Ab Initio Calculations on Large Molecules Using Molecular Fragments. Model Peptide Studies[1a]

Lester L. Shipman[1b] **and Ralph E. Christoffersen***[1c]

Contribution from the Department of Chemistry, University of Kansas,
Lawrence, Kansas 66044. Received July 27, 1972

Abstract: An *ab initio* procedure designed for the investigation of large molecules and based upon studies of molecular fragments is applied to prototype peptide systems. The molecules investigated include formamide, *N*-methylacetamide, and 2-formamidoacetamide, which were chosen as models of interactions that are expected along a peptide chain. The ability of the procedure to describe these systems is assessed by comparisons with experimental and other theoretical studies, and recommendations regarding the general applicability of the procedure to polypeptide characterizations are given.

The importance of a detailed understanding of peptide bonds and their interactions to many problems of biological interest has long been recognized. Consequently, a systematic characterization of prototype molecules which exhibit the kinds of interactions that are expected to be found in polypeptides is particularly appropriate.

The obvious importance of such studies has given rise to a number of previous quantum-mechanical studies of peptide systems. Some of these have been *ab initio* studies,[2–10] while the majority have used more approximate methods, mostly semiempirical.[11–53] In the current study, an *ab initio* procedure[54–62] (based upon the study of molecular fragments) that has been designed explicitly for use on large molecules has been employed to investigate several important aspects of peptide bonds and their interaction with their environment. Since the formulation and characterization of the model have been given in earlier studies,[54–62] they

---

(1) (a) This study was supported in part by the University of Kansas and by grants from the National Science Foundation and the Upjohn Co., Kalamazoo, Mich. 49001. (b) NSF Trainee (1969–1972). (c) Alfred P. Sloan Research Fellow, 1971–1973.

(2) H. Basch, M. B. Robin, and N. A. Kuebler, *J. Chem. Phys.*, **47**, 1201 (1967).

(3) M. A. Robb and I. G. Csizmadia, *Theor. Chim. Acta*, **10**, 269 (1968).

(4) M. A. Robb and I. G. Csizmadia, *J. Chem. Phys.*, **50**, 1819 (1969).

(5) J. B. Moffat, *J. Theor. Biol.*, **26**, 437 (1970).

(6) (a) M. Dreyfus, B. Maigret, and A. Pullman, *Theor. Chim. Acta*, **17**, 109 (1970); (b) M. Dreyfus and A. Pullman, *ibid.*, **19**, 20 (1970).

(7) D. H. Christensen, R. N. Kortzeborn, B. Bak, and J. J. Led, *J. Chem. Phys.*, **53**, 3912 (1970).

(8) C. R. Brundle, D. W. Turner, M. B. Robin, and H. Basch, *Chem. Phys. Lett.*, **3**, 292 (1969).

(9) R. Ditchfield, J. E. Del Bene, and J. A. Pople, *J. Amer. Chem. Soc.*, **94**, 703 (1972).

(10) J. A. Ryan and J. L. Whitten, *ibid.*, **94**, 2396 (1972).

(11) R. Rein, N. Fukuda, H. Win, and G. A. Clarke, *J. Chem. Phys.*, **45**, 4743 (1966).

(12) A. Julg and P. Carles, *Theor. Chim. Acta*, **7**, 103 (1967).

(13) J. A. Pople and M. Gordon, *J. Amer. Chem. Soc.*, **89**, 4253 (1967).

(14) J. E. Bloor, B. R. Gilson, and F. P. Billingsly, II, *Theor. Chim. Acta*, **12**, 360 (1968).

(15) J. F. Yan, F. A. Momany, R. Hoffmann, and H. A. Scheraga, *J. Phys. Chem.*, **74**, 420 (1970).

(16) A. Pullman and H. Berthod, *Theor. Chim. Acta*, **10**, 461 (1968).

(17) F. A. Momany, R. F. McGuire, J. F. Yan, and H. A. Scheraga, *J. Phys. Chem.*, **74**, 2424 (1970).

(18) K. N. Shaw and L. W. Reeves, *Chem. Phys. Lett.*, **10**, 89 (1971).

(19) M. G. Evans and J. Gergeley, *Biochim. Biophys. Acta*, **3**, 188 (1949).

(20) S. Nagakura, *Mol. Phys.*, **3**, 105 (1960).

(21) M. Suard, G. Berthier, and B. Pullman, *Biochim. Biophys. Acta*, **52**, 254 (1961).

(22) M. Suard, *ibid.*, **59**, 227 (1962).

(23) M. Suard, *ibid.*, **64**, 400 (1962).

(24) A. Julg and P. Carles, *Theor. Chim. Acta*, **1**, 140 (1963).

(25) S. Yomosa, *Biopolym. Symp.*, **1**, 1 (1964).

(26) K. Kaya and S. Nagakura, *Theor. Chim. Acta*, **7**, 117 (1967).

(27) D. Poland and H. A. Scheraga, *Biochemistry*, **6**, 3791 (1967).

(28) R. Hoffmann and A. Imamura, *Biopolymers*, **7**, 207 (1969).

(29) L. B. Kier and J. M. George, *Theor. Chim. Acta*, **14**, 258 (1969).

(30) A. Rossi, C. W. David, and R. Schor, *ibid.*, **14**, 429 (1969).

(31) B. Maigret, B. Pullman, and M. Dreyfus, *J. Theor. Biol.*, **26**, 321 (1970).

(32) C. H. Bushweller, P. E. Stevenson, J. Golini, and J. W. O'Neil, *J. Phys. Chem.*, **74**, 1155 (1970).

(33) A. S. N. Murthy, K. G. Rao, and C. N. R. Rao, *J. Amer. Chem. Soc.*, **92**, 3544 (1970).

(34) A. S. N. Murthy, S. N. Bhat, and C. N. R. Rao, *J. Chem. Soc. A*, 1251 (1970).

(35) B. Maigret, B. Pullman, and D. Perahia, *Biopolymers*, **10**, 107 (1971).

(36) B. Pullman, B. Maigret, and D. Perahia, *Theor. Chim. Acta*, **18**, 44 (1970).

(37) B. Maigret, D. Perahia, and B. Pullman, *J. Theor. Biol.*, **29**, 275 (1970).

(38) B. Maigret, D. Perahia, and B. Pullman, *Biopolymers*, **10**, 491 (1971).

(39) B. Pullman, B. Maigret, and D. Perahia, *C. R. Acad. Sci., Ser. D*, **270**, 1396 (1970).

(40) G. Govil, *J. Chem. Soc. A*, 386 (1971).

(41) D. Perahia, B. Maigret, and B. Pullman, *Theor. Chim. Acta*, **19**, 121 (1970).

(42) A. R. Rossi, C. W. David, and R. Schor, *J. Phys. Chem.*, **74**, 4551 (1970).

(43) M. Maigret, B. Pullman, and J. Caillet, *Biochem. Biophys. Res. Commun.*, **40**, 808 (1970).

(44) G. D. Zeiss and M. A. Whitehead, *J. Chem. Soc. A*, 1727 (1971).

(45) B. Maigret, B. Pullman, and D. Perahia, *J. Theor. Biol.*, **31**, 269 (1971).

(46) J. Almlöf and O. Mårtensson, *Acta Chem. Scand.*, **25**, 355 (1971).

(47) B. Pullman, J. L. Coubeils, P. Courriere, and D. Perahia, *Theor. Chim. Acta*, **22**, 11 (1971).

(48) F. A. Momany, R. F. McGuire, J. F. Yan, and H. A. Scheraga, *J. Phys. Chem.*, **75**, 2286 (1971).

(49) R. F. McGuire, F. A. Momany, and H. A. Scheraga, *ibid.*, **76**, 375 (1972).

