–180° and also, the surface has inversion symmetry about the point (60, 90).

The molecular fragment basis set [7–19], consisting of 30 floating spherical Gaussian functions (FSGO) [23] that are contracted to 25 basis functions, is described in Table 2. Since the methods for determining these basis orbitals and for forming large molecules from them have been described earlier, [7–19] they will not be repeated here.

## 2. $\phi$—$\psi$ Conformational Energy Surface

A grid over the conformational energy surface [24] was calculated in 15° increments in $\phi$ and $\psi$. The absolute minimum was found at the point (60, 0). Using linear interpolation, a contour map was constructed (see Fig. 2) with

Table 1. Nuclear geometry

| Bond | Experimental[a] | Constrained[b] |
|---|---|---|
| Bond distances (in Å) | | |
| C1–O1 | 1.250 | 1.250 |
| C1–O2 | 1.251 | 1.250 |
| C1–C2 | 1.526 | 1.526 |
| C2–H4 | 1.090 | 1.090 |
| C2–H5 | 1.089 | 1.090 |
| C2–N | 1.476 | 1.476 |
| N–H1 | 1.054 | 1.039 |
| N–H2 | 1.037 | 1.039 |
| N–H3 | 1.025 | 1.039 |

| Angle | Experimental[a] | Constrained[b] |
|---|---|---|
| Bond angles (in degrees) | | |
| O2–C1–O1 | 125.4 | 125.4 |
| O1–C1–C2 | 117.5 | 117.3 |
| O2–C1–C2 | 117.1 | 117.3 |
| C1–C2–N | 111.9 | 111.9 |
| C1–C2–H4 | 108.8 | 109.7 |
| C1–C2–H5 | 110.5 | 109.7 |
| H4–C2–H5 | 108.0 | 108.0 |
| N–C2–H4 | 108.5 | 108.8 |
| N–C2–H5 | 109.0 | 108.8 |
| C2–N–H1 | 112.1 | 111.4 |
| C2–N–H2 | 111.7 | 111.4 |
| C2–N–H3 | 110.4 | 111.4 |
| H1–N–H2 | 108.7 | 107.5 |
| H1–N–H3 | 107.1 | 107.5 |
| H2–N–H3 | 106.6 | 107.5 |

[a] See Ref. [20].

[b] The geometry used in the present studies.

Table 2. Molecular fragment data[a]

| Fragment type | FSGO type | FSGO Distance from "heavy" atom | FSGO radii ($\varrho$) |
|---|---|---|---|
| $CH_4 (T_d)$ | C–H | 1.23379402 | 1.67251562 |
| R(C, H) = 2.05982176 | C inner shell | 0.0 | 0.32784375 |
| $\cdot CH_3$ (planar) | C–H | 1.13093139 | 1.51399487 |
| R(C, H) = 1.78562447 | C–$\pi$ | ±0.1 | 1.80394801 |
| | C inner shell | 0.0 | 0.32682735 |
| · OH (*sp* hybrid) | O–H | 0.76467773 | 1.23671871 |
| R(O, H) = 1.54774058 | O–LP ($\sigma$) | 0.21614258 | 1.28753780 |
| | O–LP ($P$) | ±0.1 | 1.19741696 |
| | O–$\pi$ | ±0.1 | 1.12242182 |
| | O inner shell | 0.00057129[b] | 0.24028227 |
| $NH_4^+ (T_d)$ | N–H | 0.80547793 | 1.50046875 |
| R(N, H) = 1.95021656 | N inner shell | 0.0 | 0.27770068 |

[a] All distances are in Hartree atomic units, see Shull, H., Hall, G. G.: Nature **184**, 1559 (1959).

[b] This is the distance from the oxygen nucleus, along the O–H bond axis, toward the H nucleus.

