

Theoretical Study of Blocked Glycine and Alanine Peptide Analogues

Teresa Head-Gordon,^{*,†,§} Martin Head-Gordon,^{†,§} Michael J. Frisch,[‡]
Charles L. Brooks III,[†] and John A. Pople[†]

Contribution from the Department of Chemistry, Carnegie-Mellon University,
Pittsburgh, Pennsylvania 15213, and Lorentzian Inc., 127 Washington Avenue,
North Haven, Connecticut 06473. Received September 19, 1990

Abstract: We present a high-level ab initio study of the model blocked alanine and glycine dipeptide molecules, (S)- α -(formylamino)propanamide and α -(formylamino)ethanamide. Fully relaxed grids of the conventionally defined conformational space variables ϕ and ψ have been evaluated for each molecule at the HF/3-21G level. In order to obtain the best results currently feasible, HF/6-31+G* fully optimized geometries and frequencies were obtained for all minima and transition structures observed on the HF/3-21G grid, and correlation corrections were explored with MP2/6-31+G**//HF/6-31+G* single-point energy calculations. The HF/3-21G structures are in reasonable agreement with the results of the larger basis, although the relative energies are sometimes poor. At the higher level of theory, we do not find minima in the regions of ϕ, ψ space corresponding to protein secondary structures, although these regions are relatively low in energy. We report the presence of a cusp in certain regions of the relaxed (ϕ, ψ) map of both systems; the surfaces are double-valued in these regions, and the cusp occurs where the energies of the two surfaces cross. We also observe significant deviations from peptide planarity (up to 40°) in several regions of the (ϕ, ψ) map. The discontinuities and large peptide distortions indicate that the (ϕ, ψ) degrees of freedom alone do not fully define the available conformational space of these dipeptide molecules. Finally we discuss the molecular origin of the differences between the alanine and glycine fully relaxed (ϕ, ψ) surfaces, differences which are qualitatively consistent with those inferred from Ramachandran maps.

I. Introduction

The use of empirical potential functions and molecular dynamics to study the structural, dynamical, and equilibrium thermodynamic properties of biological macromolecules such as proteins is now common.¹ Assuming that the underpinnings of statistical mechanics are satisfied in practice in a molecular dynamics simulation of these molecules (i.e. the system is ergodic, the ensemble is well-defined, and the properties of interest are converged), condensed-phase simulations can in principle provide unique information not readily accessible to experiment. However, the utility of any molecular dynamics simulation will depend critically on the adequacy of the empirical potential model used. In practice, the empirical protein potential energy functions used to study protein motions on short time scales are derived from a combination of theoretical and experimental data on (smaller) mono- and dipeptide systems. Because proteins exhibit much conformational flexibility and contain stable structural fingerprints such as hydrogen-bonding conformations and secondary structural motifs, these smaller peptide systems should exhibit similar properties if they are to be useful models of the protein macromolecule. Peptides such as 1-(acetylamino)-N-methylethanamide and (S)-2-(acetylamino)-N-methylpropanamide (often referred to as glycine dipeptide (GD) and alanine dipeptide (AD), respectively) have been studied by various theoretical methods,¹⁻¹⁵ since these fragments show conformational variations which are similar to proteins, and therefore may be considered as reasonable models of the larger globular proteins.

Molecules such as (S)- α -(formylamino)propanamide (Figure 1) and α -(formylamino)ethanamide (Figure 2) are also of interest, since they are formed from GD and AD by replacing the terminal methyl groups by hydrogen atoms. We refer to the molecules in Figures 1 and 2 as the glycine dipeptide analogue (GDA) and alanine dipeptide analogue (ADA). We have recently presented a preliminary ab initio study of AD and ADA.⁷ The main conclusions reached in that report can be summarized as follows:

(1) A comparison of ab initio calculations on the AD and ADA molecules shows that the terminal methyl groups play an insignificant role in the structure and energetics of alanine dipeptide at several important conformations.⁷ Therefore all subsequent

calculations were performed on the smaller ADA molecule since it is apparently functionally equivalent to AD.

(2) The biologically relevant conformational space (i.e. excluding conformations with cis peptide units, and mirror images) of the alanine dipeptide analogue is often described in terms of the flexible backbone dihedral angles ϕ ($C_1-N_1-C_\alpha-C_2$) and ψ ($N_1-C_\alpha-C_2-N_2$) (see Figure 1). On the basis of MP2/6-31+G**//HF/6-31+G* benchmark calculations at several important minima, HF/3-21G was chosen as a reasonable level of theory with which to evaluate a fully relaxed (ϕ, ψ) map. A (ϕ, ψ) grid with 30° spacings was generated at this level of theory.

(3) From the 30° map, we determined the presence of six conformational minima, and the approximate pathways and barriers interconnecting them. The map clearly demonstrates that many of the barriers were much lower than indicated by past ab initio studies,⁵ thereby emphasizing the need to perform fully relaxed geometry optimizations. However, the resolution of this

(1) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*; John Wiley and Sons, Inc.: 1988; Advances in Chemical Physics, Volume LXXI, also see references therein.

(2) Maigret, B.; Pullman, B.; Perahia, D. *J. Theor. Biol.* **1971**, *31*, 269-285.

(3) Peters, D.; Peters, J. *J. Mol. Struct.* **1981**, *85*, 107-123.

(4) Scarsdale, J. N.; Van Alsenoy, C.; Klimkowski, V. J.; Schäfer, L.; Momany, F. A. *J. Am. Chem. Soc.* **1983**, *105*, 3438-3445.

(5) Schäfer, L.; Klimkowski, V. J.; Momany, F. A.; Chüman, H.; Van Alsenoy, C. *Biopolymers* **1984**, *23*, 2335-2347.

(6) Weiner, S. J.; Chandra Singh, U.; O'Donnell, T. J.; Kollman, P. A. *J. Am. Chem. Soc.* **1984**, *106*, 6243-6245.

(7) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L.; Pople, J. A. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1989**, *16*, 311-322.

(8) Ramakrishnan, C.; Ramachandran, G. N. *Biophys. J.* **1965**, *3*, 909-933.

(9) Peters, D.; Peters, J. *J. Mol. Struct.* **1979**, *53*, 103-119.

(10) Schäfer, L.; Van Alsenoy, C.; Scarsdale, J. N. *J. Chem. Phys.* **1982**, *76*, 1439-1444.

(11) Klimkowski, V. J.; Schäfer, L.; Momany, F. A.; Van Alsenoy, C. *J. Mol. Struct. (Theochem)* **1985**, *124*, 143-153.

(12) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J.; Marcocchia, J.; Csizmadia, I. G. Preprint.

(13) Cheam, T. C.; Krimm, S. *J. Mol. Struct.* **1989**, *193*, 1-34.

(14) Cheam, T. C.; Krimm, S. *J. Mol. Struct. (Theochem)* **1989**, *188*, 15-43.

(15) Cheam, T. C.; Krimm, S. *J. Mol. Struct. (Theochem)* **1990**, *206*, 173-203.

[†] Carnegie-Mellon University.

[‡] Lorentzian Inc.

[§] Present address: AT&T Bell Laboratories, Murray Hill, NJ 07974.

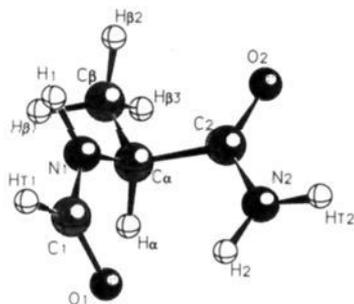


Figure 1. The structure of the $C7_{\alpha}$ conformer for the alanine dipeptide analogue (ADA).

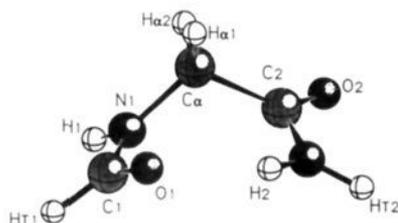


Figure 2. The structure of the $C7$ conformer for the glycine dipeptide analogue (GDA).

map is insufficient to elucidate all stationary points.

(4) Finally, our *ab initio* MP2 correlation energy calculations do not support the use of empirical dispersion energy corrections.⁶ The relative energy ordering of minima found by empirical corrections to Hartree-Fock results are different from the energy ordering determined at MP2/6-31+G**//HF/6-31+G*.

In this work we wish to provide a more extensive study of the structural and energetic properties of the ADA peptide and to extend our study to the simplest dipeptide without a side chain, GDA, using *ab initio* molecular orbital theory.¹⁶ In addition to characterizing these systems, the results we obtain will be of use in assessing the performance of empirical potential functions and semiempirical methods. Recent advances in integral evaluation techniques¹⁷ and direct *ab initio* methods¹⁸⁻²³ make such calculations feasible on systems of this size.