(50) R. Schor, H. Stymne, G. Wettermark, and C. W. David, *ibid.*, **76**, 670 (1972).

(51) G. Govil, *J. Chem. Soc. A*, 2464 (1970).

(52) G. Govil, *J. Indian Chem. Soc.*, **48**, 731 (1971).

(53) G. Govil and A. Saran, *J. Chem. Soc. A*, 3624 (1971).

(54) R. E. Christoffersen and G. M. Maggiora, *Chem. Phys. Lett.*, **3**, 419 (1969).

(55) R. E. Christoffersen, D. W. Genson, and G. M. Maggiora, *J. Chem. Phys.*, **54**, 239 (1971).

(56) G. M. Maggiora, D. W. Genson, R. E. Christoffersen, and B. V. Cheney, *Theor. Chim. Acta*, **22**, 337 (1971).

(57) R. E. Christoffersen, *J. Amer. Chem. Soc.*, **93**, 4104 (1971).

(58) R. E. Christoffersen, *Advan. Quantum. Chem.*, **6**, 333 (1972).

(59) R. E. Christoffersen, L. L. Shipman, and G. M. Maggiora, *Int. J. Quantum Chem.*, **5S**, 143 (1971).

(60) B. V. Cheney and R. E. Christoffersen, *J. Chem. Phys.*, **56**, 3503 (1972).

(61) D. W. Genson and R. E. Christoffersen, *J. Amer. Chem. Soc.*, **94**, 6904 (1972).

(62) L. L. Shipman and R. E. Christoffersen, *Chem. Phys. Lett.*, **15**, 469 (1972).

will not be repeated here. The current studies are also designed to assess the accuracy to be expected in future studies, when the procedure is applied to larger polypeptides, by comparison with more extensive basis set calculations and experimental results, where available. Only ground states have been included in the current investigations, and the many interesting questions regarding description of excited states have been deferred, in order to ascertain the adequacy of ground-state descriptions first.

Three molecules have been chosen as "model" peptides. The first of these, formamide, was selected as the simplest model of a peptide bond, and is also a molecule for which extensive experimental data and theoretical calculations are available for comparison. *N*-Methylacetamide was selected as a model for the interactions that occur between an amide unit (-CONH-) and attached saturated hydrocarbon groups, using a methyl group at each end of the amide unit as a model. Finally, 2-formamidoacetamide (formylglycinamide) was selected as a model for the interaction of two amide units, when the space and orientations in which they can interact are restricted by a linking tetrahedral carbon. Formamide is shown in Figure 1 in the microwave geometry of Costain and Dowling,[63] and *N*-methylacetamide and 2-formamidoacetamide are shown in their minimum energy conformations as calculated in the current study.

The basis orbitals for these calculations are floating spherical Gaussian orbitals (FSGO)[64] and the molecular fragments in which their nonlinear parameters were determined, along with the optimized parameter values are given in Table I.

Table I. Molecular Fragment Data[a]

Fragment type	FSGO type	FSGO distance from "heavy" atom	FSGO radii, ρ
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-π	±0.1	1.80394801
	C inner shell	0.0	0.32682735
$\cdot$OH (sp hybrid)	O–H	0.76467773	1.23671871
R(O,H) = 1.54774058	O–LP (σ)	0.21614258	1.28753780
	O–LP (P)	±0.1	1.19741696
	O-π	±0.1	1.12242182
	O inner shell	0.00057129[b]	0.24028227
$:NH_3$ (planar)	N–H	0.75201903	1.39424495
R(N,H) = 1.93131910	N-π	±0.1	1.50625972
	N inner shell	0.0	0.27684894
$:NH_3$ (T_d)	N–H	0.87735349	1.52791683
R(N,H) = 1.91242167	N–LP	0.25523498	1.58328000
	N inner shell	0.00099090	0.27732014

[a] See ref 69. [b] This is the distance from the oxygen nucleus along the OH bond axis, toward the H nucleus.

The peptide nomenclature used in this paper will be in accordance with the tentative rules established by the IUPAC–IUB Commission on Biochemical Nomenclature.[65] For that portion of the study dealing with the nonplanarity of formamide, the notation proposed by Winkler and Dunitz,[66] as a logical extension of the IUPAC–IUB nomenclature to nonplanar peptides, will be used.

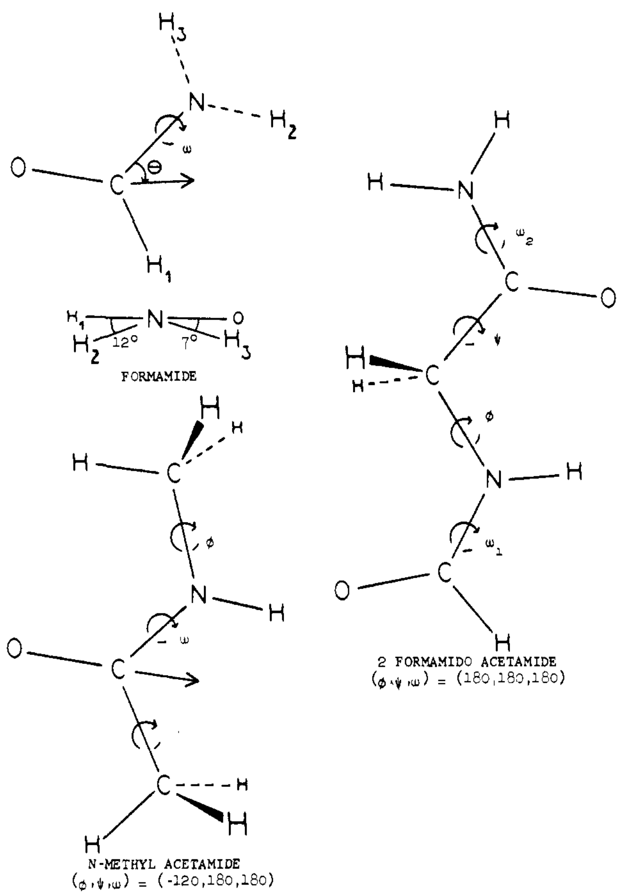

Figure 1. Model peptides studied.

Formamide

When selecting a geometry to use for formamide, one finds that there are several experimentally determined structures[63,67,68] from which to choose, including two microwave studies, one by Kurland and Wilson,[67] another later study by Costain and Dowling,[63] and one X-ray study by Ladell and Post.[68] Since the *ab initio* calculations are for molecules in a vacuum, the microwave structures are more suitable for the current calculations than is the X-ray structure. Since a more extensive body of microwave data was available in the study of Costain and Dowling, the structural parameters assigned in their study will be employed for the current calculations.

A. Electronic Structure. The orbital energies,[69] total energies, and other properties of interest that have been calculated in *ab initio* SCF studies to date are given in Table II (the current study is labeled calculation A). Also given in this table is information as to the size of

---

(63) C. C. Costain and J. M. Dowling, *J. Chem. Phys.*, **32**, 158 (1960).

(64) These basis orbitals were introduced and utilized for small molecules by A. A. Frost and coworkers, *J. Chem. Phys.*, **54**, 764 (1971), and in earlier references contained therein. Formulation and characterization of these orbitals as basis orbitals for large molecule calculations is given in ref 54–62.

(65) *Biochemistry*, **9**, 3471 (1970).

(66) F. K. Winkler and J. D. Dunitz, *J. Mol. Biol.*, **59**, 169 (1971).

(67) R. J. Kurland and E. B. Wilson, *J. Chem. Phys.*, **27**, 585 (1957).

(68) J. Ladell and B. Post, *Acta Crystallogr.*, **7**, 559 (1954).

