J. Am. Chem. Soc. 1992, 114, 5369–5377

Importance of Correlation-Gradient Geometry Optimization for Molecular Conformational Analyses

Regina F. Frey,[†] James Coffin,[‡] Susan Q. Newton,[§] Michael Ramek,[⊥] Vincent K. W. Cheng,[∥] F. A. Momany,[‡] and Lothar Schäfer*[,§]

Contribution from the IBM Corporation, 500 Maryville College Drive, St. Louis, Missouri 63141, Polygen Corp.-Molecular Simulations Inc., 200 Fifth Avenue, Waltham, Massachusetts 02254, Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, Arkansas 72701, Institut für Physikalische und Theoretische Chemie, Technische Universität Graz, A-8010, Graz, Austria, and Physics Department, University of Hong Kong, Hong Kong, Hong Kong. *Received November 7, 1991*

Abstract: The geometries of three conformations of glycine {Gly, forms **1** to **3**} and of two conformations of the diamide, *N*-formylalanineamide {Ala, forms C5 and C7eq}, were determined by MP2 gradient optimization at the MP2/6-311G** level in order to study the effects of correlation-gradient geometry refinement on the results of molecular conformational analyses. The MP2/6-311G** energy difference (ΔE) between **1** and **2** is 3.5 kJ/mol for MP2/6-311G** optimized geometries. This value increases by 1.9 kJ/mol to 5.4 kJ/mol when the MP2 calculations are performed with HF/6-311G** optimized geometries. The MP2/6-311G** ΔE for **1** and **3** is 2.9 kJ/mol for MP2/6-311G** optimized geometries, and increases to 4.8 kJ/mol for MP2/6-311G** calculations at HF/6-311G** geometries. In Ala, geometry optimization has the opposite effect; that is, ΔE for C5 and C7eq is smaller (by 1.4 kJ/mol) for the unoptimized structures. Thus, the effects of MP2 geometry optimization on MP2 energies are unpredictable and cannot be neglected, and single-point MP2 energies calculated at HF-optimized (MP2-unoptimized) geometries are not reliable. Similarly, RHF calculations performed at various levels (3-21G, 4-21G, 4-31G, 6-31G, 6-31G**, 6-311G**) demonstrate the difficulty of obtaining reliable estimates of the ΔE values and torsional angles of the most stable conformations of Gly and Ala. In the case of Gly, different levels of theory do not agree on the symmetry of the second minimum, the planar **2** or nonplanar **3**; the scatter in ΔE between **2** and **1** is 6 kJ/mol, and, most disturbingly, the most advanced HF calculations compare the worst with experimental results. The exact locations of the two most stable energy minima in the ϕ,ψ-torsional space of the diamide are significantly affected by MP2 geometry optimization (changes of >4° between MP2/6-311G** and HF/6-311G**, and changes of >15° between MP2/6-311G** and HF/3-21G). Concomitant with these effects, considerable shrinkage by electron correlation is observed for H$\cdots$X nonbonded interactions. Conformational changes in bond distances and angles also display significant variances with computational method, but the fluctuations do not preclude the identification of some clear and useful structural trends. Thus, even at simple levels of theory, such as HF/4-21G, conformational geometry maps of diamides capture essential structural trends which are not falsified by more advanced calculations. In contrast to this, RHF conformational energy maps are intrinsically inaccurate: the nonbonded interactions are incorrectly evaluated not only because dispersion effects are neglected but also because nonbonded distances are wrong due to errors in torsions.

Introduction

In recent years quantum chemistry has become an important tool in laboratory research. Applications of ab initio calculations in vibrational[1] and NMR[2] spectroscopy, in structural studies with microwave[3] and electron diffraction data,[4] and in countless problems of practical conformational analysis are convincing examples of the effective interplay between experiment and theory that is now pervading many branches of chemistry. In empirical molecular modeling (MM), results from ab initio calculations are used[5] increasingly to supplement the limited experimental data base that is available for parameter refinement. Thus, mediated by MM, ab initio theory now even impacts on matters that are at the very heart of chemistry, molecular synthesis and investigations of large molecules for which advanced calculations cannot be executed directly. A series of corrections of experimental conformational results prompted by calculations[6] demonstrates that the interplay is mutually beneficial.

When empirical decisions are influenced by computational results, questions of accuracy are of general interest. Many aspects of theory are constantly being tested by frequent use, but important questions remain open and must be addressed. There is great current interest in including electron correlation effects in the calculations of molecular properties of polyfunctional organic molecules. Typically such calculations are executed by using geometries which were optimized at the Hartree–Fock (HF) level, i.e., with a single determinant wave function without inclusion of electron correlation. This situation is similar to the late seventies, when "standard",[7] i.e., unoptimized, geometries were used in ab initio conformational analyses, even though effective HF gradients were well known and available in standard program packages.[8] The procedure was extremely approximate and often misleading, because molecular geometries are "local";[9] that is, they depend on where a molecule is in its conformational space, and deviations from "standard" geometry differ from one form to the

---

[†] IBM Corp.

[‡] Polygen Corp.

[§] University of Arkansas.

[⊥] Technische Universität Graz.

[∥] University of Hong Kong.

(1) (a) Fogarasi, G.; Pulay, P. In *Vibrational Spectra and Structure*; Durig, J. R., Ed.; Elsevier: Amsterdam, 1985; Vol. 14, p 215. (b) Hess, B. A., Jr.; Schaad, L. J.; Carsky, P.; Zahradnik, R. *Chem. Rev.* **1986**, *86*, 709. (c) Grev, R. S.; Schaefer, H. F., III *J. Am. Chem. Soc.* **1989**, *111*, 5687.

(2) (a) Kutzelnigg, W.; Fleischer, U.; Schindler, M. Ab Initio Calculation and Interpretation of NMR Chemical Shifts and Magnetic Susceptibilities by Means of the IGLO Method. In *NMR-Basic Principles and Progress*; Springer: Heidelberg, 1990. (b) Wolinsky, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251. (c) Schleyer, P. v. R.; Bühl, M. Abstract BO-02, Second World Congress of Theoretical Organic Chemists, Toronto, July 1990.

(3) (a) Norden, T. D.; Staley, S. W.; Taylor, W. H.; Harmony, M. D. *J. Am. Chem. Soc.* **1986**, *108*, 7912; **1987**, *109*, 7641. (b) Schäfer, L.; Siam, K.; Ewbank, J. D.; Caminati, W.; Fantoni, A. C. In *Modelling of Structures and Properties of Molecules*; Maksic, Z. B., Ed.; Horwood: Chichester, 1987; Chapter 4, p 79.

(4) (a) Geise, H. J.; Pyckhout, W. In *Stereochemical Applications of Gas Phase Electron Diffraction*; Hargittai, I., Hargittai, M., Eds.; VCH, 1988; Vol. A, Chapter 10, p 321. (b) Schäfer, L.; Ewbank, J. D.; Siam, K.; Chiu, N. S.; Sellers, H. L. *Ibid.*, Chapter 9, 301.

(5) (a) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551. (b) Momany, F. A.; Klimkowski, V. J.; Schäfer, L. *J. Comput. Chem.* **1990**, *11*, 654.

(6) Schäfer, L. *J. Mol. Struct.* **1991**, *230*, 5.

(7) Pople, J. A.; Gordon, M. *J. Am. Chem. Soc.* **1967**, *89*, 4253.

(8) (a) Pulay, P. *Mol. Phys.* **1969**, *17*, 197. (b) Pulay, P. In *Ab Initio Methods in Quantum Chemistry*; Lawley, K. P., Ed.; Wiley: New York, 1987; Vol. 2, p 241. (c) Schlegel, H. B.; Wolfe, S.; Bernardi, F. *J. Chem. Phys.* **1975**, *63*, 3632.

(9) Schäfer, L. *J. Mol. Struct.* **1983**, *100*, 51.

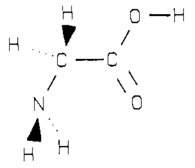
0002-7863/92/1514-5369$03.00/0

other. For ab initio calculations which include correlation effects in conformational analyses, the same problem must be explored: to what extent do calculated conformational properties depend on geometry optimization at the correlated level?