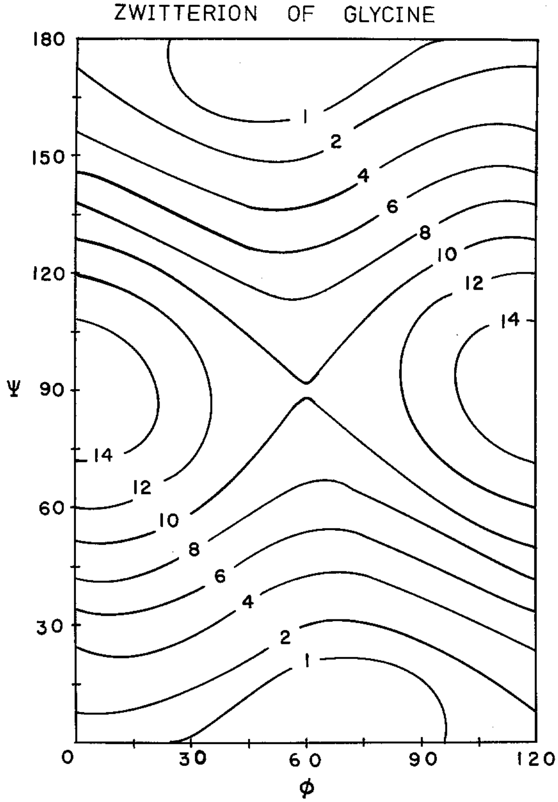

Fig. 2. $\phi - \psi$ conformational energy surface for the zwitterion of glycine. Contours are given in units of kcal/mole relative to the energy of the conformation of lowest energy

contours in kcal/mole relative to the absolute minimum. Note that the conformational energy surface is saddle-shaped with a saddle point at (60, 90). Also, the $\phi - \psi$ map indicates that a wide range of conformations is accessible. For example, within a 2 kcal/mole range, any value of $\phi$ is accessible when $\psi = 0°$. For a rotation starting and terminating at the absolute minimum energy conformation, the lowest barrier for $\phi$-rotation is 1.5 kcal/mole at (0, 0) and the lowest barrier for $\psi$-rotation is 10.1 kcal/mole at (60, 90).

Using an empirical energy function, Ponnuswamy and Sasisekharan [2] have also found the minimum to be at (60, 0), and 0.5 and 1 kcal/mole contours were plotted. It is of interest to compare the extent from the minimum of the 1 kcal/mole contour in that study and in the present study. This contour has roughly the same maximum extent in the $\phi$-direction in the two studies, with the contour of the present study having the largest extent. The greatest extent of the 1 kcal/mole contour in the $\psi$-direction in the study of Ponnuswamy and Sasisekharan is to $\psi \approx 60°$, while in the present study this contour extends to $\psi \approx 20°$.

Imamura *et al.* [3a], in a CNDO study, fixed $\phi$ at 60° and found that $\psi = 0°$ was the energetically preferred conformation. They found a barrier for $\psi$-rotation of 5.8 kcal/mole at (60, 90).

There are three crystalline forms of glycine, $\alpha$ [4, 5, 20], $\beta$ [25], and $\gamma$ [26], which differ primarily in the network of hydrogen bonds between zwitterions. $\phi$ is known to be roughly 60° in $\alpha$-glycine [20] but $\phi$ has not been determined in $\beta$ and $\gamma$-glycine. $\psi$ is known to be approximately 18.6° in $\alpha$-glycine, 24.8° in $\beta$-glycine, and 12.8° in $\gamma$-glycine [26]. Note that, for a range of values of $\phi$ (see Fig. 2), these $\psi$ values correspond to conformations that are less than 2 kcal/mole from the minimum energy conformation. In particular (58.3, 18.6), is the point in the constrained geometry closest to the experimental geometry of $\alpha$-glycine [20], and this point has an energy 0.8 kcal/mole above the minimum. However, comparison of the preferred conformation predicted by the present study with the conformation of the monomer in $\alpha$-glycine from neutron diffraction must be made with caution. The results of the present study apply to an isolated molecule, and the effects of the crystal environment on monomer conformation could be considerable in the case of $\alpha$-glycine. For example, Oegerle and Sabin [3b] have estimated that the displacement of the nitrogen atom from the plane of the heavy atoms in the crystal is a crystal packing effect, that can occur with only very little cost in energy of the zwitterion. Thus, the preferred values of $\phi$ and $\psi$ may be determined by orientational requirements of *intermolecular* hydrogen bonding in the crystal, since O1, O2, H1, H2, H3, (and possibly H4 and H5) may participate in intermolecular hydrogen bonds. However, the high monomer conformational energy associated with values of $\psi$ near 90° should be sufficient to make these values unfavorable in the crystal.