(69) All energies and distances, unless otherwise specified, are reported in Hartree atomic units. See H. Shull and G. G. Hall, *Nature (London)*, **184**, 1559 (1959).

Table II. *Ab Initio* Calculations of Molecular Properties of Formamide[a]

Property	Calcn A	B	C	D	E	F	G
ϵ_{12}	−0.069 (π_n)	−0.413 (π_n)	−0.425 (π_n)	−0.425 (π_n)	−0.266 (σ_n)	−0.276 (σ_n)	−0.145 (π_n)
ϵ_{11}	−0.085 (σ_n)	−0.427 (σ_n)	−0.441 (σ_n)	−0.445 (σ_n)	−0.429 (π_n)	−0.364 (π_n)	−0.145 (σ_n)
ϵ_{10}	−0.287	−0.598	−0.578	−0.586	−0.455	−0.462	−0.342
ϵ_9	−0.298	−0.608	−0.617	−0.618	−0.646	−0.561	−0.375
ϵ_8	−0.379	−0.677	−0.675	−0.679	−0.650	−0.606	−0.489
ϵ_7	−0.466	−0.751	−0.751	−0.758	−0.970	−0.856	−0.575
ϵ_6	−0.615	−0.878	−0.847	−0.853	−1.039	−1.127	−0.736
ϵ_5	−1.009	−1.213	−1.216	−1.222	−1.676	−1.832	−1.149
ϵ_4	−1.224	−1.396	−1.410	−1.431	−2.011	−2.368	−1.316
ϵ_3	−9.104		−11.369	−11.383	−12.042	−12.046	−11.067
ϵ_2	−12.815		−15.596	−15.602	−16.674	−16.819	−15.378
ϵ_1	−17.071		−20.537	−20.548	−21.855	−22.676	−20.233
E_{total}	−143.7746		−168.9630	−168.9155	−156.7234	−162.5272	−159.3919
$\lvert\mu\rvert$, D (θ)	3.91 (37.3°)		4.07 (44.4°)	4.14 (44.2°)			1.41 (−62.3°)
ΔE_{rot}[b]	19.72		21.72	19.89			
Ref	*g*	Extrapolation[g]	*c*	*c*	*d*	*d*	*e*
No. of Gaussians	15		138	120	25	31	21
No. of orbitals	15		78	60	25	31	21
Geometry employed	CD	CD	CD	CD	*X*, *Y*, *Z* coord given in paper	*X*, *Y*, *Z* coord given in paper	CD, plus planar, 120° angles at N
Geometry ref	*f*	*f*	*f*	*f*	*d*	*d*	*f*

[a] Hartree atomic units are employed, unless otherwise specified. See ref 69. [b] ΔE_{rot} is the torsional rotation barrier, expressed in units of kcal/mol. [c] Reference 7. [d] Reference 5. [e] Reference 3. [f] Reference 63. [g] Current study.

the Gaussian basis set and geometry that was utilized. This list is not intended to be exhaustive since, in some studies, calculations were conducted for a range of molecular geometries. The calculations included in Table II are primarily those that were carried out at "standard" geometries.

Since the symmetry of formamide is not high, comparison of the "balance" of the basis sets of various procedures by examination of the ordering of molecular orbital irreducible representation labels is not particularly useful. On the other hand, the two highest occupied molecular orbitals are identifiable as a π-type nonbonding orbital encompassing the OCN moiety (π_n), and a σ-type nonbonding orbital that is primarily an oxygen lone pair, p-type orbital (σ_n). It is also of interest to note that increasing the basis set size does not automatically guarantee better results. For example, the use of basis sets that are considerably larger than the basis set used in the current study have been employed in some cases (calculations E and F), but the correct ordering of these two important molecular orbitals was not obtained, when compared with the most extensive calculation available (calculation C).

By means of Hartree–Fock molecular orbital population analyses, further comparisons were made between the current results and those of the most extensive calculation (calculation C). It was found that the ordering of molecular orbitals produced using the molecular fragment approach is in exact agreement with the ordering predicted by the most extensive calculation. Furthermore, as shown in Figure 2, the orbital energy values calculated using the molecular fragment approach are closely related to those calculated using a near Hartree–Fock wave function. In particular, except for inner shell orbital energies (which are not related in a linear fashion to the near Hartree–Fock energies and are omitted from the figure), the plot shows a very nearly linear relationship between the orbital energies calculated from the two wave functions. It should be noted that such a result is not limited to formamide. In fact, similar linear relationships have been found to exist for a wide variety of molecules (*e.g.*, benzene, diimide, ethane, ethylene, formaldehyde, formic acid, furan, hydrogen peroxide, methanol, methylamine, naphthalene, pyrazine, pyridine, pyrrole, and water). In no case studied to date for orbital energy relationships has the relationship failed to be linear.[70] Thus it is evident that, even though the basis sets of the molecular fragment procedure are quite small, the relative importance of the various interactions in the molecule are apparently correctly assessed. Quantifying this notion slightly, a least-squares fit of the points in Figure 2 to a straight line gives rise to the equation[69]

$$\epsilon_i(\text{C}) = 0.8511\epsilon_i(\text{A}) - 0.354 \qquad (\pm 0.015) \qquad (1)$$

and this best straight-line fit is also displayed in Figure 2. Used in another fashion, the linear relation on the right-hand side of eq 1 has been used to estimate what the ϵ_i for the near Hartree–Fock calculation would be, and these values are listed as "calculation B" in Table II. The obvious closeness of these values to the actual

---

(70) These relationships can also be used for predictive purposes, *e.g.*, in isoelectronic molecular systems. Applications of this nature for cumulenes and related molecules are given in L. J. Weimann and R. E. Christoffersen, *J. Amer. Chem. Soc.*, in press.

values (calculation C) emphasizes further the ability of the molecular fragment approach to produce basis sets that are balanced in a manner similar to that found in near Hartree–Fock calculations.

B. Dipole Moment. Values for the magnitude of the dipole moment of formamide as calculated using various quantum-mechanical methods and nuclear geometries[13,17,63,67,68,71] are also given in Table II. The angle θ is the angle between the dipole moment vector ($\mathbf{\mu}$) and the C–N bond (see Figure 1). Of particular interest in this table is a comparison between the results obtained from the most extensive *ab initio* Hartree–Fock calculation (calculation C) and the results using the wave function of the current study (calculation A). In spite of the vast difference in the complexity of the two wave functions, the resulting charge distribution is very similar in both cases, since both the magnitude and direction of $\mathbf{\mu}$ are very similar in both cases (4% difference in magnitude).

It should be noted that one of the procedures that is frequently employed to examine large molecular systems, extended Hückel theory (EHT), does not give rise to acceptable charge distributions, since the dipole moment values are a factor of 2 (at least) too large (10.23 D[14] using the KW geometry,[67] and 8.75 D[15] using a "standard" geometry[71]). Another comparison of interest concerns calculations by Robb and Csizmadia,[4] G, which also employed a small Gaussian basis set. As the dipole moment results especially indicate, particular care in the selection of basis orbitals must be exercised if small Gaussian basis sets are to be employed.

The dipole moment of formamide has been measured experimentally using a variety of techniques.[67,72–75] The microwave determination of the dipole moment of formamide has been reexamined in the current study, incorporating new measurements in the analysis.[76] Although it is possible to determine μ_a and μ_b from microwave Stark shift measurements, it is not possible in practice to determine μ_c from such measurements, because the absolute value of the coefficient (C_c) of μ_c^2 in the theoretical second-order Stark shift formula (Stark shift $= C_a\mu_a^2 + C_b\mu_b^2 + C_c\mu_c^2$) is very small relative to the absolute values of the coefficients of μ_a^2 and μ_b^2. A least-squares fit of the currently available Stark shift data (see Table III) results in $\mu_a = 3.62 \pm 0.01$ D and $\mu_b = 0.86 \pm 0.13$ D ($|\mu_a + \mu_b| = 3.72 \pm 0.04$ D), where the term $C_c\mu_c^2$ has been assumed to be negligible. Also displayed in Table III are the Stark shifts calculated assuming the values of μ_a and μ_b above. The dipole moment components determined here are nearly identical with the previously determined values[67] of $\mu_a = 3.62$ D and $\mu_b = 0.85$ D.