The outline of the present paper is as follows. In section II we briefly describe the levels of *ab initio* theory used for evaluating the conformational and energetic properties of glycine and alanine dipeptide analogues. In section III we present the results of stationary point calculations on the two peptide systems. We report full geometry optimizations and frequency calculations for all minima and many transition structures at the HF/3-21G and HF/6-31+G* levels, as well as MP2/6-31+G** correlation energy calculations using the HF/6-31+G* structures. Fully relaxed (ϕ, ψ) maps for GDA and ADA have been generated at the HF/3-21G level with 15° spacing (four times the resolution of

our previous study⁷). In section IV, we discuss the general features of the HF/3-21G (ϕ, ψ) surface for the GDA molecule and compare our results with previous *ab initio* work. We also report an unusual feature on the GDA and ADA surfaces: the presence of a cusp ridge which is due to mapping many flexible degrees of freedom onto a two-dimensional surface. In section V we analyze the ADA (ϕ, ψ) surface and compare it with the GDA results to discern the structural and energetic role of the ADA methyl side chain. Finally in section VI we summarize our main results and conclusions and briefly discuss future work.

II. Theoretical Methods

All *ab initio* calculations were performed with the Gaussian 88²² and Gaussian 90²³ molecular orbital packages, and were run on Multiflow, Alliant, and Cray computers. All calculations presented here and in the previous study⁷ have been placed in the Carnegie-Mellon Quantum Chemistry Archive²⁴ and are available on request from the authors. The fully relaxed 15° (ϕ, ψ) map of the ADA molecule is the result of 576 full optimizations of all degrees of freedom except ϕ and ψ , which are frozen at each grid point. The fully relaxed GDA surface, on the other hand, requires fewer optimizations to complete a 15° grid. Inversion through the achiral α -carbon [i.e. $(\phi, \psi) \rightarrow (-\phi, -\psi)$] yields an equivalent structure; therefore only half the grid points are needed to evaluate an entire map. In actuality, additional grid points were required due to the fact that parts of the surface are double-valued, which we discuss more thoroughly in section III.

The (ϕ, ψ) map optimizations are performed at the Hartree-Fock (HF) level with the small split valence 3-21G basis.²⁵ We have already shown⁷ that HF/3-21G structures and relative energies for the lowest three ADA minima are in very good qualitative agreement with the results of larger calculations, using the methods described below. With the aid of the ADA and GDA maps, we have performed additional HF/3-21G full geometry optimizations using the structure at the closest grid point to a possible stationary point as an initial guess. To provide a further check on the HF/3-21G results, and to obtain the best results currently feasible, we have performed HF optimizations on the minima and transition structures using the larger 6-31+G* basis.²⁶ This basis set has been quite successful in studies of hydrogen-bonding systems, where the need for diffuse functions was recognized.²⁷ The optimized HF/3-21G geometries were used as starting structures for full optimizations at the HF/6-31+G* level. Standard first-order optimization methods^{28,29} were used to locate all stationary points at HF/3-21G and HF/6-31+G*. To determine the character (i.e. minimum, transition structure, etc.) of the resulting stationary points, analytical second-derivative calculations are performed at each optimized structure. The HF/6-31+G* structures are also used for single-point calculations with the 6-31+G** basis, at the second-order Møller-Plesset (MP2) level of theory.¹⁶ The MP2 method with an extended basis such as 6-31+G** recovers much of the electron correlation energy neglected in HF theory, and significantly improves HF relative energies.¹⁶ Direct methods,¹⁸⁻²² in which quartic storage of the two-electron integrals is eliminated, are used to make these large HF optimizations,¹⁸ force constant calculations,¹⁹ and MP2 correlation energy calculations^{20,21} feasible on minisupercomputers.

Finally, we comment on the optimization techniques employed in this work. In general, all optimizations are started with analytical second derivatives; the hessian is updated by a quasi-Newton method^{28,29} on subsequent geometry steps until convergence is achieved. In the cases where an expected minimum or transition structure is not found (e.g. as discussed in section III, not all HF/3-21G stationary points are located at HF/6-31+G*), we do an additional optimization where analytical second derivatives are used at every geometry step. In some cases (section III), this additional calculation does not yield the desired stationary point. Although we cannot conclusively prove that a stationary point does

(16) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(17) Head-Gordon, M.; Pople, J. A. *J. Chem. Phys.* **1988**, *89*, 5777-5786.

(18) Almlöf, J.; Faegri, K., Jr.; Korsell, K. *J. Comput. Chem.* **1982**, *3*, 385-389.

(19) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys.* **1990**, *141*, 189-196.

(20) Head-Gordon, M.; Frisch, M. J.; Pople, J. A. *Chem. Phys. Lett.* **1988**, *153*, 503-506.

(21) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275-280.

(22) Gaussian 88; Frisch, M. J.; Head-Gordon, M.; 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 Inc., Pittsburgh, PA 15213.

(23) Gaussian 90; 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 Inc., Pittsburgh, PA 15213.

(24) Carnegie-Mellon Quantum Chemistry Archive; Whiteside, R. A., Frisch, M. J., Pople, J. A., Eds.; Carnegie-Mellon University: Pittsburgh, PA, 1988.

(25) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939-947.

(26) (a) Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209-214. (b) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265-3269. (c) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257-2261. (d) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294-301.

(27) Frisch, M. J.; Pople, J. A.; Del Bene, J. E. *J. Phys. Chem.* **1985**, *89*, 3664-3669.

(28) Baker, J. *J. Comput. Chem.* **1986**, *7*, 385-395.

(29) Schlegel, H. B. *J. Comput. Chem.* **1982**, *3*, 214-218.

Table I. HF/3-21G and HF/6-31+G* Stationary Point Structures and Energies for GDA

structure	label	HF/3-21G				HF/6-31+G* ^f			
		ϕ^b	ψ^b	energy ^a	N^c	ϕ^b	ψ^b	energy ^a	N^c
C7	1G	-83.3	64.7	0.00 ^d	0	-85.2	67.4	0.58 ^e	0
C5	2G	-180.0	180.0	0.65	0	-180.0	180.0	0.00	0
β	3G	-121.9	25.2	3.27	0	-	-	-	-
$\beta \rightarrow C7$	4G	-110.4	38.4	3.55	1	-	-	-	-
C7 \rightarrow C5	5G	-97.9	119.1	4.83	1	-85.3	126.7	1.86	1
C7 \rightarrow C7	6G	0.0	0.0	8.33	1	0.0	0.0	9.74	1
C5 $\rightarrow \beta$	7G	-125.3	-96.4	8.86	1	-	-	-	-
$\beta \rightarrow \beta$	8G	-112.2	-58.6	8.95	1	-	-	-	-
C ₂ cusp	9G	-180.0	0.0	11.24	2	-180.0	0.0	8.95	2
C7 $\xrightarrow{\text{alt}}$ C7	10G	-4.0	84.6	12.32	1	-0.9	79.9	10.27	1
C ₂ max	11G	0.0	180.0	23.60	2	0.0	180.0	22.98	2

^a In units of kcal/mol. ^b In units of deg. See text for dihedral angle definitions. ^c Number of imaginary frequencies. ^d The zero of energy is -373.648 790 3 hartrees. ^e The zero of energy is -375.762 297 2 hartrees. ^f Stationary points which could not be located at the HF/6-31+G* level by using the procedure described in section II are marked with dashes.