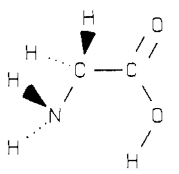
Variational methods (CI, MCSCF) with gradient optimization would be desirable to rigorously explore this question. However, for systems which are large enough to be interesting to chemists, specifically polyfunctional organic molecules with several dozen degrees of freedom, this would require very large CI expansions and forbidding computing expenses. For this reason perturbative methods are now frequently applied, such as second-order Møller–Plesset (MP2) calculations.[10] MP2 procedures are not a true substitute for CI, but, in combination with extended basis sets, they are generally considered to include a large part of the electron correlation energy that is neglected in HF theory.[10] The formulation of effective MP2 gradients[11] and their inclusion in generally applicable program packages, such as GAUSSIAN 90[12] or CADPAC,[13] make it now possible to examine how important MP2 geometry optimization is for calculated conformational properties, such as energy, geometry changes, nonbonded interactions, positions of energy minima, and the conformation dependence of atomic charges.

First proposals[14] to employ HF geometry trends as direct constraints of experimental data analyses were based on the hypothesis that errors in HF optimized relative bond distances and angles in organic compounds were sufficiently small to make this kind of procedure meaningful. The hypothesis was confirmed by many successful cases.[6] Nevertheless, to what exact extent relative bond distances and angles are modified by correlation is still an interesting question. In the related area of nonbonded interactions, the ubiquitous H···N and H···O interactions may be particularly sensitive to dispersion effects, and it is important to examine how correlation-gradient optimized hydrogen-bonded structures differ from HF geometries.

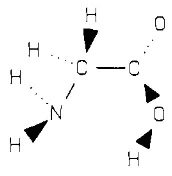
In order to study the problems identified above, we have determined the geometries of three forms of glycine (Gly, **1** to **3**) and of two forms of a diamide, the *N*-formylalanineamide (Ala, C5 and C7eq), using MP2 gradient[11] optimization. The conformations selected contain H-bonded structures, and their conformational energies, structural trends, and exact locations in torsional space are revealing to compare at the MP2 and HF levels of theory. Thus, the two systems are meaningful models for various classes of compounds, including larger ones. Since more than 300 h of supercomputer CPU time (IBM 3090) are needed for this project, the MP2 optimizations of Gly or Ala are perhaps close to the practical limits of current computational possibilities and larger systems are difficult to consider directly.



(1)



(2)



(3)

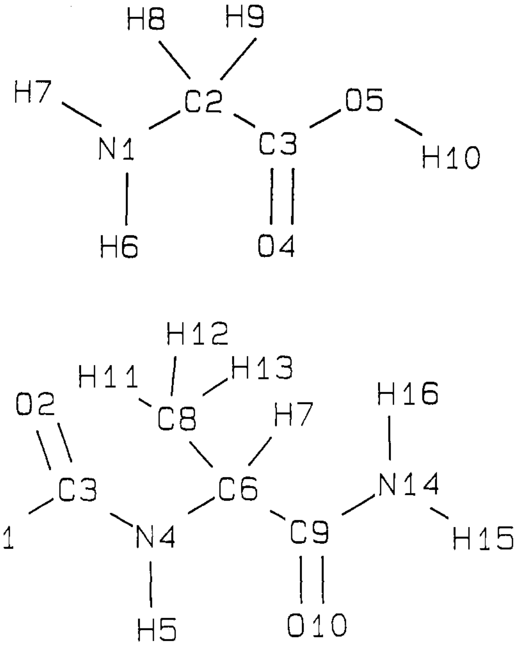


Figure 1. Atom numbering for glycine and *N*-formylalanineamide.

In the case of glycine, a particular point of interest concerns an apparent conflict[15] between the symmetry of the cyclic form suggested by the calculations and that derived from the microwave data.[16,17] The nonplanar form **3** is favored by the MP2 calculations, but the planar form **2** was observed in the microwave spectra.

In the case of Ala, a particular point of interest concerns conformational geometry trends which were established by the first HF gradient geometry optimizations of the homologous model dipeptides, the *N*-acetyl *N′*-methyl derivatives of glycine and alanine some time ago.[18,19] In these studies,[18,19] bond distances and bond angles were optimized at selected torsional angles, not necessarily at energy minima, in order to determine how the structures change from one characteristic region to another in dipeptide ϕ(N–C(α)), ψ(C(α)–C′) torsional space. Even though this intent of the optimized structures was clearly stated (e.g.,[19] "these geometries are not minimum energy conformations..."), for some reason it is at times totally disregarded. Head-Gordon et al.,[20] for example, recently falsely quoted these structures as "minima" and then spent a great deal of effort to show that some of them, indeed, were no minima. The MP2 geometry optimizations presented below confirm that our interest in differences between bond distances and angles in characteristic regions, i.e., not necessarily at the exact minima, of dipeptide ϕ,ψ-space was meaningful. Significant structural trends were found already 10 years ago which are not falsified by the more advanced calculations which are possible now. At the same time, the results will show that single-point MP2 energies obtained at HF geometries, like the MP2/6-31+G**//HF/6-31+G* energies of Ala by Head-Gordon et al.[20], are unreliable, and that HF conformational energy maps of polyfunctional organic molecules are intrinsically inaccurate.

---

(10) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. In *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986; Chapter 6, p 133.

(11) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem. Symp.* **1979**, *13*, 225.

(12) Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A. *GAUSSIAN 90*; Gaussian Inc.: Pittsburgh, PA.

(13) Amos, R. D.; Rice, J. E. *CADPAC: The Cambridge Analytical Derivative Package*; Cambridge University Press: Cambridge, England, 1987; Issue 4.0.

(14) (a) Chiu, N. S.; Ewbank, J. D.; Askari, M.; Schäfer, L. *J. Mol. Struct.* **1979**, *54*, 185. (b) Chiu, N. S.; Sellers, H. L.; Schäfer, L.; Kohata, K. *J. Am. Chem. Soc.* **1979**, *101*, 5883. (c) van Hemelrijk, D.; Van Den Enden, L.; Geise, H. J.; Sellers, H. L.; Schäfer, L. *J. Am. Chem. Soc.* **1980**, *102*, 2189. (d) Klimkowski, V. J.; Ewbank, J. D.; van Alsenoy, C.; Scarsdale, J. N.; Schäfer, L. *J. Am. Chem. Soc.* **1982**, *104*, 1476.

(15) (a) Ramek, M.; Cheng, V. K. W.; Frey, R. F.; Newton, S. Q.; Schäfer, L. *J. Mol. Struct.* **1991**, *235*, 1. (b) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 7917.

(16) Brown, R. D.; Godfrey, P. D.; Storey, J. W. V.; Bassez, M. P. *J. Chem. Soc., Chem. Commun.* **1978**, 547.

(17) (a) Suenram, R. D.; Lovas, F. J. *J. Mol. Spectrosc.* **1978**, *72*, 372. (b) Suenram, R. D.; Lovas, F. J. *J. Am. Chem. Soc.* **1980**, *102*, 7180.

(18) (a) Schäfer, L.; van Alsenoy, C.; Scarsdale, J. N. *J. Chem. Phys.* **1982**, *76*, 1439. (b) Klimkowski, V. J.; Schäfer, L.; Momany, F. A.; van Alsenoy, C. *J. Mol. Struct.* **1985**, *124*, 143. (c) Schäfer, L.; Klimkowski, V. J.; Momany, F. A.; Chuman, H.; van Alsenoy, C. *Biopolymers* **1984**, *23*, 2335.

(19) Scarsdale, J. N.; van Alsenoy, C.; Klimkowski, V. J.; Schäfer, L.; Momany, F. A. *J. Am. Chem. Soc.* **1983**, *105*, 3438.

(20) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C.; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989.