μ_a, μ_b, and μ_c have been calculated from the results of *ab initio* calculations C and A at the geometry of Costain and Dowling,[63] and the results in Debyes are $\mu_a = 3.86$, $\mu_b = 1.20$, $\mu_c = 0.46$ and $\mu_a = 3.78$, $\mu_b = 0.78$, $\mu_c = 0.59$, for calculations C and A, respectively. It is interesting that the results of the molecular fragment approach (calculation A), with its small basis set, are actually closer to the experimental results than the results of calculation C (with a much larger basis set). The results of these two *ab initio* calculations indicate that μ_c is of the order of 0.5 D. In an attempt to estimate a value for the total dipole moment of formamide, we have combined the experimental values, $\mu_a = 3.62 \pm 0.01$ D and $\mu_b = 0.86 \pm 0.13$ D, with the *ab initio* result, $\mu_c = 0.5 \pm 0.1$ D, which gives $\mu = 3.75 \pm 0.05$ D.

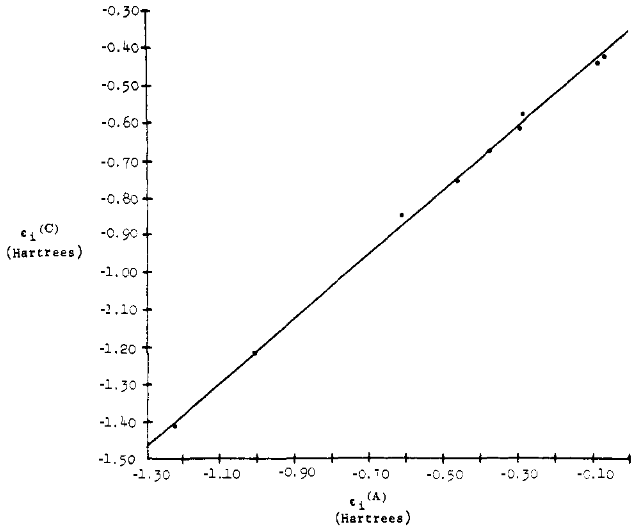

Figure 2. Comparison of formamide orbital energies. See Table II for the definition of the basis sets used.

Table III. Calculated and Measured Stark Shifts for Formamide in Units of MHz cm^4 esu^{-2} [a]

Investigators	Transition	M	Calcd[d]	Measured
Kurland and Wilson[b]	$0_{00} \rightarrow 1_{01}$	0	7.541	7.522
	$4_{14} \rightarrow 4_{13}$	4	1.670	1.630
	$3_{13} \rightarrow 4_{04}$	2	1.063	1.019
Steinmetz[c]	$3_{13} \rightarrow 4_{04}$	1	0.293	0.332
	$3_{13} \rightarrow 4_{04}$	2	1.063	1.10
	$3_{13} \rightarrow 4_{04}$	3	2.347	2.39

[a] Calculated Stark shifts are for $\mu_a = 3.62$ D and $\mu_b = 0.86$ D. [b] See ref 67. [c] See ref 76. [d] Current investigation.

C. Torsional Rotational Energy Barrier. The torsional rotational barrier, ΔE_{rot}, for formamide has usually been calculated quantum mechanically by holding all structural parameters fixed except the "twist" angle, τ, which is defined as follows[66]

$$\tau = [\omega(\text{H}_1\text{CNH}_2) + \omega(\text{OCNH}_3)]/2 \qquad (2)$$

where $|\omega(\text{H}_1\text{CNH}_2) - \omega(\text{OCNH}_3)| < \pi$. Table II presents the results of a number of quantum-mechanical calculations carried out in this manner, which are to be compared with values obtained from nmr measurements[77–79] ranging from 16.8 ± 1.0 to 21.3 ± 1.3 kcal/mol. The most precise of the measurements[79] indicates a value of 19.7 ± 0.2 to 19.2 ± 0.2 kcal/mol, depending upon the choice of solvent. Thus, the results

---

(71) H. A. Scheraga, *Advan. Phys. Org. Chem.*, 6, 103 (1968).

(72) (a) C. T. Zahn, *Phys. Z.*, 33, 525 (1932). (b) E. D. Hughes, C. G. Le Fèvre, and J. J. W. Le Fèvre, *J. Chem. Soc.*, 202 (1937).

(73) W. W. Bates and M. C. Hobbs, *J. Amer. Chem. Soc.*, 73, 2151 (1951).

(74) G. D. Burdun and P. B. Kantor, *Dokl. Akad. Nauk SSSR*, 67, 985 (1949).

(75) R. J. Kurland, *Bull. Amer. Phys. Soc.*, 1, 12 (1956).

(76) W. Steinmetz, private communication. The experimental measurements were made on a Hewlett-Packard spectrometer at Harvard University.

(77) H. Kamei, *Bull. Chem. Soc. Jap.*, 41, 2269 (1968).

(78) B. Sunners, L. H. Piette, and W. G. Schneider, *Can. J. Chem.*, 38, 681 (1960).

(79) T. Drakenberg and S. Forsen, *J. Phys. Chem.*, 74, 1 (1970).

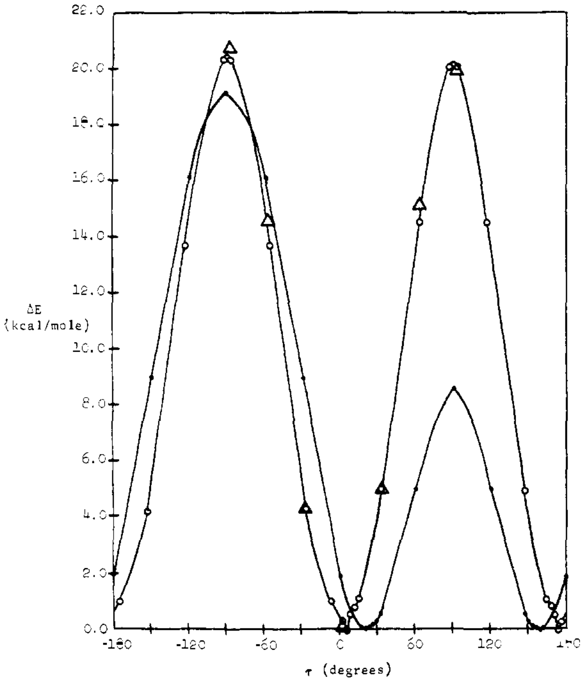

Figure 3. Calculated torsional rotational barrier of formamide. The various calculations are labeled as follows: (○) molecular fragment basis set, Costain–Dowling (CD) geometry; (●) molecular fragment basis set, CD geometry modified by imposing a tetrahedral geometry at N; and (△) basis set of calculation D, CD geometry (see Table II).

using the molecular fragment approach (19.72 kcal/mol) are again in excellent agreement with both the most extensive calculation (21.72 kcal/mol for calculation C) and the experimental results.