Table II. HF/3-21G and HF/6-31+G* Stationary Point Structures and Energies for ADA

structure	label	HF/3-21G				HF/6-31+G* ^f			
		ϕ^b	ψ^b	energy ^a	N^c	ϕ^b	ψ^b	energy ^a	N^c
C7 _{eq}	1A	-84.5	67.3	0.00 ^d	0	-85.8	78.1	0.00 ^e	0
C5	2A	-168.4	170.5	1.26	0	-155.6	160.2	0.19	0
C7 _{ax}	3A	74.1	-57.3	2.53	0	75.1	-54.1	2.56	0
β_2	4A	-128.0	29.7	3.83	0	-110.4	12.0	2.24	0
α_L	5A	63.8	32.7	5.95	0	69.5	24.9	4.73	0
α'	6A	-178.6	-44.1	7.31	0	-165.6	-40.7	5.52	0
α_D	7A	67.5	-177.3	8.16	0	-	-	-	-
C7 _{eq} $\rightarrow \beta_2$	8A	-116.7	42.5	4.04	1	-106.0	20.0	2.26	1
C7 _{eq} \rightarrow C5	9A	-113.0	114.7	4.86	1	-100.6	133.3	1.11	1
C7 _{ax} $\rightarrow \alpha_L$	10A	76.8	-4.3	7.03	1	72.3	14.4	4.76	1
C7 _{ax} $\rightarrow \alpha_D$	11A	63.2	-127.7	8.87	1	-	-	-	-
C7 _{ax} $\rightarrow \alpha'$	12A	128.2	-28.9	9.03	1	138.9	-28.7	7.15	1
C7 _{eq} \rightarrow C7 _{ax}	13A	-0.5	1.7	9.77	1	2.6	-45.3	10.07	1
C5 \rightarrow C7 _{ax}	14A	106.7	-178.6	9.99	1	120.3	-153.0	7.39	1
$\beta_2 \rightarrow \alpha'$	15A	-118.3	-55.7	10.11	1	-119.1	-48.9	7.30	1
C5 $\rightarrow \alpha'$	16A	-159.6	-98.4	10.35	1	-149.6	-94.5	6.76	1
C7 _{eq} $\rightarrow \alpha_L$	17A	-1.8	86.8	12.27	1	2.6	82.2	9.74	1
$\alpha_L \rightarrow \alpha_D/C5$	18A	84.8	106.2	13.17	1	86.5	85.1	10.59	1
$\beta_2 \rightarrow \alpha_L$	19A	113.2	76.0	13.29	1	-	-	-	-

^a In units of kcal/mol. ^b In units of deg. See text for dihedral angle definitions. ^c Number of imaginary frequencies. ^d The zero of energy is -412.474 780 0 hartrees. ^e The zero of energy is -414.799 097 3 hartrees. ^f Stationary points which could not be located at the HF/6-31+G* level by using the procedure described in section II are marked with dashes.

not exist at HF/6-31+G* with this extra calculation (since a full grid search at this level of theory is required), this fairly elaborate procedure makes it unlikely.

III. Stationary Point Results for Glycine and Alanine Dipeptide

In our previous study⁷ of ADA, the choice of HF/3-21G for generating the fully relaxed (ϕ, ψ) surface was based on a comparison of this level of theory with HF/6-31+G* structures and MP2/6-31+G** energies for *three* ADA minima. Since that time we have generated a fully relaxed map in 15° intervals at HF/3-21G for ADA, from which we observe 20–25 possible stationary points (see Figure 4), and for GDA, which exhibits comparatively fewer stationary points (see Figure 3). We have also fully optimized all possible minimum and saddle point structures that are evident in Figures 3 and 4, at HF/3-21G and HF/6-31+G*. (We have chosen not to optimize what we feel are uninteresting stationary structures, such as most of the maxima in Figures 3 and 4.) Thus a comparison of results of full geometry optimizations at HF/3-21G and HF/6-31+G* (along with MP2/6-31+G**//HF/6-31+G* single-point energies) provides a larger database with which to assess the overall reliability of the smaller basis set and the Hartree-Fock approximation. The results of this set of calculations are given in Tables I–IV for the GDA and ADA systems. The HF/3-21G stationary point structures are numbered as an aid for finding them in Figures 3 and 4. The additional designation of A or G, depending on whether we are referring to an alanine or glycine structure, will be used in the text and tables.

Table III. HF/3-21G, HF/6-31+G*, and MP2/6-31G**//HF/6-31G* Energies for GDA

structure	label	HF/3-21G ^a	HF/6-31+G* ^b	MP2 ^a /HF ^{a,f}
C7	1G	0.00 ^b	0.58 ^c	0.00 ^d /0.60 ^e
C5	2G	0.65	0.00	1.11/0.00
β	3G	3.27	-	-
$\beta \rightarrow C7$	4G	3.55	-	-
C7 \rightarrow C5	5G	4.83	1.86	1.76/1.98
C7 \rightarrow C7	6G	8.33	9.74	9.13/9.78
C5 $\rightarrow \beta$	7G	8.86	-	-
$\beta \rightarrow \beta$	8G	8.95	-	-
C ₂ cusp	9G	11.24	8.95	9.60/8.90
C7 $\xrightarrow{\text{alt}}$ C7	10G	12.32	10.27	9.62/10.58
C ₂ max	11G	23.60	22.98	22.69/23.07

^a In units of kcal/mol. ^b The HF/3-21G zero of energy is -373.648 790 3 hartrees. ^c The HF/6-31+G* zero of energy is -375.762 297 2 hartrees. ^d The MP2/6-31+G**//HF/6-31+G* zero of energy is -376.878 277 7 hartrees. ^e The HF/6-31+G**//HF/6-31+G* zero of energy is -375.779 057 6 hartrees. ^f Stationary points which could not be located at the HF/6-31+G* level by using the procedure described in section II are marked with dashes.

In Tables I and II, we compare the GDA and ADA ϕ, ψ values and energies of all minima and transition structures determined at the HF/3-21G and HF/6-31+G* level of theory. We have chosen the ϕ and ψ variables as geometric probes since they are thought to describe the complete conformational space of the

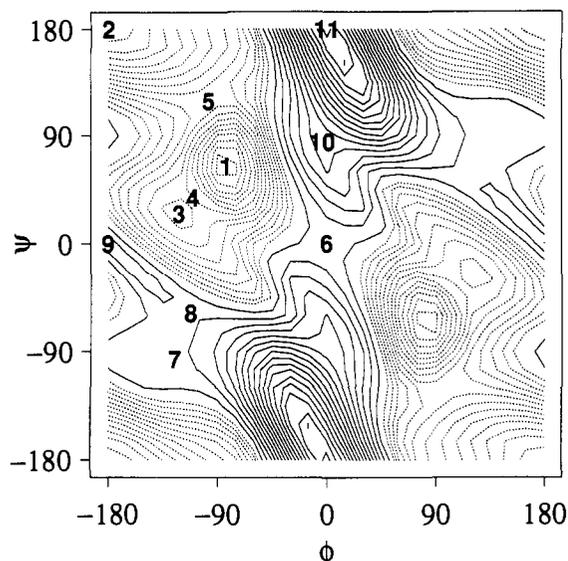


Figure 3. The (ϕ, ψ) surface of the glycine dipeptide analogue, at the HF/3-21G level of theory. Geometry optimizations of all variables except ϕ and ψ were performed on a grid with 15° spacing. The dashed energy contours are drawn every 0.5 kcal/mol and extend from the zero of energy, corresponding to the C7 conformer (labeled as 1, and illustrated in Figure 2), to 7.0 kcal/mol. Solid contours are drawn every 1.0 kcal/mol thereafter. Each stationary point that has been isolated by full geometry optimization and characterized by analytical second derivatives is marked on the figure in ascending order of energy. The numbers correspond directly to the entries in Tables I and III and are also used in the text as a connection with this figure. Since the map is symmetric with respect to inversion through the origin, only a non-redundant set of points are labeled.

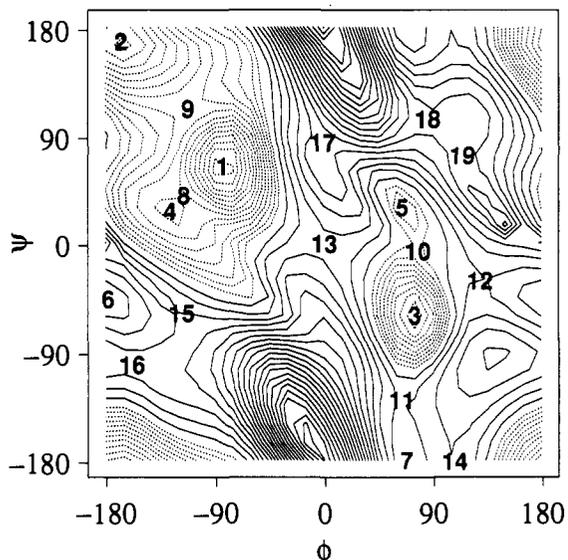


Figure 4. The HF/3-21G (ϕ, ψ) surface of the alanine dipeptide analogue, obtained from a grid of points with 15° spacing. The dashed energy contours are drawn every 0.5 kcal/mol and extend from the zero of energy, corresponding to the C7_{eq} conformer (labeled as 1, and illustrated in Figure 1), to 7.0 kcal/mol. Solid contours are drawn every 1.0 kcal/mol thereafter. The stationary points which we have characterized in Tables II and IV are labeled on the figure in order of increasing relative energy.

dipeptide molecules considered here.⁸ We also note that these degrees of freedom are particularly "soft", i.e. reasonably large displacements of these variables can result in small changes in energy (relative to bonds and angles and other dihedrals, for example).