## 3. Dipole Moment

The dipole moment, $\mu$, of the zwitterion of glycine has been determined in the present study in the minimum energy conformation, (60, 0). $\mu$ is constrained by symmetry to lie in the C1–C2–N plane. $\mu$ has a magnitude of 13.33 D [27] and forms an angle of 29.1° with the C1–C2 bond. This result is in excellent agreement with experimental value of 13.3 D for the zwitterion of glycine in aqueous solution as measured by Buckingham [28], and in moderate agreement (11%) with the experimental value of 15.0 D for the zwitterion of glycine in water-alcohol solution as measured earlier by Kirkwood [29]. Ryan and Whitten [6], in an *ab initio* study, have calculated the dipole moment of the zwitterion of glycine in the X-ray crystallographic geometry to be 12.17 D, while Oegerle and Sabin [3b] have calculated a value of 13.410 D for a geometry optimized using the CNDO procedure.

## 4. Electronic Structure

The elctronic structure of the zwitterion of glycine has been studied at the minimum energy conformation, (60, 0), and at the point (in the constrained geometry) closest to the neutron diffraction geometry [20], (58.3, 18.6). These results are presented in Table 3, along with the results from the *ab initio* study of Ryan and Whitten [6] at the X-ray crystallographic geometry [4, 5]. In the present study, each molecular orbital is characterized by means of a population analysis over the symmetric orthonormalized basis functions [18]. As has been generally observed in previous studies [7–19], the ordering of the valence molecular orbitals is identical to that observed in more extensive basis set studies. Further-

Table 3. Electronic structure[a]

| ($\phi$, $\psi$) Energy minimum[b] | | Experimental geometry[c] | | | |
|---|---|---|---|---|---|
| Present calc. | | Present calc.[e] | | Ryan and Whitten[f] | |
| MO type[d] | $\varepsilon_i$ | MO type[d] | $\varepsilon_i$ | MO type[d] | $\varepsilon_i$ |
| O2 (1*s*) | −16.893 | O2 (1*s*) | −16.893 | O2 (1*s*) | −20.363 |
| O1 (1*s*) | −16.865 | O1 (1*s*) | −16.865 | O1 (1*s*) | −20.348 |
| N (1*s*) | −13.219 | N (1*s*) | −13.220 | N (1*s*) | −15.885 |
| C2 (1*s*) | − 9.318 | C2 (1*s*) | − 9.318 | C1 (1*s*) | −11.488 |
| C1 (1*s*) | − 9.126 | C1 (1*s*) | − 9.126 | C2 (1*s*) | −11.448 |
| N–H | − 1.322 | N–H | − 1.323 | N–H | − 1.500 |
| C–O ($\sigma$) | − 1.127 | C–O ($\sigma$) | − 1.127 | C–O ($\sigma$) | − 1.344 |
| C–O ($\sigma$) | − 1.022 | C–O ($\sigma$) | − 1.021 | C–O ($\sigma$) | − 1.221 |
| C–H, C–C | − 0.894 | C–H, C–C | − 0.895 | $\sigma$ | − 1.123 |
| N–H | − 0.664 | N–H | − 0.665 | N–H | − 0.963 |
| N–H | − 0.659 | N–H | − 0.659 | N–H | − 0.946 |
| C–N | − 0.539 | C–N | − 0.540 | $\sigma$ | − 0.807 |
| C–H, C–C | − 0.477 | C–H, C–C | − 0.477 | $\sigma$ | − 0.751 |
| C–H | − 0.452 | C–H | − 0.451 | $\sigma$ | − 0.736 |
| O($sp-lp$), C–O ($\sigma$) | − 0.250 | O($sp-lp$), C–O ($\sigma$) | − 0.249 | O1, O2 ($lp$) | − 0.575 |
| $\pi_b$ | − 0.217 | $\pi_b$ | − 0.217 | $\pi_b$ | − 0.523 |
| O($sp-lp$) | − 0.160 | O($sp-lp$) | − 0.159 | O1, O2 ($lp$) | − 0.504 |
| O($P-lp$) | + 0.025 | O($p-lp$) | + 0.026 | O1, O2 ($lp$) | − 0.358 |
| O($P-lp$) | + 0.037 | O($P-lp$) | + 0.038 | O1, O2 ($lp$) | − 0.339 |
| $\pi_n$ | + 0.087 | $\pi_n$ | + 0.088 | $\pi_n$ | − 0.289 |
| $\pi_a$ | + 0.680[g] | $\pi_a$ | + 0.654[g] | | |

[a] All energies are reported in Hartree atomic units. See Shull, H., Hall, G. G.: Nature **184**, 1559 (1959).