Several other calculations of ΔE_{rot} have been carried out using procedures other than the one just described that are worth noting. In particular, Christensen, *et al.*,[7] found ΔE_{rot} to be 19.40 kcal/mol by calculating the energy difference between a geometry optimized with respect to R(C,N) and R(C,O) at $\tau = 2.5°$, and a geometry optimized with respect to the same two distances at $\tau = 92.5°$ (all other structural parameters were fixed at the Costain–Dowling values). They also obtained $\Delta E_{\mathrm{rot}} = 20.42$ kcal/mol as the energy difference between a planar geometry with 120° angles at N, $R(\mathrm{C,N}) = 1.36$ Å, $R(\mathrm{C,O}) = 1.212$ Å, and $R(\mathrm{N,H}) = 0.9950$ Å, and a geometry identical to this geometry except that $\tau = 90°$ (all other structural parameters were fixed at the Costain–Dowling values). Also, Shaw and Reeves[18] have calculated (using CNDO/2) that the energy difference between a planar geometry at $\tau = 0°$ and a tetrahedral geometry at $\tau = 90°$ is 9.86 kcal/mol (the basic geometry was that of Kurland–Wilson).

In order to obtain information regarding the shape of the ΔE_{rot} *vs.* τ curve, several studies have been carried out. Figure 3 shows the results of the current molecular fragment calculations using the basis set of calculation A (see Tables I and II) and two different geometries, one of which is the Costain–Dowling experimental geometry, and the other is the Costain–Dowling geometry modified by imposing a tetrahedral geometry at N. In both these studies, as the geometry was varied with increasing values of τ from $\tau = 0°$, the molecule was reflected through the OCN plane for the specific values of $\tau = 90$ and $-90°$, and the calculations were continued at the reflected geometry. This resulted in a calculated barrier that is symmetrical about $\tau = 90°$, and another barrier symmetrical about $\tau = -90°$, and also assured that the nuclear geometries at $\tau = 2.5$ and 177.5° would be exactly the Costain–Dowling geometry. Also included in Figure 3 is that portion of the results of Christensen, *et al.*,[7] in which the nuclear geometry was varied in exactly the same fashion as in the current calculations. In their studies, the extensive basis set of calculation D (see Table II) was utilized. As Figure 3 indicates, the calculated points of Christensen, *et al.*, fall very nearly on the corresponding curve of the present study. Consequently, the wave functions obtained using the molecular fragment approach not only predict barriers whose magnitudes are very close to those of more extensive calculations, but also reproduce the shape of the barrier as well.

The data in Figure 3 concerning the modified Costain–Dowling geometry (T_d at N) indicate that there is a dependence of the position of the torsional rotational minima and barrier heights on the nuclear configuration at N. In particular, for a planar geometry at N (not shown), the minima appear at $\tau = 0$ and 180° and the barriers at $\tau = \pm 90°$ are of equal height. On the other hand, Figure 3 shows that the shallow pyramidal structure of Costain–Dowling produces minima at $\tau \sim 5.5$ and $\sim 174.5°$ and the barrier at $\tau = 90°$ is slightly lower than the barrier at $\tau = -90°$. The assumption of a tetrahedral geometry at N produces minima at $\tau \sim 21$ and $\sim 159°$ and the barrier at $\tau = 90°$ is 10.6 kcal/mol lower than the barrier at $\tau = -90°$.

Shaw and Reeves[18] have calculated (using CNDO/2) that, for a tetrahedral geometry at N (basic geometry of Kurland–Wilson), the barrier heights at ±90° are almost equal in height. However, this apparent conflict with the results using the molecular fragment approach may be due only to the difference in basic geometries utilized.

N-Methylacetamide (NMA)

The current investigation of NMA (see Figure 1) is the first *ab initio* study of the molecule. The nuclear geometry was taken from a compilation by Scheraga[71] of "standard" peptide bond angles and lengths that were derived from a review of the literature.

A. Rotational Studies. In order to investigate several aspects of the geometric structure of NMA, the total energy was considered to be a function of the dihedral angles ϕ, ψ, and ω (see Figure 1). The values of ω were restricted to 0 and 180°, corresponding to *cis*- and *trans*-NMA, respectively. The energy variation as a function of ϕ and ψ at $\omega = 0$ and 180° is given in Table IV. Also given in Table IV are the results of EHT and CNDO/2 calculations that were carried out by Yan, *et al.*,[15] who also used the geometric data from the previously mentioned compilation of peptide structural parameters by Scheraga.[71] Cis–trans energy differences calculated using all three methods are also given in Table IV.

Several other semiempirical investigations have been carried out on NMA. In particular, Govil[40] found (using EHT) that the trans form of NMA was 14 kcal/mol more stable than the cis form. This is in conflict with the results of Table IV, especially in view of the

Table IV. Rotational Studies[a] of NMA

	ϕ, deg				
ψ, deg	180	−150	−120	−90	Barrier height
			ω = 180° (trans)[b]		
180	0.69	0.35	0.00	0.35	0.69
	(0.30, 0.24)		(0.00, 0.00)		(0.30, 0.24)
−150	1.20	0.86	0.50	0.85	0.70
−120	1.82	1.46	1.10	1.46	0.72
	(0.48, 0.54)		(0.18, 0.30)		(0.30, 0.24)
−90	1.20	0.85	0.50	0.86	0.69
Barrier height	1.13	1.11	1.10	1.11	
	(0.18, 0.30)		(0.18, 0.30)		
			ω = 0° (cis)[c]		
180	0.00	0.68	1.07	0.68	
	(0.00, 0.00)		(3.78, 0.84)		
−150	1.48	8.20	8.80	2.13	
−120	2.76	9.70	20.23	9.70	
	(4.72, 1.28)		(15.11, 9.10)		
−90	1.48	2.13	8.80	8.20	
$E(180,180,0) - E(-120,180,180)$ =			3.66 kcal/mol (*ab initio*) 2.92 kcal/mol (EHT) −0.09 kcal/mol (CNDO/2)		

[a] All energies are reported in kcal/mol, and are relative to the lowest energy conformer. For each (ϕ,ψ) pair, the format of the energies is: *ab initio* results of current study (without parentheses), and (A,B), where A = extended Hückel results (see ref 15), B = CNDO/2 results (see ref 15). [b] The entries for ω = 180° correspond to $E(\phi,\psi,180) - E(-120,180,180)$, and are in units of kcal/mol. [c] The entries for ω = 0° correspond to $E(\phi,\psi,0) - E(180,180,0)$, and are in units of kcal/mol.

fact that the same basic geometry was utilized.[80] In another investigation, Maigret, *et al.*,[31] reported rotational barriers of 0.8 kcal/mol for ϕ and 1.0 kcal/mol for ψ using the PCILO method on *trans*-NMA. The position of the ψ minimum using PCILO is identical with that obtained by each of the three methods in Table IV, but the ϕ minimum was found at ϕ = 180°. This is the position at which the other procedures found a maximum. This difference in results is especially interesting in view of the fact that PCILO uses the same parameterization as CNDO/2. However, the difference in results may be due entirely to differences in peptide geometry utilized.[81] Finally, an additional study by Murthy, *et al.*,[33] found that the cis form is less stable than the trans form by 1.9 kcal/mol using EHT, and by 7.0 kcal/mol using CNDO/2, although the peptide geometry utilized in this study was not reported.

Available experimental evidence is consistent with the results of the molecular fragment approach presented in Table IV. For example, La Planche and Rogers[82] concluded from an nmr study of NMA that all of the NMA molecules were in the trans form. Similarly, Barker and Boudreaux[83] concluded in another nmr study that 97% of NMA molecules in water are in the trans form.