A comparison of the structural probes for a given stationary point in Tables I and II suggests that the agreement between the HF/3-21G and HF/6-31+G* structures is generally quite good

Table IV. HF/3-21G, HF/6-31+G*, and MP2/6-31+G**//HF/6-31+G* Energies for ADA

structure	label	HF/3-21G ^a	HF/6-31+G* ^{a,f}	MP2 ^d /HF ^{e,f}
C7 _{eq}	1A	0.00 ^b	0.00 ^c	0.00 ^d /0.00 ^e
C5	2A	1.26	0.19	1.13/0.14
C7 _{ax}	3A	2.53	2.56	2.19/2.53
β_2	4A	3.83	2.24	2.67/2.29
α_L	5A	5.95	4.73	4.46/4.82
α'	6A	7.31	5.52	5.83/5.69
α_D	7A	8.16	—	—
C7 _{eq} \rightarrow β_2	8A	4.04	2.26	2.71/2.32
C7 _{eq} \rightarrow C5	9A	4.86	1.11	1.82/1.18
C7 _{ax} \rightarrow α_L	10A	7.03	4.76	4.53/4.85
C7 _{ax} \rightarrow α_D	11A	8.87	—	—
C7 _{ax} \rightarrow α'	12A	9.03	7.15	7.56/7.32
C7 _{eq} \rightarrow C7 _{ax}	13A	9.77	10.07	9.74/10.28
C5 \rightarrow C7 _{ax}	14A	9.99	7.39	7.91/7.40
β_2 \rightarrow α'	15A	10.11	7.30	8.01/7.31
C5 \rightarrow α'	16A	10.35	6.76	6.70/6.78
C7 _{eq} \rightarrow α_L	17A	12.27	9.74	9.35/10.01
α_L \rightarrow α_D /C5	18A	13.17	10.59	10.64/10.65
β_2 \rightarrow α_L	19A	13.29	—	—

^aIn units of kcal/mol. ^bThe HF/3-21G zero of energy is -412.4747800 hartrees. ^cThe HF/6-31+G* zero of energy is -414.7990973 hartrees. ^dThe MP2/6-31+G**//HF/6-31+G* zero of energy is -416.0674595 hartrees. ^eThe HF/6-31+G**//HF/6-31+G* zero of energy is -414.8188004 hartrees. ^fStationary points which could not be located at the HF/6-31+G* level by using the procedure described in section II are marked with dashes.

($\Delta\phi \approx 10^\circ$ and $\Delta\psi \approx 10^\circ$), with the exception of structures involving the ADA C₅ [2A], β_2 [4A], and α_D [7A] minima, and the GDA β [3G] minimum. However, in Figures 3 and 4 we observe that the topology of the HF/3-21G ϕ, ψ surface in the region of these minima is quite broad and flat, a feature that probably carries over to a HF/6-31+G* (ϕ, ψ) surface as well. We thus believe that the larger differences in the structures in the vicinity of the ADA C₅ [2A] and β_2 [4A] conformers are due to the softness of the ϕ and ψ variables in this region. The ADA α_D [7A] minimum is a mere dimple on the HF/3-21G surface since the $\alpha_D \rightarrow$ C7_{ax} [11A] barrier is only 0.7 kcal/mol; therefore its disappearance at HF/6-31+G* is not surprising. Similar conclusions apply to the disappearance of the GDA β structure at HF/6-31+G*, since the $\beta \rightarrow$ C7 [4G] barrier at HF/3-21G is only 0.28 kcal/mol. We conclude that the HF/3-21G structures are reasonably reliable overall.

However, the agreement between HF/3-21G and HF/6-31+G* relative energies is not good; the relative energies of most minima and transition structures are lower at HF/6-31+G* than HF/3-21G. In fact, the HF/6-31+G* results suggest that the true surfaces may be more flat and featureless than the HF/3-21G surfaces. The energy ordering has changed between the ADA minima C7_{ax} [3A] and β_2 [4A]. The global minimum for the GDA molecule has also changed on going from HF/3-21G (C7 [1G]) to HF/6-31+G* (C5 [2G]), although the energy differences are less than 1 kcal/mol.

Clearly, we must consider levels of theory that include correlation in order to resolve such discrepancies. In Tables III and IV we compare the GDA and ADA MP2 single-point energies obtained at the HF/6-31+G* geometries with the HF/3-21G and HF/6-31+G* energies. An initial overview of these tables reveals that the energies determined at HF/6-31+G* are in far better agreement with the MP2 single-point energies than the HF/3-21G energies. In fact, the correlation energy correction ranges only from 0.3 to 1.1 kcal/mol over the HF/6-31+G* energies for all stationary points found. Therefore in accord with conventional wisdom, HF/3-21G is found to be a less adequate level of theory for energies than for structures. HF/6-31+G* appears reliable for these systems when stationary point relative energies differ by more than about 1 kcal/mol.

IV. Glycine Dipeptide Results

The GDA HF/3-21G, 15° (ϕ, ψ) grid is shown in Figure 3. The symmetry with respect to inversion through the $(0^\circ, 0^\circ)$ origin

discussed in section II is evident. There are hence three unique HF/3-21G minima, C7 [1G], C5 [2G], and β [3G], which are listed in Table I. The very shallow β [3G] minimum disappears at HF/6-31+G*, so it is likely that the only stable minima on the gas-phase GDA surface correspond to the hydrogen-bonding conformations C7 [1G] and C5 [2G].

This implies that any secondary structure-like conformation which may exist in solution for this dipeptide is most likely driven and stabilized by solvent interactions, where perhaps the driving force is the hydrophobic effect.³⁰ The initiation stage of secondary structure formation in longer peptides and proteins is thus not driven by intramolecular interactions in the dipeptide unit, although the relevant α -helical region is of low energy and conformationally accessible. This of course indicates the limitations of glycine dipeptide as a model of globular proteins, since no secondary structure features are present in such a small peptide.

Previous molecular orbital theory work on glycine dipeptide includes semiempirical PCILO calculations,² unrelaxed HF/STO-3G (ϕ, ψ) maps,⁹ fully relaxed HF/4-21G geometries at several relevant minima,^{10,11} and vibrational force fields and dipole moment derivatives at HF/4-21G at the C7 and C5 minima.¹³ The problems with the first two procedures are well-recognized,^{4,5,10,11} so we will not critique them here. However, the HF/4-21G study of Schäfer et al.^{10,11} is of interest since we have found differences between their GD study and the GDA HF/3-21G results presented above. This is somewhat surprising since the HF/3-21G level of theory is essentially identical with that of HF/4-21G, when we make the additional assumption that the terminating methyl groups of the blocked GD molecule play no role at all stationary points.

In our previous study of the AD and ADA system,⁷ several HF/3-21G barriers between ADA minima were found to be much lower than that suggested by the Schäfer study⁵ of AD at HF/4-21G, where geometries were not fully relaxed in the latter work. Schäfer et al. argued⁵ that full geometry optimization is not necessary since the large (~ 18 – 29 kcal/mol) barriers found in partially optimized structures would only be lowered by several kcal/mol. However, we found that these same barriers are lowered by as much as 15 kcal/mol when geometries are fully relaxed.⁷ Although stricter tolerances were used in their more recent HF/4-21G GD study,¹¹ where geometries are optimized to within a few tenths of a kcal/mol, we have found this too lax to be reliable for determining stationary points (we have found barriers which are of this size between minima; see Tables I–IV). While Klimkowski and co-workers find four GD minima, the minimum designated IV (α_R) in their paper¹¹ has not been found in this work for GDA. Although they “point to the possibility that no barrier exists between α_R and β ,”¹¹ it is possible that the terminating methyl groups may contribute enough stabilization so that a very shallow HF/4-21G α_R minimum exists for GD, and not for GDA. To investigate this possibility we have performed Hartree-Fock geometry optimization on GD with Pulay's 4-21G basis,³¹ using the geometry in ref 10 as a starting point; we did not find a minimum corresponding to α_R . This is probably due to the extreme flatness of the surface near the region of β [3G] and IV, where quite small nuclear forces can give rise to large displacements. Thus with further optimization, we found that structure IV collapsed to the β structure.