[b] ($\phi$, $\psi$) = (60, 0)

[c] The neutron diffraction geometry (Ref. [20]), X-ray crystallographic geometry (Refs. [4] and [5]), and constrained geometry are quite similar. See the text for a discussion of the differences between the constrained and neutron diffraction geometry. The X-ray geometry is essentially the same as the neutron diffraction geometry except that the heavy atom-hydrogen bond distances are 13–17% smaller.

[d] Where a given molecular orbital has a contribution of at least 50% (1.0 electron) of a particular type (as measured by a population analysis over the symmetric orthonormalized basis functions), then that contribution is indicated symbolically. If no single contribution comprises 50% of the molecular orbital, then the primary and secondary contributions are listed, separated by a comma.

[e] ($\phi$, $\psi$) = (58.3, 18.6).

[f] These results are from an *ab initio* SCF calculation of Ryan and Whitten (Ref. [6]) at the X-ray crystallographic geometry, using a basis set of 152 spherical Gaussian functions. The notation used to label the molecular orbitals has been changed from that found in Ref. [6] to an equivalent notation, in some instances.

[g] Lowest unoccupied molecular orbital.

more, the detailed nature of the molecular orbitals is seen to be identical at the level of characterization in the Ryan and Whitten study, except that the ordering of the inner shell molecular orbitals involving C1 and C2 is reversed. Disappointingly, the highest three occupied molecular orbitals have positive orbital energies in the present study. However, a plot of the valence molecular orbital energies, $\varepsilon_i$ (R-W), for the Ryan and Whitten study vs. the corresponding orbital energies, $\varepsilon_i$, for the present study at (58.3, 18.6) is approximately linear. This type of behavior has been observed in a wide variety of examples using molecular fragment basis sets [30], and appears to be an indication of the good balance of basis sets obtained using the molecular fragment procedure. A least squares fit of the current data to a straight line results in the following equation,

$$\varepsilon_i(\text{R-W}) = 0.8571 \cdot \varepsilon_i - 0.362\,, \tag{1}$$

where the RMS deviation from this straight line is 0.016. Thus, even though the molecular orbital energies have been shifted upward and the spacings increased slightly, the proper ordering of valence molecular orbitals is maintained. Note also that the character of the various molecular orbitals is the same for the (60, 0) and (58.3, 18.6) conformations, and that the molecular orbital energies are nearly the same, with the largest change occurring for the orbital energy of the lowest unoccupied molecular orbital.

## 5. Discussion and Summary

One of the more striking aspects of this study is the appearance of a rather wide range of conformations that are accessible to the isolated zwitterion, that are in the vicinity of the lowest energy conformer (60, 0). For example, the expected ethane-type rotational barrier for rotation about $\phi$ (for $\psi = 0°$) is substantially lower (1.5 kcal/mole) than is calculated in ethane itself (5.60 kcal/mole [8]) using the molecular fragment procedure. This rather flat energy surface for $\phi$-rotation is apparently due to the presence of an additional effect in the case of the glycine zwitterion that is not present in ethane itself. In particular, the nuclear positions for an optimum intramolecular hydrogen bond between H3 and O1 are present at approximately (0, 0). Thus, the stabilization of the molecules by an intramolecular hydrogen bond offsets the destabilization that is due to an "ethane-type" eclipsing at (0, 0), and results in a rather flat ($\phi$, 0) curve.

Because of the low conformational energies associated with a $\phi$-rotation and the fact that the intramolecular hydrogen bonded conformation, (0, 0), lies only 1.5 kcal/mole above the lowest calculated energy conformation, the zwitterion of glycine probably assumes those values of $\psi$ in aqueous solution and in the crystalline state that facilitate *intermolecular* hydrogen bonding to the solvent water molecules in the former case and to other zwitterions in the latter case. Because of the higher energies associated with $\psi$-rotations, it is anticipated that the values of $\psi$ will remain in the $0° \leqq \psi \leqq 30°$ range. However, within this range, those values of $\psi$ which facilitate intermolecular hydrogen bonding in aqueous solution and in the crystalline state would be expected to be favored, as in the case of $\phi$.

The dipole moment, $\boldsymbol{\mu}$, was found to have a magnitude of 13.33 D, in good agreement with the experimental values, and $\boldsymbol{\mu}$ was found to form an angle of 29.1° with the C1–C2 bond. The molecular orbital ordering obtained in the present study was in agreement with the ordering obtained by Ryan and Whitten [6], and an approximately linear relationship was found to hold between the valence molecular orbital energies calculated by Ryan and Whitten and those calculated in the present study.