Table I. Comparison of the Ab Initio Optimized Geometries[a] (Bond Lengths, Å; Angles, deg) of Several Conformations of Glycine Determined at the HF/4-21G, HF/6311G**, and MP2/6-311G** Levels

	HF/4-21G			HF/6-311G**			MP2/6-311G**		
	1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
				Bond Lengths					
N1—C2	1.4554	1.4744	1.4741	1.4393	1.4568	1.4546	1.4498	1.4677	1.4657
N1—H6	1.0010	1.0001	0.9998	1.0001	0.9975	0.9978	0.0152	1.0121	1.0122
N1—H7	1.0010	1.0001	1.0010	1.0001	0.9975	0.9997	1.0152	1.0121	1.0139
C2—C3	1.5149	1.5362	1.5353	1.5148	1.5271	1.5246	1.5172	1.5346	1.5321
C2—H8	1.0811	1.0808	1.0807	1.0853	1.0837	1.0841	1.0940	1.0926	1.0932
C2—H9	1.0811	1.0808	1.0814	1.0853	1.0837	1.0849	1.0940	1.0926	1.0933
C3=O4	1.2033	1.2024	1.2021	1.1815	1.1785	1.1778	1.2076	1.2044	1.2040
C3—O5	1.3645	1.3452	1.3460	1.3291	1.3168	1.3193	1.3534	1.3373	1.3396
O5—H10	0.9657	0.9748	0.9743	0.9459	0.9500	0.9490	0.9672	0.9819	0.9801
				Bond Angles					
N1—C2—C3	113.42	110.26	110.04	115.19	112.58	111.98	115.02	110.90	110.50
N1—C2—H8	110.23	111.79	113.52	110.19	112.08	114.71	109.99	112.34	114.91
N1—C2—H9	110.23	111.79	110.08	110.19	112.08	109.22	109.99	112.34	109.50
C2—N1—H6	113.28	114.48	114.53	110.43	112.43	111.96	108.54	111.43	111.31
C2—N1—H7	113.28	114.48	114.11	110.43	112.43	111.78	108.54	111.43	110.85
C2—C3=O4	126.63	122.47	122.64	125.49	121.60	122.12	125.33	122.26	122.73
C2—C3=O5	110.52	113.67	113.58	111.67	115.19	114.94	111.25	113.35	113.11
C3—C2—H8	107.95	107.71	108.12	107.41	106.54	107.49	107.67	106.98	107.92
C3—C2—H9	107.95	107.71	107.31	107.41	106.54	105.81	107.67	106.98	106.18
C3—O5—H10	112.10	108.40	108.56	108.59	107.99	108.46	105.32	102.94	103.29
O4=C3—O5	122.85	123.86	123.77	122.84	123.21	122.92	123.42	124.40	124.14
H6—N1—H7	110.62	111.41	111.29	106.32	108.19	107.84	104.80	107.44	107.32
H8—C2—H9	106.80	107.38	107.54	106.01	106.60	107.17	106.11	106.98	107.42
				Torsional Angles					
H6-N1-C2-C3	−63.52	−114.83	130.20	−58.65	−118.80	148.67	−56.69	−120.00	144.39
H6-N1-C2-H8	57.65	4.96	−108.48	63.03	1.30	−88.48	65.05	−0.34	−93.20
H6-N1-C2-H9	175.31	125.38	12.12	179.67	121.11	31.82	−178.43	120.34	27.78
H7-N1-C2-C3	63.52	114.79	−99.90	58.65	118.80	−90.21	56.69	120.00	−96.23
H7-N1-C2-H8	−175.31	−125.42	21.42	−179.67	−121.11	32.63	178.43	−120.34	26.18
H7-N1-C2-H9	−57.65	−5.01	142.02	−63.03	−1.30	152.94	−65.05	0.34	147.16
N1-C2-C3-O4	0.0	179.95	171.47	0.0	180.0	163.61	0.0	180.0	166.45
N1-C2-C3-O5	180.0	−0.05	−9.59	180.0	0.0	−18.08	180.0	0.0	−15.21
H8-C2-C3-O4	−122.45	57.72	46.97	−123.17	56.76	36.77	−122.99	57.19	40.04
H8-C2-C3-O5	57.55	−122.28	−134.08	56.83	−123.24	−144.93	57.01	−122.81	−141.63
H9-C2-C3-O4	122.45	−57.82	−68.76	123.17	−56.76	−77.49	122.99	−57.18	−74.89
H9-C2-C3-O5	−57.55	122.18	110.19	−56.83	123.24	100.81	−57.01	122.82	103.44
C2-C3-O5-H10	180.0	−0.01	2.33	180.0	0.0	3.31	180.0	0.0	3.56
O4-C3-O5-H10	0.0	179.99	−178.74	0.0	180.0	−178.40	0.0	180.0	−178.13

[a] All structures were fully optimized at the indicated levels. Largest residual forces in the optimized geometries were <0.001 aJ/Å for bond stretches, and <0.001 aJ/rad for angle bends.

Computational Procedures

MP2 gradient geometry optimizations were executed with CADPAC[13] on the IBM 3090 at the IBM computing center in Dallas. HF gradient calculations were executed with the same program, but also with GAMESS[21] at various institutions, including the EDV-Zentrum of the University of Vienna (IBM 3090), with an older version of GAUSSIAN,[22] and with an updated version of TEXAS[23] obtained from P. Pulay, using one of the IBM RISC/6000 systems at the University of Arkansas. The standard basis sets defined in the quoted programs were employed; the 4-21G basis was taken from ref 24.

Gradient geometry optimizations for Gly and Ala included HF/3-21G, HF/4-21G, HF/4-31G, HF/6-31G, HF/6-31G**, HF/6-311G**, and MP2/6-311G**. Some of the results of Gly were taken from ref 15. Vibrational frequencies were used without scaling.

Results and Discussion

1 is the stretched form of Gly and the global energy minimum. It is characterized by a bifurcated nonbonded interaction between the NH_2 hydrogens and the carbonyl oxygen. **2** is the cyclic form of Gly, the heavy atom framework is planar, only the methylene and amino hydrogen atoms are symmetrically out of the heavy atom plane, the C–H and N–H bonds are eclipsed, and there is a characteristic hydrogen bond between H10 and N1. (All atom numbering is given in Figure 1). In **3** the same H···N interaction exists as in **2**, but the heavy atom framework is not planar (N—C—C=O angle not equal to 180°) and the NH_2 group is twisted away from the arrangement in which it is eclipsed with CH_2.

The forms of Ala are characterized by the ϕ(N–C_α, C3–N4–C6–C9) and ψ(C_α–C′, N4–C6–C9–N14) torsional angles. In C7eq, ϕ and ψ are in the vicinity of (−80°, +70°); in C5 (−160°, +170°). C7eq is the global energy minimum of Ala. It is characterized by a seven-membered ring closed by a hydrogen bond between O2 and H15 in O2=C3—N4—C6—C9—N14—H15. C5 is the next stable minimum on the ϕ,ψ surface. It is characterized by a five-membered ring closed by a nonbonded interaction between H5 and O10 in H5–N4–C6–C9–O10.

The HF/4-21G, HF/6-311G**, and MP2/6-311G** gradient-optimized geometries of Gly and Ala are listed in Tables I and II. From the numerous calculations performed for these systems, the HF/4-21G geometries were selected for this presentation because they are a good representative of HF/double-zeta nonpolarization calculations. The HF/6-311G** results are good representatives for computations using basis sets with polarization functions.

Calculated energies and the positions of the conformational energy minima of Gly and Ala are listed in Tables III and IV. Conformational differences between structural parameters ob-

---

(21) Schmidt, M. W.; Boatz, J. A.; Baldridge, K. K.; Koseki, S.; Gordon, M. S.; Elbert, S. T.; Lam, B. *QCPE Bull.* **1987**, *7*, 115.

(22) Frisch, M. J.; Binkley, J. S.; Schlegel, H. B.; Raghavachari, K.; Martin, R.; Stewart, J. J. P.; Bobrowicz, F.; DeFrees, D. J.; Seeger, R.; Whiteside, R.; Fox, D. J.; Fluder, E.; Pople, J. A. *GAUSSIAN 86, Release C*; Carnegie Mellon University: Pittsburgh, PA, 1986.

(23) Pulay, P. *Theor. Chim. Acta* **1979**, *50*, 299.

(24) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550.