We next consider an unusual feature of both the GDA and ADA relaxed surfaces, though for clarity we will only consider the GDA case in the following discussion. In Figure 3, in the region of $(-180^\circ, 0^\circ)$, we note the presence of a cusp ridge running diagonally through this point. The cusp is shown more clearly in Figure 5, which gives the energy as a function of moving perpendicular to the ridge through $(-180^\circ, 0^\circ)$. In the vicinity of this point, the surface is double-valued, with the cusp occurring at the point where the curves cross. Such discontinuities result

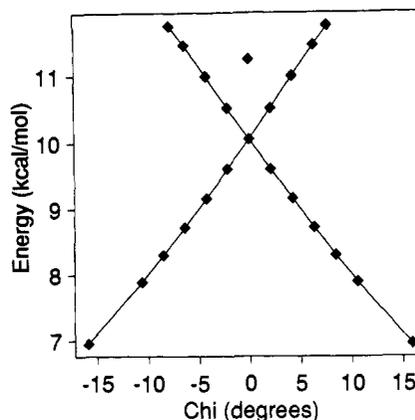


Figure 5. A two-dimensional cut through the glycine dipeptide analogue (ϕ, ψ) surface in the vicinity of the cusp. The variable χ which is plotted on the x axis, corresponds to moving perpendicular to the cusp ridge with $\phi = 2^{-1/2}\chi - 180^\circ$ and $\psi = 2^{-1/2}\chi$. Both the upper and lower surfaces are illustrated, and it can be seen that the region in which the surface is double-valued is quite small, about 15° in width. Additionally the relative HF/3-21G energy of the C_s (undistorted) $(-180^\circ, 0^\circ)$ structure is shown for comparison and is about 1.5 kcal mol⁻¹ above the properly relaxed curves.

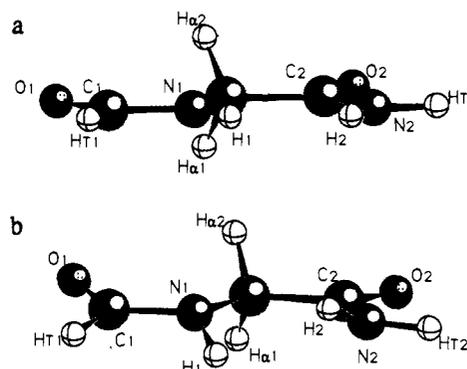


Figure 6. (a) The HF/3-21G undistorted C_s geometry of the $(-180^\circ, 0^\circ)$ conformation of the glycine dipeptide analogue. The structure is a six-membered ring with significant steric interactions between the two amide hydrogens, making it relatively high in energy (see Figure 5). (b) One of the two equivalent C_1 structures at the $(-180^\circ, 0^\circ)$ point ($\chi = 0^\circ$ on Figure 5). Distortions of the amide hydrogens from planarity partly relieve the steric interaction present in the C_s structure. Concerted (ϕ, ψ) changes that bring the two amide hydrogens closer to coplanar are energetically unfavorable and correspond to the high-energy tails of the curves in Figure 5. On the other hand (ϕ, ψ) changes that further separate the hydrogens are stabilizing and correspond to the low-energy branches in Figure 5.

from the reduction of the full space of many (flexible) degrees of freedom to a reduced two-dimensional space. In the case of ADA and GDA, conventional wisdom asserts that the dihedrals ϕ and ψ are the only soft dihedrals along the backbone, and the remaining torsions around the C–N peptide bond, with a calculated barrier of 18 kcal/mol³² (formamide model of the backbone), should be neglected. As Figure 5 shows, there are evidently two sets of values for the peptide torsion variables which are minima in the partial (ϕ, ψ) optimization.

In Figure 6a, we display the C_s structure of GDA at $(-180^\circ, 0^\circ)$, i.e. a conformation where the peptide units are planar. At this six-membered-ring geometry, large steric forces between the amide hydrogens might be expected. In fact, frequency calculations performed on this structure result in two imaginary frequencies—one in the space of the ϕ, ψ variables, and one in the space of all other variables. Hence this structure is not a valid point on the fully relaxed (ϕ, ψ) surface, since the energy is not a minimum in the space of the other variables. If we release the

(30) Ben-Naim, A. *Water and Aqueous Solutions: Introduction to a Molecular Theory*; Plenum: New York, 1974.

(31) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550–2560.

(32) Wiberg, K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1987**, *109*, 5935–5943.

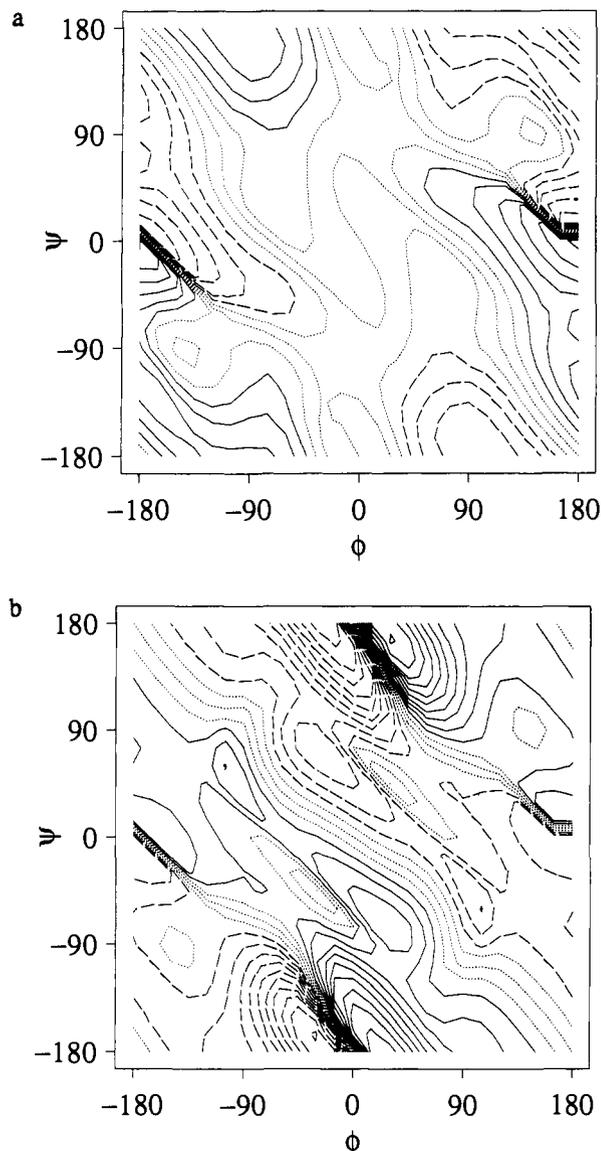


Figure 7. A contour map of dihedral deviations from planarity for rotations around the C_1-N_1 bond of GDA, where the atom labeling is that of Figure 2. The dihedral angles considered are (a) $\omega_1 = H_{\tau 1}-C_1-N_1-H_1$, which involves the amide hydrogen, and (b) $\omega'_1 = O_1-C_1-N_1-C_\alpha$, which does not. In both parts, the contours are in steps of 3° . The three dotted contours correspond to distortions of -3° , 0° , 3° while the dashed contours represent larger negative distortions (i.e. -6° , -9° , etc) and the solid contours represent larger positive distortions (i.e. 6° , 9° , etc).

C_s symmetry constraint (while still constraining ϕ and ψ), we find significant stabilizing distortions of the peptide backbone from planarity to form the C_1 structure, which is exhibited in Figure 6b. The energy of the C_s structure is higher than that of the distorted C_1 structure, (marked in Figure 5). A graphical illustration of the distortion from planarity for the GDA molecule is illustrated in Figures 7 and 8. In Figure 7a we display the deviation of the dihedral angle ω_1 ($H_{\tau 1}-C_1-N_1-H_1$) as a function of ϕ and ψ , which should be compared to the deviation in the angle ω'_1 ($O_1-C_1-N_1-C_\alpha$) as a function of ϕ and ψ displayed in Figure 7b, for the GDA molecule. Parts a and b of Figure 8 exhibit a similar comparison of dihedrals ω_2 ($C_\alpha-C_2-N_2-H_2$) and ω'_2 ($C_\alpha-C_2-N_2-H_{\tau 2}$) for the other peptide moiety of GDA. As is evident from these figures and Figure 6b, the cusp region shows large deviations from planarity for these peptide units in the region of $(-180^\circ, 0^\circ)$, by as much as 26° , with discontinuities at the cusp itself.