Thus, even for molecules in which rather large charge polarization is encountered, the molecular fragment procedure appears to provide a well balanced description of several molecular properties of interest, and appears to be appropriate for description of molecular systems of considerable size and variety.

*Acknowledgement.* The authors would like to express their appreciation to the University of Kansas for partial support of the computing costs required for these studies.

## References

1. Hückel, W.: Theoretical principles of organic chemistry, Vol. II, 150–157. New York: Elsevier 1958
2. Ponnuswamy, P. K., Sasisekharan, V.: Biopolymers **7**, 624 (1969)

3a Imamura, A., Fujita, H., Nagata, C.: Bull. Chem. Soc. Japan **42**, 3118 (1969)
3b Oegerle, W. R., Sabin, J. R.: J. Molec. Struct. **15**, 131 (1973)

4. Albrecht, G., Corey, R. B.: J. Am. Chem. Soc. **61**, 1087 (1939)
5. Marsh, R. E.: Acta Cryst. **11**, 654 (1958)
6. Ryan, J. A., Whitten, J. L.: J. Am. Chem. Soc. **94**, 2396 (1972)
7. Christoffersen, R. E., Maggiora, G. M.: Chem. Phys. Lett. **3**, 419 (1969)
8. Christoffersen, R. E., Genson, D. W., Maggiora, G. M.: J. Chem. Phys. **54**, 239 (1971)
9. Christoffersen, R. E.: J. Am. Chem. Soc. **93**, 4104 (1971)
10. Christoffersen, R. E.: Adv. Quan. Chem. **6**, 333 (1972)
11. Maggiora, G. M., Genson, D. W., Christoffersen, R. E., Cheney, B. V.: Theoret. chim. Acta (Berl.) **22**, 337 (1971)
12. Christoffersen, R. E., Shipman, L. L., Maggiora, G. M.: Int. J. Quant. Chem. **5S**, 143 (1971)
13. Cheney, B. V., Christoffersen, R. E.: J. Chem. Phys. **56**, 3503 (1972)
14. Genson, D. W., Christoffersen, R. E.: J. Am. Chem. Soc. **94**, 6904 (1972)
15. Shipman, L. L., Christoffersen, R. E.: Chem. Phys. Lett. **15**, 469 (1972)
16. Genson, D. W., Christoffersen, R. E.: J. Am. Chem. Soc. **95**, 362 (1973)
17. Shipman, L. L., Christoffersen, R. E.: Proc. Nat. Acad. Sci. US **69**, 3301 (1972)
18. Shipman, L. L., Christoffersen, R. E.: J. Am. Chem. Soc. **95**, 1408 (1973)
19. Weiman,, L. J., Christoffersen, R. E.: J. Am. Chem. Soc. **95**, 2074 (1973)
20. Jönsson, P.-G., Kvick, Å.: Acta Cryst. B **28**, 1827 (1972)
21. For a 15° grid in $\phi$ and $\psi$, 576 calculations would be required for the experimental geometry, while only 50 calculations are required for the constrained geometry
22. Throughout this manuscript conformations will be specified by the notation, $(\phi, \psi)$
23. These basis orbitals were introduced and utilized for small molecules by Frost, A. A., *and coworkers*: J. Chem. Phys. **54**, 764 (1971), and in earlier references contained therein. Formulation and characterization of these orbitals for large molecule calculations is given in [7–19]
24. Since the $\phi - \psi$ surface has inversion symmetry about the point (60, 90), the discussion of this surface will be limited to the region $0° \leqq \phi < 120°$, $0° \leqq \psi \leqq 90°$, without loss of generality
25. Iitaka, Y.: Acta Cryst. **13**, 35 (1960)
26. Iitaka, Y.: Acta Cryst. **14**, 1 (1961)
27. The dipole moment has been calculated according to the conversion factor, 2.54158 D/a.u.
28. Buckingham, A. D.: Aust. J. Chem. **6**, 323 (1953)
29. Kirkwood, J. G.: J. Chem. Phys. **2**, 351 (1934)
30. This approximately linear relationship between orbital energies from extensive basis set calculations and those of the molecular fragment procedure has been observed in 30 molecules to date, and is apparently not restricted to any specific type of molecule. For example, the cases studied to date include linear and non-linear molecules, saturated and unsaturated molecules, and molecules with or without hetero-atoms

Prof. Dr. R. E. Christoffersen
Department of Chemistry
University of Kansas
Lawrence, Kansas 66044
USA