Table II. Ab Initio Optimized Geometries (Bond Lengths, Å; Angles, deg)[a] of the C5 and C7eq Conformations of *N*-Formylalanineamide

	C5			C7		
	HF/4-21G	HF/6-311G**	MP2/6-311G**	HF/4-21G	HF/6-311G**	MP2/6-311G**
Bond Lengths						
H1—C3	1.0826	1.0925	1.1046	1.0816	1.0915	1.1031
O2=C3	1.2227	1.1923	1.2200	1.2271	1.1944	1.2234
C3—N4	1.3455	1.3433	1.3565	1.3447	1.3435	1.3554
N4—H5	0.9974	0.9947	1.0114	0.9952	0.9938	1.0098
N4—C6	1.4545	1.4437	1.4425	1.4747	1.4592	1.4596
C6—H7	1.0817	1.0843	0.0963	1.0800	1.0821	1.0942
C6—C8	1.5444	1.5348	1.5325	1.5272	1.5203	1.5161
C6—C9	1.5272	1.5256	1.5254	1.5401	1.5361	1.5364
C8—H11	1.0823	1.0853	1.0939	1.0798	1.0831	1.0914
C8—H12	1.0789	1.0823	1.0909	1.0814	1.0848	1.0924
C8—H13	1.0813	1.0851	1.0930	1.0808	1.0838	1.0921
C9=O10	1.2232	1.1950	1.2209	1.2211	1.1937	1.2199
C9—N14	1.3524	1.3495	1.3628	1.3503	1.3501	1.3643
N14—H15	0.9922	0.9910	1.0059	0.9980	0.9950	1.0130
N14—H16	0.9946	0.9933	1.0081	0.9942	0.9939	1.0096
Bond Angles						
H1—C3=O2	122.40	122.18	123.17	121.90	121.95	112.68
H1—C3—N4	113.32	112.69	112.40	113.84	113.00	112.80
O2=C3—N4	124.29	125.13	124.42	124.26	125.05	124.52
C3—N4—H5	123.00	120.43	121.79	119.92	118.35	118.96
C3—N4—C6	121.23	122.40	121.26	122.14	123.26	122.06
N4—C6—H7	109.22	108.37	108.70	106.73	106.65	106.40
N4—C6—C8	111.48	112.19	111.55	110.00	110.25	110.09
N4—C6—C9	106.31	107.60	107.19	110.03	109.62	109.21
H5—N4—C6	115.72	116.96	116.59	117.93	118.26	118.59
C6—C8—H11	111.36	111.32	111.55	109.87	110.24	110.03
C6—C8—H12	108.18	109.14	108.88	110.15	109.89	110.22
C6—C8—H13	109.78	110.25	109.55	110.22	110.84	109.97
C6—C9=O10	121.67	121.89	122.02	121.85	122.04	122.07
C6—C9—N14	115.08	115.26	114.74	113.68	114.35	113.37
H7—C6—C8	109.38	108.86	108.94	111.08	110.09	110.79
H7—C6—C9	109.55	109.22	109.79	108.17	108.48	109.04
C8—C6—C9	110.83	110.54	110.64	110.74	111.61	111.18
C9—N14—H15	122.61	122.20	120.57	121.04	120.42	117.58
C9—N14—H16	118.64	118.30	117.13	118.66	117.51	115.99
O10=C9—N14	123.24	122.83	123.24	124.47	123.61	124.55
H11—C8—H12	109.00	108.53	108.67	109.19	108.85	109.11
H11—C8—H13	109.03	108.56	108.85	107.46	107.43	107.43
H12—C8—H13	109.46	109.00	109.30	109.91	109.53	110.03
H15—N14—H16	118.74	118.94	118.10	120.23	118.91	118.10
Torsions						
H1—C3—N4—H5	1.27	3.22	4.41	4.07	5.82	8.04
H1—C3—N4—C6	178.77	177.74	177.24	−175.16	−178.42	−179.23
O2=C3—N4—H5	−178.62	−176.82	−175.82	−175.87	−174.46	−172.93
O2=C3—N4—C6	−1.11	−2.29	−2.99	4.91	1.30	−0.20
C3—N4—C6—H7	−48.47	−38.77	−41.14	32.47	31.79	36.06
C3—N4—C6—C8	72.52	81.45	78.99	153.08	151.30	156.16
C3—N4—C6—C9	−166.59	−156.75	−159.75	−84.66	−85.45	−81.50
H5—N4—C6—H7	129.21	135.94	132.05	−146.77	−152.45	−151.19
H5—N4—C6—C8	−109.80	−103.84	−107.82	−26.16	−32.94	−31.09
H5—N4—C6—C9	11.09	17.95	13.44	96.10	90.31	91.25
N4—C6—C8—H11	−179.57	−178.27	−178.50	−177.88	−178.24	−177.15
N4—C6—C8—H12	−59.81	−58.50	−58.58	−57.56	−58.27	−56.78
N4—C6—C8—H13	59.59	61.19	60.91	63.89	62.94	64.70
H7—C6—C8—H11	−58.68	−58.34	−58.51	−59.94	−60.84	−59.75
H7—C6—C8—H12	61.08	61.44	61.41	60.39	59.14	60.63
H7—C6—C8—H13	−179.52	−178.88	−179.10	−178.17	−179.66	−177.89
C9—C6—C8—H11	62.21	61.62	62.27	60.28	59.68	61.68
C9—C6—C8—H12	−178.02	−178.60	−177.81	−179.39	179.66	−177.95
C9—C6—C8—H13	−58.63	−58.91	−58.32	−57.95	−59.13	−56.47
N4—C6—C9=O10	−11.10	−19.69	−18.00	−113.35	−101.71	−97.96
N4—C6—C9—N14	169.95	162.24	162.08	66.83	78.34	82.46
H7—C6—C9=O10	−129.00	−137.12	−135.90	130.42	142.19	146.16
H7—C6—C9—N14	52.05	44.81	44.18	−49.41	−37.75	−33.42
C8—C6—C9=O10	110.21	103.13	103.83	8.47	20.73	23.72
C8—C6—C9—N14	−68.94	−74.94	−76.09	−171.35	−159.21	−155.86
C6—C9—N14—H15	−3.10	−8.68	10.83	−2.07	−11.79	−19.15
C6—C9—N14—H16	178.23	−179.99	167.37	−178.91	−171.50	−166.92
O10=C9—N14—H15	177.96	173.27	−169.09	178.12	168.27	161.29
O10=C9—N14—H16	−0.71	1.96	−12.55	1.27	8.56	13.51

[a] All structures were optimized without any constraint at the indicated levels. Largest residual forces in the optimized geometries were <0.001 aJ/Å in bond stretches and <0.001 aJ/rad in angle bends.

Table III. Calculated Energies for Conformations **1** to **3** of Glycine

	1	**2**			**3**		
basis[a]	−*E*(tot)[b]	*E*(rel)[c]	*E*′(rel)[d]	ν(N–C–C=O)[e]	*E*(rel)[f]	*E*′(rel)[g]	N–C–C=O
HF/3-21G	281.247498	7.5	8.9	63.9i	7.3	9.4	169.5
HF/4-21G	282.158115	9.1	10.5	46.2i	9.0	10.9	171.5
HF/4-31G	282.409311	11.9	13.3	16.7			
HF/6-31G	282.698013	12.4	13.8	27.1			
HF/6-31G**	282.848341	13.6	14.1	105.8i	12.6	14.2	163.8
HF/6-311G**	282.917964	13.5	13.9	106.1i	12.4	13.9	163.6
MP2/6-311G**//HF/6-311G**[h]	283.883723	5.4			4.8		
MP2/6-311G**[i]	283.886760	3.5			2.9		166.5

[a] Standard basis sets were used as described in the text. Results are from this study or from ref 15. [b] Total energies, hartrees. [c] RHF energy differences, **2** − **1**, kJ/mol. All N-C-C=O angles are 180°. [d] (RHF + ZPE) energy differences, **2** − **1**, kJ/mol. [e] Torsional mode, in 1/cm; imaginary frequencies signified by "i". [f] RHF energy differences, **3** − **1**, kJ/mol. [g] (RHF + ZPE) energy differences, **3** − **1**, kJ/mol. [h] MP2 energies calculated at the HF optimized geometries. [i] MP2 energies for fully optimized MP2 geometries.

tained at different computational levels are given in Tables V (for Gly) and VI and VII (for Ala). The basis set and correlation dependence of the hydrogen-bonded structures are listed in Table VIII, and of Mulliken populations in Tables IX and X.

Importance of MP2 Gradient Optimization for Conformational Energies

In 1978 the ab initio study of Gly[25] caused the first incident in which an ab initio geometry predicted the existence of a hidden molecular conformational state, and at the same time effectively guided new experiments which led to the discovery of that state.[17b,26] In the first microwave (mw) investigation of Gly, Brown et al.[16] assigned the observed spectrum to **2** and claimed that this observation was in conflict with the ab initio study by Vishveshwara et al.,[27] in which **1** was identified as the global energy minimum on the basis of qualitative "standard geometry" HF/4-31G calculations.