Evidently, the peptides twist about the C-N bond in order to relieve the steric contact between the two amide hydrogens. As one moves along the cusp ridge, these correspond to values of ϕ

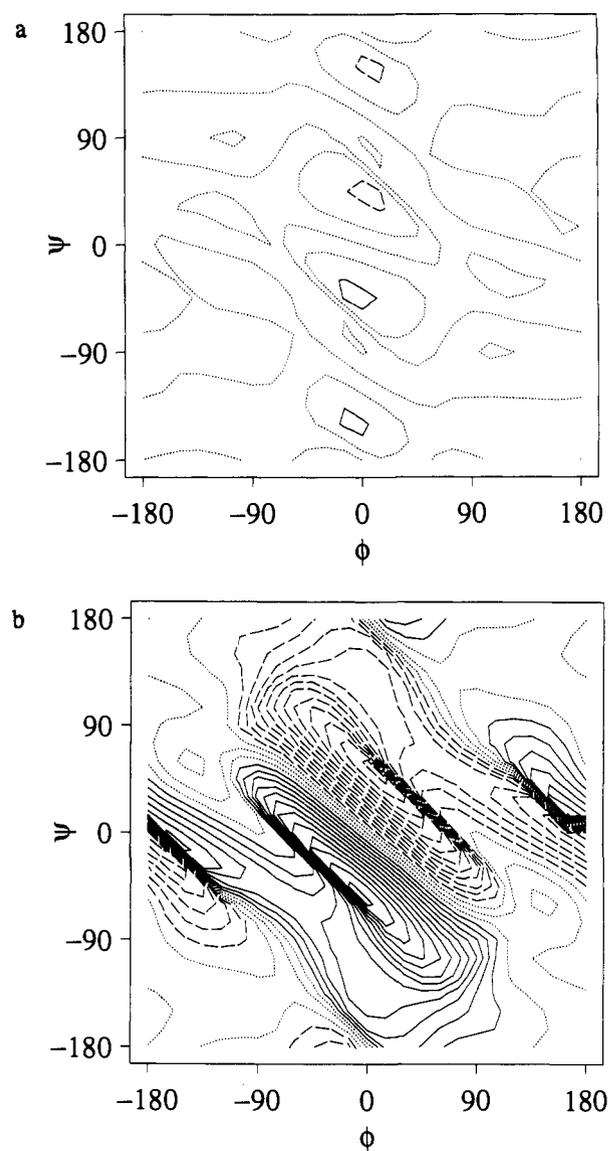


Figure 8. A contour map of dihedral deviations from planarity for rotations around the C_2-N_2 bond of GDA. The atom labeling is that of Figure 2, and the contours are as described in Figure 7. The dihedral angles plotted are (a) $\omega_2 = C_\alpha-C_2-N_2-H_{\tau 2}$ and (b) $\omega'_2 = C_\alpha-C_2-N_2-H_2$.

and ψ which result in bad hydrogen contacts, and where peptide distortions from planarity are large. There is also a symmetry to this distortion moving perpendicular to the ridge—in one case the hydrogen of the first peptide can distort above the plane of the C_s structure, while the other peptide hydrogen moves below the plane, and an equivalent second case where the two peptide hydrogens simply reverse their distortions. The cusp is where the curves corresponding to these distortions cross. In order to obtain the extent of the two different surfaces, a finer grid of geometry optimization was performed around the cusp region. As shown in Figure 5, the cusp is localized and the higher surfaces quickly collapse within 15° of the discontinuity back to a single minimized surface for the following reason. The ϕ and ψ values that correspond to moving perpendicular to the ridge on the high-energy curve tend to reintroduce the bad amide hydrogen contact, and further distortions of the peptide units from planarity are required to relieve this poor steric interaction. Finally, the distortions of the peptide groups on the high-energy curve becomes so large that the hydrogens flip positions relative to the plane of the ring. At this point ($\sim 7.5^\circ$ from the ridge) the low-energy curve is recovered, where the ϕ and ψ displacements away from the ridge relieve the steric interaction.

The unusual cusp present on the GDA map is also evident on the ADA surface as well (Figure 4). Furthermore, it does not

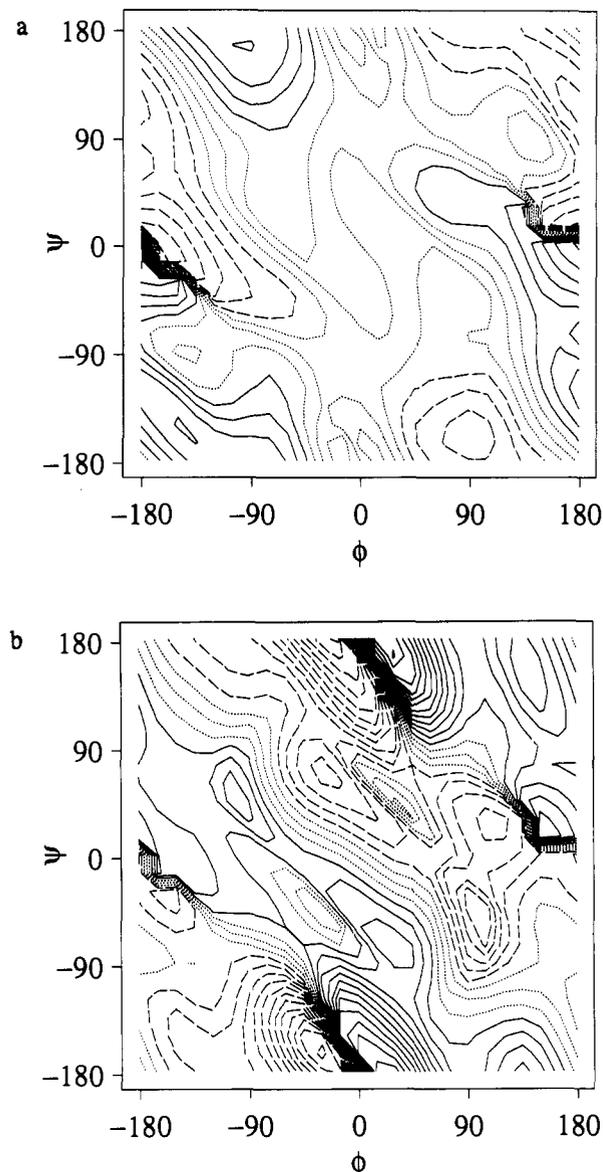


Figure 9. A contour map of dihedral deviations from planarity for rotations around the C_1-N_1 bond of ADA. The atom labeling is that of Figure 1, and the contours are as described in Figure 7. The dihedral angles plotted are (a) $\omega_1 = H_{11}-C_1-N_1-H_1$ and (b) $\omega'_1 = O_1-C_1-N_1-C_\alpha$.

seem to be an artifact of the HF/3-21G basis set. Geometry optimizations and frequency calculations at HF/6-31+G* on the GDA C_s structure also result in two imaginary frequencies—one in the space of ϕ and ψ , and one in the space of all other variables. Therefore we are confident that the structural distortion of the peptide backbone is a real effect.

Since conventional wisdom asserts that deviations from peptide planarity should be small, it is interesting to discuss other regions of (ϕ, ψ) space where significant peptide distortions occur. At the $(0^\circ, -60^\circ)$ point we observe a 40° deviation from peptide planarity (Figures 8b and 10b)—the largest found in this study. We also observe a 31° peptide distortion at the $(-15^\circ, 180^\circ)$ structure (see Figures 7b and 9b). Physically the distortions may arise as follows, although further study is required to fully confirm the following hypotheses. The $(0^\circ, 0^\circ)$ transition structure is a seven-membered ring that is stabilized by an attractive electrostatic interaction between O_1 and H_2 but destabilized by a strong steric interaction between these same two atoms. While the steric interaction is relieved at the $(0^\circ, -60^\circ)$ structure, there is also loss of the favorable electrostatic interaction. It is possible that the attractive interaction is recovered by allowing the H_2 atom to follow the oxygen atom, i.e. rotation about the C_2-N_2 bond by -40° . This demonstrates that some trade-off exists between the

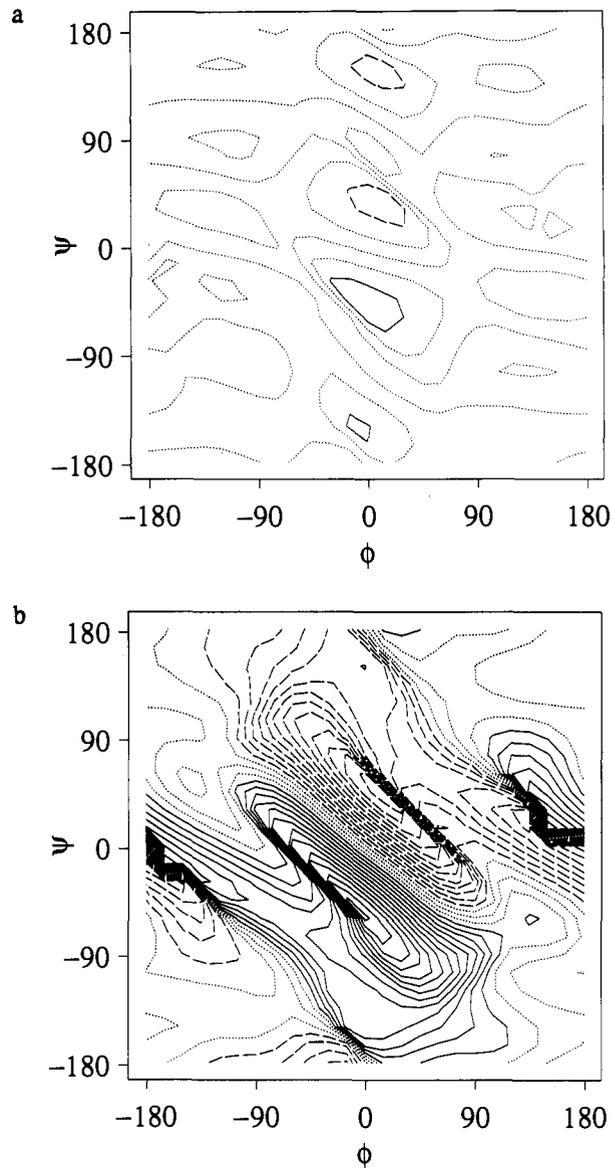


Figure 10. A contour map of dihedral deviations from planarity for rotations around the C_2-N_2 bond of ADA. The atom labeling is that of Figure 1, and the contours are as described in Figure 7. The dihedral angles plotted are (a) $\omega_2 = C_\alpha-C_2-N_2-H_{12}$ and (b) $\omega'_2 = C_\alpha-C_2-N_2-H_2$.