In summary, the trans form of NMA is favored by a few kilocalories per mole. In addition, quantum-mechanical calculations indicate that the ϕ and ψ rotations are essentially energetically uncoupled in the trans form. In the trans conformation, the results of all quantum-mechanical calculations (except the PCILO calculation) agree that the energetically favored conformation for the ϕ rotation ($\phi = -120°$) is the conformation in which one of the CH bonds eclipses the CN bond. Also, there is agreement that the energetically favored conformation for the ψ rotation (ψ = 180°) is the conformation in which one of the CH bonds eclipses the CO bond. The numerical values for the barriers to rotation differ slightly in the various quantum-mechanical calculations, but there is general agreement that the barriers are small.

B. Electronic Structure of NMA. The Hartree–Fock orbital energies that were obtained from the current calculations are given in Table V for *cis*- and *trans*-NMA, at the energy minima with respect to ϕ and ψ rotations. Although the "σ–π" notation employed is not strictly applicable due to the nonplanarity of the terminal CH_3 groups, it is still particularly informative as to the primary contribution to a given MO.

Table V. Orbital Energies[a] of *cis*- and *trans*-NMA at ϕ–ψ Conformational Energy Surface Minima

	Cis	Trans
ϵ_{21}	+0.5312 (π)	+0.5336 (π)
ϵ_{20}	−0.1115 (π)	−0.1094 (π)
ϵ_{19}	−0.1207 (σ)	−0.1172 (σ)
ϵ_{18}	−0.3078 (σ)	−0.3056 (π)
ϵ_{17}	−0.3091 (π)	−0.3115 (σ)
ϵ_{16}	−0.3883 (σ)	−0.3747 (σ)
ϵ_{15}	−0.4061 (σ)	−0.4038 (σ)
ϵ_{14}	−0.4174 (π)	−0.4408 (π)
ϵ_{13}	−0.4580 (σ)	−0.4441 (σ)
ϵ_{12}	−0.4844 (σ)	−0.4712 (π)
ϵ_{11}	−0.4871 (π)	−0.5081 (σ)
ϵ_{10}	−0.6407 (σ)	−0.6352 (σ)
ϵ_{9}	−0.8506 (σ)	−0.8294 (σ)
ϵ_{8}	−0.8836 (σ)	−0.9032 (σ)
ϵ_{7}	−1.1460 (σ)	−1.1426 (σ)
ϵ_{6}	−1.2623 (σ)	−1.2601 (σ)
ϵ_{5}	−9.1715 (σ)	−9.1780 (σ)
ϵ_{4}	−9.2363 (σ)	−9.2229 (σ)
ϵ_{3}	−9.3259 (σ)	−9.3281 (σ)
ϵ_{2}	−13.0605 (σ)	−13.0615 (σ)
ϵ_{1}	−17.0987 (σ)	−17.0963 (σ)

[a] See ref 69.

Previous experience[54–61] (see the section on Electronic Structure) indicates that the ordering of valence molecular orbitals is, in general, in excellent agreement with the results that are obtained at the Hartree–Fock limit. Thus, it is expected that the ordering of valence molecular orbitals in Table V is an accurate ordering, although the orbital energies are expected to be shifted to more positive values and the spacings slightly larger than those obtained from a near Hartree–Fock wave function.

The highest pair of occupied MO's and the lowest unoccupied MO in NMA are highly analogous to the corresponding MO's in formamide, since the contributions from the methyl groups are small. The next to highest occupied MO is a σ-nonbonding orbital that is primarily a p orbital on O, perpendicular to the CO bond. The highest occupied MO is the π-nonbonding

---

(80) G. Govil, private communication. Although not reported in ref 36, the "standard" peptide geometry of ref 63 was used, but the values of ϕ and ψ assumed in the calculations on *cis*- and *trans*-NMA were not specified.

(81) B. Maigret, private communication. Although not reported in ref 29, this geometry is almost identical with the "standard" geometry of ref 63, except for large deviations in bond angles at N. In particular, the "standard" bond angles are $\tau(C',N,H)$ = 123°, $\tau(C',N,C^{\alpha})$ = 123°, $\tau(C^{\alpha},N,H)$ = 114°, and the angles utilized by Maigret, *et al.* (ref 29) were $\tau(C',N,H)$ = 107°, $\tau(C',N,C^{\alpha})$ = 117°, $\tau(C^{\alpha},N,H)$ = 136°.

(82) L. A. La Planche and M. T. Rogers, *J. Amer. Chem. Soc.*, **86**, 337 (1964).

(83) R. H. Barker and G. J. Boudreaux, *Spectrochim. Acta*, *Part A*, **23**, 727 (1967).

Table VI. ϕ-ψ Energy[a] Surface Sample for 2-Formamidoacetamide (ω_1 = 180°)

	ϕ																								
ψ	−180	−165	−150	−135	−120	−105	−90	−75	−60	−45	−30	−15	0	15	30	45	60	75	90	105	120	135	150	165	180
180	0.00	0.29			2.79		5.37		7.64				271.60				7.64		5.37		2.79			0.29	0.00
165	0.93	0.60		1.51		3.76																		1.66	0.93
150																									
135		5.17		4.29		5.63																			
120	8.14				6.19				11.21				22.87				15.97				9.77				8.14
105		8.67		7.32		7.33						11.21	12.84	20.63											
90							7.22		10.27		21.92	17.67	11.09	12.32	15.48										
75							5.85	4.84	8.06			36.37	18.43	10.47	11.77	14.36	15.56								
60	8.34				5.45	5.52	5.08	3.34	9.83		98.01		41.70	15.72	10.54	12.40	14.05				15.14				8.34
45							5.16	5.45	20.07						12.88	10.78	12.10								
30							5.29		24.42																
15																									
0	31.79				9.64		6.80		11.02				3058.00				11.02		6.80		9.64				31.79
−15																									
−30																	24.42		5.29						
−45									12.10	10.78	12.88						20.07	5.45	5.16						
−60	8.34				15.14				14.05	12.40	10.54	15.72	41.70		98.01		9.83	3.34	5.08	5.52	5.45				8.34
−75									15.56	14.36	11.77	10.47	18.43	36.37			8.06	4.84	5.85						
−90											15.48	12.32	11.09	17.67	21.92		10.27		7.22						
−105												20.63	12.84	11.11						7.33		7.32		8.67	
−120	8.14				9.77				15.97				22.87				11.21				6.19				8.14
−135																				5.63		4.29		5.17	
−150																									
−165	0.93	1.66																		3.76		1.51		0.60	0.93
−180	0.00	0.29			2.79		5.37		7.64				271.60				7.64		5.37		2.79			0.29	0.00

[a] Energy units are kcal/mol, and the values given in the table are calculated from $E(\phi,\psi,180) - E(180,180,180)$.

MO, and is located primarily on O and N. The lowest unoccupied MO is the π-antibonding MO, and has primary contributions on C′, O, and N. In the lower filled π orbitals, the contributions from the methyl groups are substantial, and the π molecular orbitals are no longer localized in the OCN region.

C. Dipole Moment of *trans*-NMA. Using the current molecular fragment *ab initio* procedure, the dipole moment vector of *trans*-NMA at the ϕ–ψ rotational minima, $(\phi,\psi) = (-120,180)$, was found to have a magnitude of 3.73 D, and formed an angle of 54.2° with the CN bond (see Figure 1). For comparison, several experimental measurements are available. Mizushima, *et al.*,[84] found the dipole moment of NMA in carbon tetrachloride to be 4.39 D, and Meighan and Cole[85] calculated the dipole moment of NMA to be 3.82 D in benzene and 3.71 D in the gas phase, from dielectric constant measurements. The gas-phase value of 3.71 D is the value that is most appropriate for comparison with the quantum-mechanical results, and it is seen that the agreement is better than 1%.