In contrast to this work, Sellers et al.[25] and Schäfer et al.[26] optimized the geometries of several forms of Gly at the HF/4-21G level and, in agreement with an independent mw study by Suenram et al.,[17a] concluded that the exclusive observation of **2** in the mw spectrum did not allow any inference regarding its energy relative to **1**. The calculated dipole moment of **2** was much larger than that of **1**. Since intensities of transitions in the mw region are proportional to the square of the dipole moment components, it was possible that the mw spectrum of an equilibrium mixture of Gly was dominated by the spectrum of the less populated conformer **2**, while the most populated one, **1**, remained undetected.

The HF/4-21G geometry of **2** reproduced the experimental rotational constants of Gly like an r_0 structure refined from the data. Therefore, the spectroscopic search for **1** was continued,[17b,26] with spectral predictions based on its 4-21G structure, to guide the experiments. Indeed, these efforts led to the discovery of **1** and the determination of its energy, 5.9 ± 1.8 kJ/mol below that of **2**.[17b,26]

When the alleged contrast between the ab initio work and the mw investigations of Gly had been dissolved in this way, it seemed that the essential conformational properties of this molecule had been determined. A large number of conformational calculations exist for Gly,[25-28] both ab initio and semiempirical. Nevertheless, it was recently pointed out[15] that a particular aspect of Gly potential energy, namely, the exact shape of the energy surface around **2** and **3**, has not been discussed exhaustively, and, in fact, contradictory results are obtained with HF calculations using different basis sets.

It is seen from Table III that the HF energy differences between **1** and **2** are rather uncertain, ranging from 7.5 to 13.6 kJ/mol for the calculations shown. In addition, they are contradictory regarding the exact location of the minimum of the cyclic form, at **2** or **3**. In some calculations, **2** is the energy minimum. In others, like HF/3-21G, HF/4-21G, HF/6-31G**, and HF/6-311G**, the nonplanar cyclic form **3** is the energy minimum while the planar **2** is a saddle point, characterized by an imaginary vibrational frequency. Additional HF results, equally contradictory, were discussed in a recent study.[15]

It is a disturbing aspect of the HF energies that the more advanced the calculations are, the more they deviate from the experimental value, 27% at HF/3-21G and 129% at HF/6-311G**. Thus, enlarging the basis set does not necessarily lead to convergence of HF energies toward experimental values, as is often implied. In order to determine whether vibrational effects are contributing to this trend, zero-point energies (ZPE) of **1**, **2**, and **3** were calculated at various HF levels and energies corrected for ZPE. It is seen from Table III that ZPE corrected energy differences between **1** and **2** are in even poorer agreement with experiment than the uncorrected values.

In view of the contradictions in the HF results in this case, the addition of electron correlation is necessary. It is seen from Table III that **3** is the MP2 energy minimum of the cyclic form, approximately 0.5 kJ/mol more stable than **2**, and the energy difference between **2** and **1** is smaller than the experimental value. For the sake of this study the change in energy by geometry optimization is most important to consider. MP2/6-311G** energy differences calculated for MP2/6-311G** gradient optimized geometries are 3.5 kJ/mol for **2** − **1**, and 2.9 kJ/mol) for **3** − **1**. *There is a 50 to 60% shift in these values when MP2/6-311G** energies are calculated for unoptimized geometries, i.e., for the HF/6-311G** structures.* Thus, it is seen that the effect of correlation-gradient geometry optimization can be considerable on calculated energy differences of simple polyfunctional organic molecules. The generally accepted procedure, of calculating correlated energies at HF geometries, is potentially highly inaccurate. The results are of the same limited utility as HF energies of "standard" geometries.[7]

It is a disappointing aspect of the MP2 calculations that the optimized energy of **2**, 3.5 kJ/mol, is not closer to the experimental value (5.9 ± 1.8 kJ/mol) than the HF/3-21G result, and that the best agreement is obtained for the unoptimized value, 5.4 kJ/mol, which is conceptually the most questionable one. However, we consider the latter agreement as fortuitous. Furthermore, the MP2 energies, like the others, may be increased, i.e., get closer to the experimental value, by zero-point contributions which are too time-consuming to calculate. The large experimental error estimate, nearly 2 kJ/mol, must also be considered in this comparison.

The calculated energy differences for Ala, C5 minus C7eq, Table IV, display similar trends. There is considerable scatter in the HF results, from 5.79 kJ/mol at HF/4-21G to 0.84 kJ/mol at HF/6-31+G*. Enlarging the basis set leads to the divergence of energies from, rather than their convergence to, the MP2 energy (6.9 kJ/mol). Again, the effect of geometry optimization at the

---

(25) Sellers, H. L.; Schäfer, L. *J. Am. Chem. Soc.* **1978**, *100*, 7728.

(26) Schäfer, L.; Sellers, H. L.; Lovas, F. J.; Suenram, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 6566.

(27) Vishveshwara, S.; Pople, J. A. *J. Am. Chem. Soc.* **1977**, *99*, 2422.

(28) Wright, L. R.; Borkman, R. F. *J. Am. Chem. Soc.* **1980**, *102*, 6207. (b) Palla, P.; Petrongolo, C.; Tomasi, J. *J. Phys. Chem.* **1980**, *84*, 435. (c) Dykstra, C. E.; Chiles, R. A.; Garrett, M. D. *J. Comput. Chem.* **1981**, *2*, 266. (d) Millefiori, S.; Millefiori, A. *J. Mol. Struct.* **1983**, *91*, 391. (e) Bonaccorsi, R.; Palla, P.; Tomasi, J. *J. Am. Chem. Soc.* **1984**, *106*, 1945. (f) Siam, K.; Klimkowski, V. J.; Ewbank, J. D.; van Alsenoy, C.; Schäfer, L. *J. Mol. Struct.* **1984**, *110*, 171. (g) Ramek, M. *Int. J. Quantum Chem. Biol. Symp.* **1990**, *17*, 45. (h) Laurence, P. R.; Thomson, C. *Theor. Chim. Acta* **1981**, *58*, 121. (i) Masamura, M. *J. Mol. Struct.* **1987**, *152*, 293.

Table IV. Calculated Energies for the C5 and C7eq Conformations of *N*-Formylalanineamide

	C7eq		C5	
basis[a]	$-E$(tot)[b]	ϕ(N–C(α)) ψ(C(α)–C′)	E(rel)[c]	ϕ(N–C(α)) ψ(C(α)–C′)
HF/3-21G[d]	412.474 779	−84.5 67.3	5.24	−168.4 170.5
HF/3-21G[e]	412.474 780	−84.5 67.3	5.27	−168.4 170.5
HF/4-21G	413.830 881	−84.7 66.8	5.79	−166.6 169.9
HF/4-31G	414.173 491	−85.5 69.4	1.88	−161.5 164.5
HF/6-31G	414.599 943	−85.2 69.8	1.36	−160.9 164.0
HF/6-31G**	414.805 988	−85.3 76.0	1.24	−157.9 162.6
HF/6-311G**	414.900 685	−85.5 78.3	1.04	−156.8 162.2
MP2/6-311G**// HF/6-311G**[f]	416.385 739	−85.5 78.3	5.52	−156.8 162.2
MP2/6-311G**[g]	416.390 074	−81.5 82.5	6.92	−159.8 162.1
HF/6-31+G*[e]	414.799 097	−85.8 78.1	0.84	−155.6 160.2
MP2/6-31+G**// HF/6-31+G*[h]	416.067 460	−85.8 78.1	4.73	−155.6 160.2

[a] Standard basis sets were used as described in the text. [b] Hartrees. [c] RHF energy differences, C5 − C7eq, kJ/mol. [d] This study. [e] Results from Head-Gordon et al.;[20] for a discussion of other results for C5 by Head-Gordon et al.,[29] see text. [f] This study: MP2 energies calculated at the HF optimized geometries. [g] This study: MP2 energies for fully optimized MP2 geometries. [h] Single point MP2 energy at HF/6-31+G* geometry, by Head-Gordon et al.[20]

MP2 level is nonnegligible. The MP2/6-311G** energy difference for the MP2/6-311G** optimized geometries is 1.4 kJ/mol larger than the MP2 energy obtained for the HF/6-311G** optimized structures. This change is in opposite direction compared to Gly; i.e., in Ala the optimized value is larger than the unoptimized one. Thus, the consequences of optimization are not systematic and impossible to predict a priori.