energy required to distort from peptide planarity and the energy gained in maintaining the favorable electrostatic interaction. This trade-off becomes apparent when we consider the $(-15^\circ, 180^\circ)$ point and $(0^\circ, 180^\circ)$ maximum; this region is high in energy due to a large steric interaction between atoms O_1 and O_2 . This bad contact is relieved by 31° distortions of the peptide groups at the $(-15^\circ, 180^\circ)$ structure, although there is considerable angle strain as well. However, at the maximum [11G] the peptide groups are found to be planar, and the molecule exhibits much greater angle strain (up to 14°); apparently a trade-off has been reached so that the required rotation about the C-N bonds is now too large (i.e. $>40^\circ$) and energetically expensive to aid in relieving the poor contact. Together with the cusp region, these results provide some indication that the shape of the peptide rotational potential may have some square-well-like character, i.e. distortions of up to 40° may be preferred to angle distortions, with a sudden onset of a large energy penalty for larger distortions.

It is interesting to note that many molecular mechanics (ϕ, ψ) surfaces of gas-phase ADA and GDA do not show a cusp feature, suggesting that the parametrized barrier to rotation about the C-N bond is too large, or that the absence of harmonic terms higher than $\cos(2\omega)$ in the peptide torsional potential yields an incorrect shape. It has been argued that parametrizing the gas-phase

properties of biological systems is not of direct physical interest,¹ and preference should instead be given to modeling solvated biomolecules. In this case, the empirical potentials may incorporate a "solvent effect": a stiffer barrier to rotation about the C–N bond to mimic the favorable delocalization of the planar form of the peptide in water. The use of such potentials becomes confusing when true gas-phase results are required, for example, when evaluating the gas-phase side of a closed thermodynamic cycle for a free energy of solvation calculation, when nonpolar solvents are being considered, or more generally whenever the effect of aqueous solvation is explicitly of interest.

V. The Alanine Dipeptide Surface and the Role of the Side Chain

The fully relaxed surface at HF/3-21G in 15° intervals for the ADA molecule in the direction of each dihedral angle is displayed in Figure 4. In our previous study⁷ we determined the presence of six conformational minima at HF/3-21G; since that time we have also isolated a seventh minimum at HF/3-21G which we have designated α_D [7A].⁷ Previous HF/4-21G studies of alanine dipeptide^{4,5} report an additional minimum denoted as α_R at (–92°, –6°), although these workers "indicate that no barrier exists at this computational level between α_R and β_2 ",⁵ suggesting it is not a true minimum. To clarify this point, we have performed a Hartree–Fock geometry optimization using Pulay's 4-21G basis set³¹ for the AD system using the α_R geometry in ref 4 as an initial guess. We find this structure is not a stationary point and that the optimization leads to the β_2 structure [4A], presumably for the same reasons discussed in section IV for GD. Although a reported force field analysis at the α_R geometry¹⁵ is therefore formally ill-defined, it is still useful for the intended purpose of analyzing the high-frequency bond and angle modes.¹⁵

We also note that the transition structure at (0°, 0°) in Figure 4 is a saddle point at HF/3-21G, just as it is for GDA. However, no such transition structure could be found at the HF/6-31+G* level. On the basis of the greater reliability of the HF/6-31+G* energies, we believe that the (0°, 0°) structure is probably a maximum flanked by the two saddle points [13A] and [17A] and not a saddle point itself, as predicted by HF/3-21G. Furthermore, the α_D [7A], α_L [5A], and β_2 [4A] minima characterized at HF/3-21G effectively vanish at the more reliable HF/6-31+G* level of theory; the α_D [7A] minimum disappears completely, while miniscule barriers separate α_L [5A] and β_2 [4A] from other minima. The $C7_{ax}$ to α_L [10A] and $C7_{eq}$ to β_2 [8A] barriers are 0.03 and 0.02 kcal/mol, respectively. We therefore regard the hydrogen-bonding conformations $C7_{eq}$ [1A], $C7_{ax}$ [3A], and C5 [2A] and the α' [6A] structure as the only stable minima on the ADA surface. Thus as for the GDA system there are no significant minima in the regions of conformation space corresponding to secondary structure in larger peptides. We note here that the α' conformer is perhaps a poor choice of name,⁴ since its position in (ϕ, ψ) space is outside acceptable low-energy regions corresponding to secondary structure.

We next turn to the role of the side chain and how it manifests itself in stabilizing or destabilizing particular conformations. Simple analyses, such as Ramachandran maps,⁸ suggest that glycine dipeptide (with no side chain) has a significantly larger accessible conformation space than alanine dipeptide (with a methyl side chain). Thus the role of the methyl side chain can be assessed by contrasting our results for this pair of amino acids. In addition, we are concerned with the physical origins of any differences between the two surfaces. By first assessing the maps themselves and then going on to a closer examination of the distortions of internal degrees of freedom from their optimal values as a function of ϕ and ψ , we hope to gain insight as to the role of the methyl side chain on conformational accessibility. Such effects may be even more dramatic for other amino acids with bulkier side chains.

A general description of the GDA (ϕ, ψ) surface (Figure 3) is a large low-energy region comprising the C7 [1G] and C5 [2G] structures, with very small barriers to interconversion. The side chain modifies the GDA surface by stabilizing new stationary

points such as α' [6A] and destabilizing one of the degenerate C7 regions of GDA to give rise to the less stable ADA $C7_{ax}$ [3A] minimum. The $C7_{eq}$ [1A] and C5 [2A] regions of the two surfaces at HF/3-21G are seen to be equivalent. In addition, the HF/6-31+G* results predict the presence of the seven-membered-ring and five-membered-ring hydrogen-bonding minima only, for both GDA and ADA. Therefore, the two surfaces in the region of $C7_{eq}$ [1A, 1G] and C5 [2A, 2G] may reliably be considered equivalent between the ADA and GDA molecules, so that the methyl side chain plays no significant role in this region. Inspection of the ADA (ϕ, ψ) surface reveals that the main differences are in the axial region ($\phi \geq 0^\circ$).

A comparison of the dihedral geometries as a function of ϕ and ψ is useful to elucidate the effects of the methyl group on the glycine surface. While deviations from planarity of *both* peptide groups were found to contribute to the formation of the cusp, only the peptide group dihedrals involving the dihedrals defined by the C_1-N_1 bond are found to differ significantly in the high-energy island region of ADA. This is demonstrated by a comparison of these dihedrals for GDA (Figure 7) and ADA (Figure 9). It is clear from a comparison of the C_2-N_2 dihedrals for GDA (Figure 8) and ADA (Figure 10) that no steric contacts are present between the methyl group and the other peptide moiety in this region. It is again illustrated in Figures 11 and 12, where we compare the change in the ADA improper dihedrals ($C_\alpha-C_2-N_1-C_\beta$) and ($C_\alpha-C_2-N_1-H_\alpha$), with the GDA improper dihedrals ($C_\alpha-C_2-N_1-H_{\alpha 1}$) and ($C_\alpha-C_2-N_1-H_{\alpha 2}$), which describe distortions away from sp^3 hybridization for carbon. Deviations of up to 6° from optimal hybridization are evident in the $C7_{ax}$ [3A] region of (ϕ, ψ) space for ADA, which are not present for the GDA improper dihedral. The deviations from peptide planarity and sp^3 hybridization exhibited in Figures 7–12 are due to the steric interaction between the O_1 oxygen carbonyl and a side chain methyl hydrogen at the ϕ position of 120° for all ψ values.