Comparisons with semiempirical calculations are also possible. Yan, *et al.*,[15] have calculated the dipole moment of *trans*-NMA using EHT, CNDO/2, and CNDO/2 (corrected). They found values of 8.80, 2.40, and 4.12 D, respectively. For the CNDO/2 (corrected) calculation, the calculated dipole moment vector formed an angle of 44.3° with the CN bond.

2-Formamidoacetamide (Formylglycinamide)

The current *ab initio* study of 2-formamidoacetamide (2FA) is the first quantum-mechanical study of this molecule. The nuclear geometry was taken from a compilation of standard peptide structural parameters by Scheraga,[71] augmented by a choice of 1.102 Å for the carbonyl CH bond distance. The energy was studied as a function of the dihedral angles ϕ, ψ, and ω_1 (see Figure 1), although the values of ω_1 were restricted to ω_1 = 180 and 0°.

For the case of ω_1 = 180°, an investigation of the entire ϕ–ψ energy surface was carried out by first sampling the surface at a coarse grid of 60° intervals in ϕ and ψ. Then, for those regions in which "standard" structures[86] (fully extended chain, left- and right-handed helices, and parallel and antiparallel-chain pleated sheets) are found, the energy surface was sampled at 30° increments in ϕ and ψ. Finally, the regions in which minima were expected (from the results of other theoretical investigations and experimental studies of closely related molecules) were sampled at 15° increments in ϕ and ψ until the positions of the minima were obtained. All energy results for these investigations are displayed in Table VI, and the conformational energy surface minima (as revealed using the resolution stated above) are characterized in Table VII. Note that the minima correspond to the fully extended conformation, two equivalent seven-membered ring conformers, and two equivalent conformations ($C^{\delta+}\cdots O^{\delta-}$) which are stabilized by favorable electrostatic and orbital interac-

---

(84) S.-I. Mizushima, T. Simanouti, S. Nagakura, K. Kuratani, M. Tsuboi, H. Baba, and O. Fujioka, *J. Amer. Chem. Soc.*, **72**, 3490 (1950).

(85) R. M. Meighan and R. H. Cole, *J. Phys. Chem.*, **68**, 503 (1964).

(86) J. T. Edsall, P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Nemethy, G. N. Ramachandran, and H. A. Scheraga, *Biopolymers*, **4**, 121, 1149 (1966); *J. Biol. Chem.*, **241**, 1004, 4167 (1966); *J. Mol. Biol.*, **15**, 339 (1966); *ibid.*, **20**, 589 (1966).

Table VII. Characterization of Energy Minima from *Ab Initio* Calculations on 2-Formamidoacetamide

ϕ	ψ	ω_1	$E - E_0$, kcal/mol	μ, D	$d_{H\cdots O}$ or $d_{C\cdots O}$, Å	Description
180	180	180	0.00	3.93	2.18	Fully extended chain
75	−60	180	3.34	3.34	1.79	7-Membered ring $(\phi,\psi,\omega) = (-75,60,180)$ is identical
−15	−75	180	10.47	6.35	2.42	$C^{\delta+}\cdots O^{\delta-}$ $(\phi,\psi,\omega) = (15,75,180)$ is identical
180	180	0	4.03	6.57	1.79	Planar

tions between the carbonyl oxygen of the first amide unit and the carbonyl carbon of the second amide unit.

For $\omega_1 = 0°$ (a cis conformation for the first peptide unit), neither the seven-membered ring or the $C^{\delta+}\cdots O^{\delta-}$ conformers are possible, and so the energy search in ϕ and ψ was restricted to regions around $(\phi,\psi) = (180, 180)$, and $(\phi,\psi) = (0,180)$. The search in the latter region investigated the possibility of the formation of a six-membered ring stabilized by a weak hydrogen bond between the carbonyl hydrogen of the first peptide unit and the carbonyl oxygen of the second peptide unit. The results of these investigations are given in Table VIII, and the minimum energy conformer is characterized in Table VII, along with the minima for $\omega_1 = 180°$.

Table VIII. ϕ–ψ Energy Surface Sample for 2-Formamidoacetamide ($\omega_1 = 0°$)[a]

	ϕ					
ψ	−165	−30	0	+30	+165	±180
±180	4.31	14.23	33.78	14.23	4.31	4.03
+165	4.67				5.55	4.92
+150		12.30	21.38	15.48		
−150		15.48	21.38	12.30		
−165	5.55				4.67	4.92

[a] Energy units are kcal/mole, and calculated from $E(\phi,\psi,0) - E(180,180,180)$.

In order to characterize the minimum energy conformers further, the orbital energies for each of the minima are listed in Table IX. As mentioned earlier, the ordering of valence molecular orbitals is expected to be correct, although the values of the energies are expected to be shifted toward more positive values, and the orbital energy spacings slightly larger compared with the results that would be calculated from a near Hartree–Fock wave function.

Another aspect of the (ϕ,ψ) energy surface for $\omega_1 = 180°$ that should be noted is that the $C^{\delta+}\cdots O^{\delta-}$ minimum actually corresponds to a shallow "dip" in an energy "pass" that runs approximately along the line

$$\psi = -(\phi + 90) \qquad -45 \leq \phi \leq 15 \tag{3}$$

The largest energy variation in this range was 0.64 kcal/mol (see Table VI). The range of the "pass" may actually be larger than indicated in eq 3, but this possibility was not investigated in the current study. Since the conformational energy surface has inversion symmetry about $(\phi,\psi) = (0,0)$, it follows that there is another identical "pass" that is defined by replacing ϕ and $-\phi$ and ψ by $-\psi$ in eq 3.

Table IX. Orbital Energies[a] of 2-Formamidoacetamide at the ϕ–ψ Conformational Energy Surface Minima

	(ϕ, ψ, ω_1)			
	(180,180,180)	(75,−60, 180), (−75, 60,180)	(−15,−75, 180), (15, 75,180)	(180,180,0)
ϵ_{28}	+0.5314 (π)	+0.5151	+0.5454	+0.5267 (π)
ϵ_{27}	−0.0801 (π)	−0.0752	−0.0580	−0.0799 (π)
ϵ_{26}	−0.0864 (σ)	−0.0795	−0.0695	−0.0881 (σ)
ϵ_{25}	−0.1079 (π)	−0.1257	−0.0870	−0.1120 (π)
ϵ_{24}	−0.1269 (σ)	−0.1285	−0.0933	−0.1311 (σ)
ϵ_{23}	−0.2826 (π)	−0.2855	−0.2644	−0.2810 (σ)
ϵ_{22}	−0.2843 (σ)	−0.2895	−0.2798	−0.2873 (π)
ϵ_{21}	−0.3132 (σ)	−0.3226	−0.3012	−0.3170 (σ)
ϵ_{20}	−0.3250 (π)	−0.3327	−0.3176	−0.3288 (π)
ϵ_{19}	−0.3697 (σ)	−0.3790	−0.3759	−0.4004 (σ)
ϵ_{18}	−0.4255 (σ)	−0.4180	−0.3947	−0.4170 (σ)
ϵ_{17}	−0.4681 (π)	−0.4439	−0.4448	−0.4715 (σ)
ϵ_{16}	−0.4774 (σ)	−0.4798	−0.4699	−0.4819 (π)
ϵ_{15}	−0.5197 (σ)	−0.5485	−0.5369	−0.5049 (σ)
ϵ_{14}	−0.6198 (σ)	−0.6007	−0.5883	−0.6333 (σ)
ϵ_{13}	−0.6431 (σ)	−0.6681	−0.6575	−0.6576 (σ)
ϵ_{12}	−0.8603 (σ)	−0.8627	−0.8554	−0.8627 (σ)
ϵ_{11}	−1.0867 (σ)	−1.0568	−1.0476	−1.0919 (σ)
ϵ_{10}	−1.1086 (σ)	−1.1398	−1.1246	−1.1119 (σ)
ϵ_{9}	−1.2143 (σ)	−1.2273	−1.2020	−1.2147 (σ)
ϵ_{8}	−1.2593 (σ)	−1.2529	−1.2319	−1.2636 (σ)
ϵ_{7}	−9.1593 (σ)	−9.1917	−9.1712	−9.1574 (σ)
ϵ_{6}	−9.2346 (σ)	−9.2274	−9.2195	−9.2476 (σ)
ϵ_{5}	−9.3285 (σ)	−9.3115	−9.3027	−9.3377 (σ)
ϵ_{4}	−12.9183 (σ)	−12.8904	−12.8772	−12.9234 (σ)
ϵ_{3}	−13.0274 (σ)	−13.0594	−13.0415	−13.0249 (σ)
ϵ_{2}	−17.0564 (σ)	−17.0631	−17.0320	−17.0567 (σ)
ϵ_{1}	−17.0949 (σ)	−17.0858	−17.0782	−17.1005 (σ)

[a] See ref 69.