The HF/3-21G ϕ,ψ energy surface of Ala calculated by Head-Gordon et al. was described in two reports.[20,29] There is a confusing contradiction in the two papers regarding the location of C5 in ϕ,ψ space. In ref 29, the HF/3-21G optimized torsions of C5 are given as $\phi = -191.6°$ and $\psi = +189.5°$, in ref 20, as $\phi = -168.4°$ and $\psi = +170.5°$. A similar conflict exists for C5 in refs 20 and 29 and other basis sets. We suspected a confusion of chiralities as the cause of the discrepancy. However, since the work of ref 29 was not fully documented, an independent calculation seemed necessary. Our results (Table IV) are in agreement with ref 20, but not ref 29; i.e., the location of C5 is in the usual region.[30]

Importance of MP2 Gradient Optimization for Conformational Geometry Trends

In 1979 some of us proposed to use HF optimized differences between bond distances and angles as constraints of data analysis in gas electron diffraction (GED).[14] GED data contain a great deal of information on conformational equilibria which often cannot be clearly deciphered because closely neighbored bond distances and angles are not resolved. Such differences, however, can be calculated accurately for many classes of molecules.

In spite of the advantages of the method,[4,14] it was frequently criticized because it was felt that HF geometries were not reliable. The MP2 geometries for Gly and Ala (Tables I and II) allow one to test in a novel way the reliability of calculated geometry trends for polyfunctional organic compounds. When experimental information is sparse or not available, the invariance of calculated results at different levels of computation can be used as a measure of accuracy.

Table V. Differences between the Bond Distances and Bond Angles in Conformations **1** and **2** of Glycine Obtained at the HF/4-21G, HF/6-311G**, and MP2/6-311G** Levels

	HF/4-21G	HF/6-311G**	MP2/6-311G**
	Bond Lengths (Å)		
N1—C2	−0.0190	−0.0175	−0.0179
N1—H6	+0.0009	+0.0026	+0.0031
C2—C3	−0.0213	−0.0123	−0.0174
C2—H8	+0.0003	+0.0016	+0.0014
C3═O4	+0.0009	+0.0030	+0.0032
C3—O5	+0.0193	+0.0123	+0.0161
O5—H10	−0.0091	−0.0041	−0.0147
	Bond Angles (deg)		
N—C2—C3	+3.16	+2.61	+4.12
N1—C2—H8	−1.56	−1.89	−2.35
C2—N1—H6	−1.20	−2.00	−2.89
C2—C3═O4	+4.16	+3.89	+3.07
C2—C3—O5	−3.15	−3.52	−2.10
C3—C2—H8	+0.24	+0.87	+0.69
C3—O5—H10	+3.70	+0.60	+2.38
O4═C3—O5	−1.01	−0.37	−0.98
H6—N1—H7	−0.79	−1.87	−2.64
H8—C2—H9	−0.58	−0.59	−0.87

Considering *absolute* values of bond distances and bond angles first, it is seen from Tables I and II that they vary considerably with basis set and computational methods. The C3═O4 bond distance in **1**, for example, is 1.203 Å at SCF/4-21G, 1.182 Å at HF/6-311G**, and 1.208 Å at MP2/6-311G**. Similarly, C–H bond distances are typically 1.081 Å at HF/4-21G, 1.085 Å at HF/6-311G**, and 1.094 Å at MP2/6-311G**. Among bond angles, those are particularly variable that contain hydrogen atoms and have a heteroatom at the apex. For example, C2–N1–H6 in **1** is 113.3° at HF/4-21G, 110.4° at HF/6-311G**, and 108.5° at MP2/6-311G**. Similar features are noticed in the HF and MP2 geometries of Ala (Table II), documenting significant dependence of absolute values on computational method.

When *differences* between bond distances and angles are considered, it is seen that, like absolute values, they also vary with the method of calculation, but the variations do not preclude the identification of some clear and important structural trends. Parameter differences are of two kinds: between the values of a given bond distance and angle in different conformations of the same molecule, or, alternatively, between comparable bond distances and angles in different parts of a molecule, or in series of homologous compounds. Differences of the first type are listed in Table V for **1** and **2**, and Table VI for C5 and C7eq. Ala is large enough to allow for comparisons (Table VII) within a single conformation.

It is seen from Tables V to VII that, for some of the values, the changes with computational method are virtually negligible. For example (Table V), the difference between N1–C2 in **1** and **2** is −0.019 Å for HF/4-21G, and −0.018 Å for both HF/6-311G** and MP2/6-311G**. For other parameters the changes are larger. For example, the values of C2–C3 in **1** and **2** differ by −0.021 Å at HF/4-21G, −0.012 Å at HF/6-311G**, and −0.017 Å at MP2/6-311G**. In spite of these fluctuations, the HF values alone still would allow one to establish a noticeable trend; that is, the C2–C3 bond length in **2** is longer than in **1** by 0.01 Å to 0.02 Å. The discrepancies between the MP2 and HF results suggest uncertainties of <0.01 Å in this trend, which is comparable to typical experimental errors in parameters of this type.

For bond angles, similar trends are observed. In some cases, such as H1—C3═O2, H1—C3—N4, or O2═C3—N4 (Table VI), changes in parameter differences with method are insignificant. In others, like N1–C2–C3 (Table V), the values span

---

(29) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. A. *Int. J. Quant. Chem.: Quant. Biol. Symp.* **1989**, *16*, 311.

(30) (a) Lewis, P. N.; Momany, F. A.; Scherage, H. A. *Isr. J. Chem.* **1973**, *11*, 121. (b) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J. L.; Marcoccia, J. F.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1991**, *113*, 6256.

Table VI. Differences between the Bond Distances and Bond Angles in the C5 and C7eq Geometries of *N*-Formylalanineamide

	HF/ 4-21G	HF/ 6-311G**	MP2/ 6-311G**
	Bond Lengths (Å)		
H1—C3	+0.0010	+0.0010	+0.0015
O2=C3	−0.0044	−0.0021	−0.0034
C3—N4	+0.0008	−0.0002	+0.0011
N4—H5	+0.0022	+0.0009	+0.0016
N4—C6	−0.0202	−0.0155	−0.0171
C6—H7	+0.0017	+0.0022	+0.0021
C6—C8	+0.0172	+0.0145	+0.0164
C6—C9	−0.0129	−0.0105	−0.0110
C8—H11	+0.0025	+0.0022	+0.0025
C8—H12	−0.0025	−0.0025	−0.0015
C8—H13	+0.0005	+0.0013	+0.0009
C9=O10	+0.0021	+0.0013	+0.0010
C9—N14	+0.0021	−0.0006	−0.0015
N14—H15	−0.0058	−0.0040	−0.0071
N14—H16	+0.0004	−0.0006	−0.0015
	Bond Angles (deg)		
H1—C3=O2	+0.50	+0.23	+0.49
H1—C3—N4	−0.52	−0.31	−0.40
O2=C3—N4	+0.03	+0.08	−0.10
C3—N4—H5	+3.08	+2.08	+2.83
C3—N4—C6	−0.91	−0.86	−0.80
N4—C6—H7	+2.49	+1.72	+2.30
N4—C6—C8	+1.48	+1.94	+1.46
N4—C6—C9	−3.72	−2.02	−2.02
H5—N4—C6	−2.21	−1.30	−2.00
C6—C8—H11	+1.49	+1.08	+1.52
C6—C8—H12	−1.97	−0.75	−1.34
C6—C8—H13	−0.44	−0.59	−0.42
C6—C9=O10	−0.18	−0.15	−0.05
C6—C9—N14	+1.40	+0.91	+1.37
H7—C6—C8	−1.70	−1.23	−1.85
H7—C6—C9	+1.38	+0.74	+0.75
C8—C6—C9	+0.09	−1.07	−0.54
C9—N14—H15	+1.57	+1.78	+2.99
C9—N14—H16	−0.02	+0.79	+1.14
O10=C9—N14	−1.23	−0.78	−1.31
H11—C8—H12	−0.19	−0.32	−0.44
H11—C8—H13	+1.56	+1.13	+1.42
H12—C8—H13	−0.45	−0.53	−0.73
H15—N14—H16	−1.49	+0.03	−0.00

a larger range, from +2.6° (HF/6-311G**) to +3.2° (HF/4-21G) and +4.1° (MP2/6-311G**). However, even though the largest discrepancy in this case is considerable, 1.5° between HF/6-311G** and MP2/6-311G**, the HF results alone still seem to establish a clear geometrical trend.