The anticipated effect of this bad contact on the right-hand side of the GDA map is to raise the energy of all structures with $\phi = 120^\circ$ by some amount. Comparing the GDA and ADA maps (Figures 3 and 4) we see that the differences are qualitatively resolved in this way: the GDA β [3G] region is divided by the $\phi = 120^\circ$ high-energy ribbon, giving rise to the ADA α' [6A] structure. Additionally, the second C7 structure ($C7_{ax}$ [3A] of ADA) is more localized, and cut off from the C5 [2A] region because of the formation of the bad contact in the transition structure. Ramachandran maps qualitatively show this effect, although they greatly exaggerate it:⁸ in one instance, the $C7_{ax}$ [3A] structure vanished.³

VI. Conclusion

In this paper we have presented a detailed ab initio study of the conformational preferences of model glycine and alanine dipeptides. Our results include fully relaxed (ϕ, ψ) maps for both peptides at a fairly low level of theory (HF/3-21G) plus detailed investigation of minima and transition structures at higher levels of theory (HF/6-31+G*) with inclusion of correlation effects for relative energies (MP2/6-31+G**). Our principal conclusions are as follows.

Optimized structures obtained at the HF/3-21G level are in generally good agreement with the HF/6-31+G* results although several shallow minima corresponding to secondary structure disappear at the higher level of theory. Thus, the only minima on the GDA (ϕ, ψ) surface are the C5 and C7 hydrogen-bonded structures; on the ADA surface there are the hydrogen-bonding conformers C5, $C7_{eq}$, and $C7_{ax}$, as well as the high-energy α' minimum discussed below. Barriers to interconversion are also significantly smaller with the larger basis set. In general, the HF/6-31+G* surface seems to be somewhat flatter and more featureless than the HF/3-21G results. Correlation effects on relative energies are not large—less than 1.1 kcal/mol.

The results obtained here, particularly the higher level calculations, form a useful database against which to test aspects of the performance of various empirical and semiempirical methods.

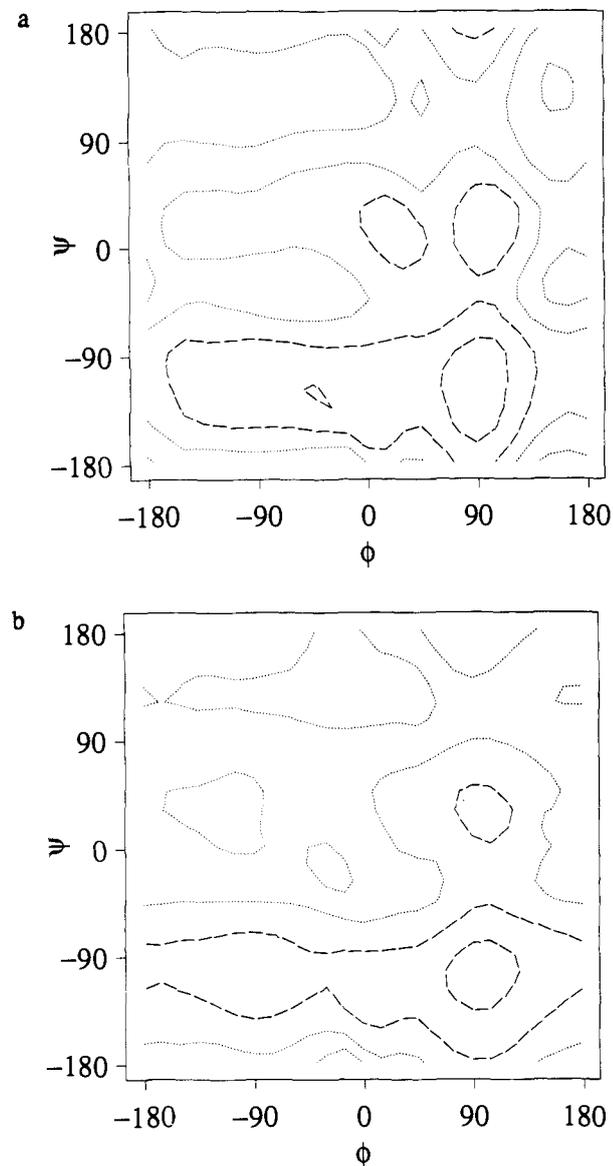


Figure 11. A contour map of dihedral deviations from optimal carbon sp^3 hybridization for ADA. The atom labeling is that of Figure 1, and the contours are as described in Figure 7. The dihedral angles plotted are (a) $\tau = C_\alpha-C_2-N_1-C_\beta$ and (b) $\tau' = C_\alpha-C_2-N_1-H_\alpha$.

We shall report such a study shortly, involving the semiempirical AM1 method and selected molecular mechanics potentials. We note that while these dipeptides do exhibit intramolecular hydrogen-bonded structures, real tests of secondary structure formation require longer peptides.

Our fully relaxed (ϕ, ψ) surfaces reveal for the first time in these systems the presence of a cusp (which is also confirmed at the higher HF/6-31+G* level of theory), due to portions of the relaxed map being double-valued. The discontinuity occurs where the two surfaces cross near $(-180^\circ, 0^\circ)$, corresponding to large distortions of the peptide group to relieve amide hydrogen steric interactions. We also observe deviations from peptide planarity of up to 40° in other regions of the (ϕ, ψ) map as well. The discontinuity and peptide distortion results presented here show the limitations of regarding the conformational space of these molecules as two dimensional.

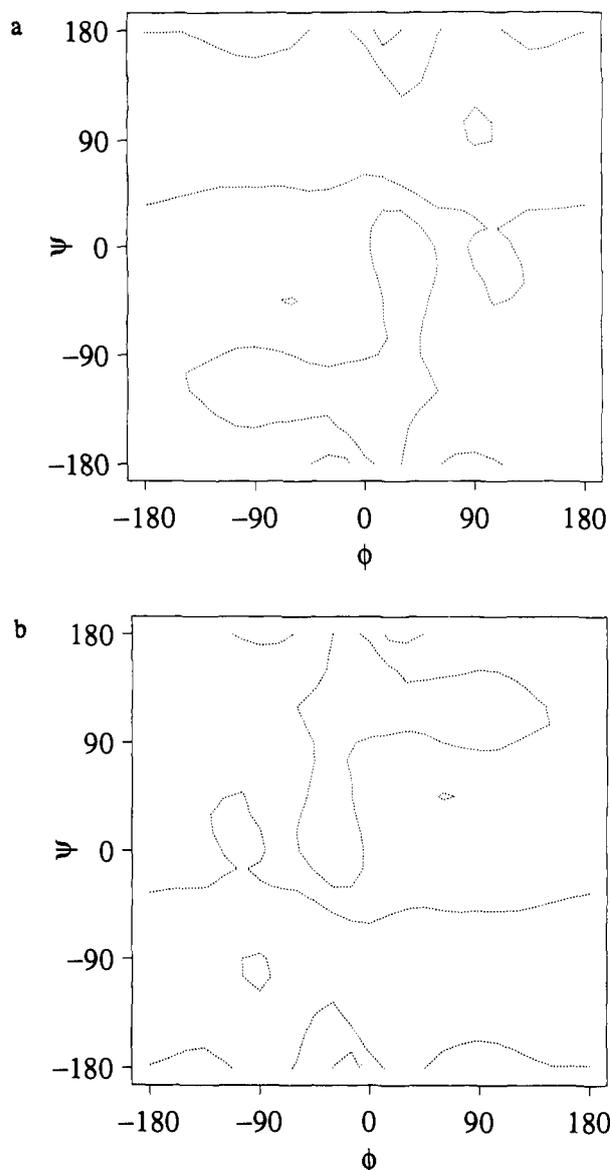


Figure 12. A contour map of dihedral deviations from optimal carbon sp^3 hybridization for GDA. The atom labeling is that of Figure 2, and the contours are as described in Figure 7. The dihedral angles plotted are (a) $\tau = C_\alpha-C_2-N_1-H_{\alpha 1}$ and (b) $\tau = C_\alpha-C_2-N_1-H_{\alpha 2}$.

Comparisons of our results for the alanine and glycine dipeptide systems allows us to assess the steric role of the alanine methyl side chain. A steric interaction between the methyl group and the terminal carbonyl oxygen destabilizes all structures with $\phi \cong 120^\circ$, as predicted in exaggerated form by unrelaxed Ramachandran maps. This interaction sterically generates the high-energy α' minimum of alanine dipeptide and partially isolates the axial seven-membered-ring structure from the other low-energy minima.

Acknowledgment. C.L.B. and T.H.G. gratefully acknowledge the support of the NIH (GM-37554-02). J.A.P. and M.H.G. thank the NSF for support (CHE-84-09405). We also thank the Pittsburgh Supercomputing Center and Cray Research for computational resources for a portion of the results presented.

Registry No. (S)-OHCNHCH(Me)CONH₂, 54046-46-7; OHCNH-CH₂CONH₂, 4238-57-7.