In larger polypeptides, it is a well known experimental fact that helical structures are frequently low energy minima. However, in 2FA the right- and left-handed α-helix regions,[86] $(\phi,\psi) \sim (-48,-57)$ and $(\phi,\psi) \sim (48,57)$, respectively, are energetically considerably above the lowest energy region, but lie very near or in an energy "pass" as just described. Of course, the experimental and calculated results are not in conflict, since the vertical hydrogen bonds that add greatly to the stability of the α-helical structures are not present in 2FA, and are not encountered until a polypeptide chain containing at least four amide units (–C(=O)NH–) is formed. Other structures that are frequently found to be energetically favored in larger polypeptides include the parallel-chain pleated sheet,[86] $(\phi,\psi) \sim (-119, 113)$, and the antiparallel-chain pleated sheet,[86] $(\phi,\psi) \sim (-142,145)$. As Table VI reveals, the energy of these structures is above that of the lowest energy conformer, although not as high as the α-helix conformation.

This also is not inconsistent with experimental data on larger polypeptides since, in a manner similar to the helical structures, hydrogen bonding is an important factor in stabilizing pleated sheet structures (in this case *via* interchain hydrogen bonding).

While there are no previous quantum-mechanical studies on 2FA for comparison, there have been studies[31,45,48] on two closely related dipeptides: *N*-acetyl-*N*′-methylglycineamide (I) and *N*-acetyl-*N*′-methylalanineamide (II). For I, PCILO studies[31] indicate minima for fully extended, seven-membered ring, and $C^{\delta+}\cdots O^{\delta-}$ structures, while EHT studies[48,51] find only fully extended (or closely related) structures as energy minima. For II, PCILO studies[31,45] indicate minima for fully extended, seven-membered ring, antiparallel-chain pleated sheet, and α-helix structures, while CNDO/2 studies[48] indicate minima for fully extended, seven-membered ring, and $C^{\delta+}\cdots O^{\delta-}$ structures. EHT studies[48] indicate a pleated sheet and several other structures [$(-100,-50)$, $(-120,40)$] as minima on the conformational surfaces.

Experimental studies are also available for comparison for these two molecules. For example, Mizushima, *et al.*,[87-90] concluded from ir studies that both I and II exist in two forms in nonpolar solvents. These forms are the fully extended form and a seven-membered ring form. Koyama and Shimanouchi,[91] reasoning from the results of an ir study and a normal coordinate analysis, concluded that I can exist in two forms in the crystalline state, one at $(\phi,\psi) \sim (-60,0)$ and the other at $(\phi,\psi) \sim (-60,180)$.

From the results of ir and nmr studies of methyl esters of alanine dipeptides and their *N*-methyl derivatives, Portnova, *et al.*,[92,93] concluded that approximately 70% of the molecules were in seven-membered ring conformations in both CCl_4 and the binary mixture $CCl_4 + CHCl_3$. In particular, the seven-membered ring in alanine dipeptides is found at $(\phi,\psi) = (60,-60)$ in nonpolar solvents. They concluded that there is a considerable increase in the population of conformers of L-alanine dipeptides at $\phi = -60°$ in polar solvents. Finally, Avignon, *et al.*,[94,95] and Marraud, *et al.*,[96] concluded from ir studies that molecule I exists in both fully extended and seven-membered ring conformations.

In summary, with the exception of the EHT results, there is general agreement among both the theoretical and experimental studies as to the importance of seven-membered ring structures as populated conformations of dipeptides. Apparently, the inability of EHT to account satisfactorily for hydrogen bond formation[17,34,48,53] is the cause of unsatisfactory results, which greatly limits its utility in studies of polypeptides. There is also general agreement as to the importance of the fully extended (or nearly fully extended) chain conformers, but disagreement among the quantum-mechanical studies as to the importance and position of the $C^{\delta+}\cdots O^{\delta-}$ minima. However, each of the $C^{\delta+}\cdots O^{\delta-}$ minima found in the *ab initio*, CNDO/2, and PCILO studies does fall in one of the energy "passes" mentioned earlier.

Conclusions

Reasoning from detailed comparisons between the results of the current calculations and the results of both experimental and other theoretical studies, we conclude that the *ab initio* molecular fragment approach satisfactorily describes several important aspects of peptides and polypeptides. In particular, cis–trans energy differences, torsional rotational energy barriers, conformational (ϕ,ψ) energy surfaces, charge distributions (as measured by the magnitude and direction of the dipole moment vector), and the ordering of the valence molecular orbitals appear to be well described using the molecular fragment approach. Consequently, it is anticipated that the *ab initio* molecular fragment approach will provide a valuable tool for both qualitative and quantitative characterizations of polypeptide systems.

Acknowledgments. The authors would like to express their appreciation to the University of Kansas for partial support of the computing required for this work. Several informative private communications from Professor E. B. Wilson concerning the microwave determination of the dipole moment of formamide are acknowledged, along with several helpful discussions with Professor M. D. Harmony on the same subject. Also, the authors would like to thank G. Govil and B. Maigret for supplying information concerning their respective quantum-mechanical studies of *N*-methylacetamide.

---

(87) S.-I. Mizushima, T. Shimanouchi, M. Tsuboi, T. Sugita, E. Kato, and E. Kondo, *J. Amer. Chem. Soc.*, 73, 1330 (1951).

(88) S.-I. Mizushima, T. Shimanouchi, M. Tsuboi, and R. Souda, *ibid.*, 74, 270 (1952).

(89) S.-I. Mizushima, T. Shimanouchi, M. Tsuboi, T. Sugita, K. Kurosaki, N. Mataga, and R. Souda, *ibid.*, 74, 4639 (1952).

(90) M. Tsuboi, T. Shimanouchi, and S.-I. Mizushima, *ibid.*, **81**, 1406 (1959).

(91) Y. Koyama and T. Shimanouchi, *Biopolymers*, **6**, 1037 (1968).

(92) S. L. Portnova, V. F. Bystrov, V. I. Tsetlin, V. T. Ivanov, and Y. A. Ovchinnikov, *Zh. Obshch. Khim.*, **38**, 428 (1968).

(93) V. F. Bystrov, S. L. Portnova, V. I. Tsetlin, V. T. Ivanov, and Y. A. Ovchinnov, *Tetrahedron*, **25**, 493 (1969).

(94) M. Avignon, P. V. Huong, and J. Lascombe, *Biopolymers*, **8**, 69 (1969).

(95) M. Avignon and P. V. Huong, *Biopolymers*, **9**, 427 (1970).

(96) M. Marraud, J. Neel, M. Avignon, and P. V. Huong, *J. Chim. Phys. Physiochim. Biol.*, 67, 959 (1970).