As to the performance of individual basis sets in HF calculations, it is seen that the HF/4-21G parameter differences are in better agreement than HF/6-311G** with MP2/6-311G** in some instances (e.g., N1-C2-C3, O4=C3-O5, Table V; C3-N4-H5, C6-C8-H11, O10-C9-N14, Table VI), but in worse agreement in others (e.g., H6-N1-H7, Table V; H15-N14-H16, C9-N14-H15, N4-C6-C9, and others Table VI).

For HF calculations with basis sets that do not contain polarization functions, it is well known that errors in relative bond lengths increase with electron delocalization,[31] and errors in bond angles are particularly large when heteroatoms are at the apex[32] and H-atoms are involved. In agreement with this experience, particularly large errors in the HF/4-21G calculations of Ala and Gly involve parameters like C3-O5-H10 (Table V), H15-N14-H16 (Table VI), and other parameters of the amide groups. At the same time it is interesting to note that, when differences between conjugated groups of the same type within the same conformation of a given molecule are determined, such as between the two amide groups within C5 or within C7eq of Ala, there is additional cancellation of errors and the trends obtained (Table VII) by different methods are no longer contradictory. Thus it is seen that even the simple HF/4-21G geometry optimizations of Gly and Ala contain clear and important structural trends which are not falsified by more advanced procedures.

In general, the data of Tables V to VII convey the impression that HF geometries allow one to derive structural trends (parameter differences) that can be of use in practical applications because their error limits are sufficiently small. The comparisons show for what type of parameter the error limits very likely are insignificant, and where larger errors must be expected. These findings are entirely consistent with recent comparisons of HF C-H[33] and C-X[32] bond distances (X = C, N, O, and F) and their experimental counterparts.

Importance of MP2 Gradient Optimization for the Location of Conformational Energy Minima

In past studies changes in internal coordinates as functions of conformation have been found to be important to empirical methods of molecular modeling.[5b] Specifically, differences between internal coordinates at different combinations of torsional angles in a given molecule have been useful in parameterizing empirical force fields. The availability of MP2-gradient optimized geometries of polyfunctional systems now allows one to explore a much more difficult area of study, i.e., variations in the exact locations of minima of dihedral space.

An example of this is given in Table IV where the dihedral angles of Ala, ϕ and ψ, show considerable variances between different basis sets at the HF level, and even greater change in going from HF to MP2. The ϕ dihedral angle in C7eq is relatively constant in all HF calculations, about −85°, but it changes to −81° in the MP2 geometry of Ala. At the same time, changes in ψ in C7eq are from 67° at HF/3-21G to 78° at HF/6-311G**, and 82° at MP2/6-311G**. In C5, ϕ varies between −168° and −156°, and ψ between 170° and 160°.

Differences of almost 15° in torsional angles of this kind are potentially highly significant. For example, the result is important because the H···X nonbonded interactions in these conformations are a determining factor in setting the dihedral angles. Clearly, changing the electronic character of the system by including electron correlation has a significant effect on dipeptide conformation. To examine this effect further, some H···X distances are given in Table VIII. It is seen that the H···O distances in C5 and C7eq are smaller at the MP2/6-311G** level than in the HF/6-311G** geometries. However, the magnitude of the shrinkage is much greater in C7eq (0.17 Å), in which the N-H bond is nearly in line with the electron-lone pair on O, than in C5 (0.04 Å), in which the interacting C=O and N—H bonds are approximately parallel.

In view of the shrinkage of H···X bonds by electron-correlation, it is interesting to note (Table X) that the net atomic charges in C5 and C7eq are very similar and essentially independent of the H···X interactions. If charges do not change with conformation but hydrogen-bond lengths do, the importance of electron-correlation effects on hydrogen bonding may lie in the exact shape and orientation of lone-pair orbitals. Similar results are found for Gly (Table IX).

The examples of Gly and Ala allow for an additional interesting observation. Chemical bonds can be characterized by a stationary point in the electron densities, which is absent in nonbonded interactions.[34,35] Electron density plots in the plane of the N···H-O and N-H···O atoms contain the characteristic stationary point in the case of the cyclic conformations **2** and **3**, but not in the case of **1**. Based on this criterion, the N-H···O interactions in **1** are nonbonded, whereas H10 and N1 in **2** and **3** form a true bond.

---

[31] Caminati, W.; di Bernardo, S.; Schäfer, L.; Kulp-Newton, S. Q.; Siam, K. *J. Mol. Struct.* **1990**, *240*, 263.

[32] (a) Boggs, J. E.; Cordell, F. R. *J. Mol. Struct.* **1981**, *76*, 329. (b) Schäfer, L.; van Alsenoy, C.; Scarsdale, J. N. *J. Mol. Struct.* **1982**, *86*, 349. (c) De Smedt, J.; van Houteghem, F.; van Alsenoy, C.; Geise, H. J.; Schäfer, L. *J. Mol. Struct.*, in press.

[33] (a) McKean, D. C.; Boggs, J. E.; Schäfer, L. *J. Mol. Struct.* **1984**, *116*, 313. (b) Schäfer, L.; Siam, K. *J. Chem. Phys.* **1988**, *88*, 7255.

[34] Villaveces, J. L.; Daza, E. E. *Int. J. Quant. Chem.: Quant. Chem. Symp.* **1990**, *24*, 97.

[35] Ramek, M.; Flock, M.; Kelterer, A.-M.; Cheng, V. K. W. *J. Mol. Struct.*, in press.

Table VII. Differences between Comparable Bond Lengths within the Same Conformation of *N*-Formylalanineamide

	C5			C7eq		
	HF/ 4-21G	HF/ 6-311G**	MP2/ 6-311G**	HF/ 4-21G	HF/ 6-311G**	MP2/ 6-311G**
(C3=O2) - (C9=O10)	-0.0005	-0.0027	-0.0009	+0.0060	+0.0007	+0.0035
(C3—N4) - (C9—N14)	-0.0069	-0.0062	-0.0063	-0.0056	-0.0066	-0.0089
(C6—C8) - (C6—C9)	+0.0172	+0.0092	+0.0071	-0.0129	-0.0158	-0.0203
(N4—C6) - (N4—C3)	+0.1090	+0.1004	+0.0860	+0.1300	+0.1157	+0.1042

Table VIII. Some Nonbonded H···X Distances (Å) in Conformations **1**–**3** of Glycine and C5 and C7eq of *N*-Formylalanineamide

	1 (H6—O4)	**1** (H10—O4)	**2** (H10—N1)	**3** (H10—N1)	C5 (H5—O10)	C7eq (H15—O2)
HF/3-21G	2.859	2.376		1.932	2.114	2.023
HF/4-21G	2.864	2.383	1.930	1.939	2.123	2.041
HF/4-31G	2.946	2.388	2.056	2.063	2.174	2.157
HF/6-31G	2.961	2.392	2.079		2.183	2.174
HF/6-31G**	2.783	2.279		2.051	2.206	2.227
HF/6-311G**	2.782	2.273	2.016	2.051	2.222	2.265
MP2/6-311G**	2.745	2.276	1.849	1.873	2.184	2.095

Table IX. Mulliken Populations for Three Conformations of Glycine

	1	**2**	**3**
	HF/6-311G**		
N1	7.466	7.576	7.562
C2	6.065	6.093	6.095
C3	5.535	5.501	5.499
(C=)O4	8.454	8.452	8.447
(C—)O5	8.402	8.380	8.379
(N—)H6	0.808	0.778	0.779
(N—)H7	0.808	0.778	0.786
(C—)H8	0.868	0.862	0.867
(C—)H9	0.868	0.862	0.865
(O—)H10	0.727	0.720	0.720
	MP2/G-311G**		
N1	7.423	7.531	7.523
C2	6.143	6.181	6.180
C3	5.713	5.664	5.664
(C=)O4	8.321	8.328	8.324
(C—)O5	8.315	8.295	8.295
(N—)H6	0.813	0.776	0.777
(N—)H7	0.813	0.776	0.781
(C—)H8	0.853	0.850	0.856
(C—)H9	0.853	0.850	0.849
(O—)H10	0.754	0.750	0.750

Table X. Mulliken Populations for the C5 and C7eq Conformations of *N*-Formylalanineamide

	C5		C7eq	
	MP2/ 6-311G**	HF/ 6-311G**	MP2/ 6-311G**	HF/ 6-311G**
H1	0.927	0.924	0.915	0.915
(C=)O2	8.360	8.487	8.375	8.500
C3	5.738	5.571	5.735	5.565
N4	7.389	7.503	7.359	7.485
(N4—)H5	0.755	0.732	0.765	0.753
C(α)6	6.105	6.019	6.153	6.077
(Cα)H7	0.847	0.862	0.842	0.851
C(β)8	6.252	6.204	6.278	6.204
C(′)9	5.650	5.485	5.628	5.458
(C9=)O10	8.357	8.498	8.350	8.489
(Cβ—)H11	0.909	0.916	0.872	0.889
(Cβ—)H12	0.853	0.867	0.885	0.902
(Cβ—)H13	0.886	0.900	0.875	0.891
N14	7.434	7.517	7.429	7.517
(N14—)H15	0.770	0.759	0.759	0.738
(N14—)H16	0.768	0.752	0.782	0.767

In parallel with this classification, it is seen that, Table VIII, the nonbonded distances in Gly, H6···O4 and H10···O4, are not greatly affected by electron correlation, in contrast to the hydrogen bond, H10···N1, for which the change is 0.17 Å. The shrinkage in H10···N1 is accompanied by large changes in the C–O–H angle, comparing the HF/6-311G** and MP2/6-311G** structures of **2** or **3**. For Ala, a similar analysis of electron densities shows that H15 and O2 in C7eq form a true bond, but H5 and O10 do not. Therefore, it is possible to expect that this criterion is generally valid.

In the analysis above, it was often implied that MP2 energies and structures are "better", i.e., more accurate, than HF geometries. In the absence of an extensive body of MP2 optimized geometries of polyfunctional molecules, we note the tentative nature of this assumption. Even MP2 gradient optimized results should be received with some suspicion. The perturbation series expansion in these calculations is terminated too early; hence the MP2 energy values are nonvariational. In addition, other open questions remain, such as complications by intramolecular basis set superposition effects which may be worse for MP2 than HF calculations. Deficiencies like these may be the reason why the MP2 energy of **2** is in no better agreement with experiment than the HF/3-21G result.

Conclusions

The data presented in Tables III and IV convey the impression that HF calculations have to be used with caution when accurate estimates of conformational energy differences are needed for simple systems like glycine (Table III) and *N*-formylalanineamide (Table IV). In the case of Gly, different levels of theory do not agree on the symmetry of the second energy minimum, the planar **2** or nonplanar **3**; the scatter in the energy difference between **2** and **1** calculated with different basis sets is 6 kJ/mol, and it is a particularly disturbing feature of the results that, the more advanced the HF calculations are, the more they deviate from the experimental value (5.9 ± 1.8 kJ/mol) or the MP2 result (3.5 kJ/mol). When corrections are made at the HF level for zero-point energies, divergence from the experiment is further enhanced.

For Ala, Table IV, the results are equally discouraging. The scatter in the energy difference between the two most stable conformations, C5 and C7eq, is 5 kJ/mol, and the HF values of ΔE deviate more from the MP2 results, the more advanced the calculations are. These observations are in direct contrast to the conclusion by Head-Gordon et al.[20] that the HF/6-31+G* results of Ala are "the best results currently feasible". As it turns out (Tables III and IV), the so-called "higher level of theory" can be linked with higher inaccuracy.

The analyses of Gly and Ala also show clearly that geometry optimization at the correlation level can have a significant effect on calculated energy. The MP2/6-311G** gradient optimized energy difference between **1** and **2** is 3.5 kJ/mol; the corresponding MP2 energy calculated at the HF/6-311G** geometries is 1.9 kJ/mol higher, i.e., 5.4 kJ/mol. Similarly, the effect of MP2 geometry optimization on conformational energy is 1.9 kJ/mol for **3**, and 1.4 kJ/mol for the difference between C5 and C7eq. Such energy increments cannot simply be neglected for moieties whose intended use it is[20] to serve as models for residues in peptides and proteins, where the frequent repetition of a given residue and its neglected energy increment will quickly accumulate forbidding errors. For example, the ΔG for the deformation of an α-helix to the coiled-coil state has been estimated[36] to be of the order of <0.5 kJ/mol per residue. When conditions are that stringent, the conclusion must be that the procedure of using HF structures for single-point energy calculations at the MP2 level of theory, as applied by Head-Gordon et al. to Ala,[20] is highly questionable, and the results are unreliable. Thus, it is difficult to support the claim by Head-Gordon et al.[29] that MP2/6-31+G** energies of Ala calculated at HF/6-31+G* geometries are "benchmark" calculations.

Considerable dependence on computational procedure is also found for the exact locations of conformational energy minima. The HF/3-21G ψ value of C7eq by Head-Gordon et al.[20] differs by more than 15° from the MP2/6-311G** minimum (Table IV). The deviation of the HF/6-311G** minimum from the latter is about 4°. Thus, it is a peculiar predicament of the HF conformational energy maps of diamides that, at the important locations of the two most stable conformations, C7eq and C5, small basis sets like 3-21G and 4-21G yield good energies, but the torsional angles are poor, whereas, for large basis sets like 6-31+G* or 6-311G**, the locations of the minima are acceptable, but the energies are poor. Differences of >10° in the ϕ and ψ values of energy minima, as found in Table IV for HF calculations of Ala with different basis sets correspond to uncertainties which are potentially of great importance. Incorporated into the modeling of an extended peptide chain, for example, and propagated along several hundreds of residues, such errors may easily lead to secondary structures which are meaningless.

Our study reveals why HF conformational energy maps are intrinsically inaccurate. One of the most important contributions to conformational energy in large organic molecules is derived from nonbonded interactions. At the HF level these interactions are incorrectly evaluated twice over: the dispersion effects are neglected which are an important factor in determining them, and the nonbonded distances are wrong, because the torsions are wrong. For the same reason MP2 single-point energies without MP2 geometry optimization are an intrinsically inaccurate product. If geometry is not optimized, nonbonded interactions are not evaluated correctly. But it is particularly in these interactions that the dispersion forces come to the fore.

For polyfunctional organic molecules, all calculated conformational properties are basis set and method dependent. However, the gravity of this handicap and the uncertainties connected with it have different consequences for energy than geometry. For geometrical parameters of many classes of compounds, relatively constant empirical corrections can be defined by comparisons with experimental structures.[32] In contrast to this, when errors in relative energies can be >5 kJ/mol at random and advanced methods of calculation are in worse agreement than approximate ones with experimental values or MP2 results, then the impression is easily derived that no generally valid precise criteria exist for estimating the accuracy of ab initio energies.

Conformationally dependent geometry changes, like those recorded in Tables V to VII, also display significant dependence on computational method, but clear structural trends can be identified whose potential utility in practice is not impaired by the variances, because the concomitant errors are not destructive. At least, such parameter differences, many of them simple HF/4-21G, have in the past been used with considerable success in a large number of experimental studies.[3,4,14] Thus, even the simple HF/4-21G calculations of diamide conformational geometry maps[37] capture many essential structural trends which form an important part of the conformational properties of these systems.

As to the conflicting HF results concerning the symmetry of the cyclic form of Gly,[15] the MP2 calculations favor the nonplanar **3** over the planar **2**. We do not think that this result is in conflict with the experimental observation[16,17] of **2**. If the vibrational ground state of **3** is above the barrier of a dual potential well, the average structure derived from ground-state data will be planar, even though the equilibrium structure is not. Extensive studies of excited states in the microwave region, which do not exist, will be needed to solve this question, if it can be solved at all.

Acknowledgment. Part of this work was supported by the Austrian Funds zur Förderung der Wissenschaftlichen Forschung (Project P6856). L.S. is grateful to Polygen/MSI and IBM Corp. for support.

Registry No. Glycine, 56-40-6; *N*-formylalanine, 10512-86-4.

---

(36) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601.

(37) Siam, K.; Kulp, S. Q.; Ewbank, J. D.; Schäfer, L. *J. Mol. Struct.* **1989**, *184*, 143.