

that nonadiabatic transitions become probable events even at low collision energy.<sup>39,40</sup>

Our study of acetone cation CID provides evidence for analogous behavior of a polyatomic ion at low collision energy. In this system, we identified an unexpectedly long-lived excited state that is scattered inelastically and dissociated in low-energy collisions.<sup>17</sup> We have measured its lifetime and the threshold energy<sup>42</sup> for inducing its nonadiabatic transition to the ground state, which then dissociates. We have also demonstrated the microscopic reversibility of the process as inelastic scattering in both bound-bound and bound-unbound transitions.<sup>43</sup> We suggested that a nonadiabatic curve crossing mechanism at small He-M<sup>+</sup> distances was responsible for this CID reaction at low collision energy and that higher excited states were excited in a similar mechanism at high collision energy. The nitromethane cation CID reactions reported here appear to be a second example of this kind of reaction dynamics for polyatomic ions. These results further demonstrate the hitherto unexpected importance of electronic excitation mechanism in the collisional activation/dissociation of polyatomic ions at moderate collision energies.

### Conclusions

This investigation of nitromethane ion CID demonstrates that the formation of NO<sub>2</sub><sup>+</sup> and NO<sup>+</sup> proceeds via different reaction pathways. NO<sub>2</sub><sup>+</sup> ions are mostly formed from ions excited by an impulsive mechanism occurring via small impact parameter collisions and accompanied by large angular momentum exchange. NO<sup>+</sup> ions are produced from at least two quite different reaction

channels. The lowest energy dissociation path generates NO<sup>+</sup> with modest energy deposition at zero scattering angle; we suggest this mechanism involves rearrangement of internally excited ions having structures approaching that of the transition state for forming methyl nitrite ions prior to the collisional activation step. The high-energy mechanism for generating NO<sup>+</sup> by deposition of 5.5-5.7 eV proceeds from an excited-state surface of the nitromethane ions.

We conclude that the CID of nitromethane ion to NO<sup>+</sup> and CH<sub>3</sub>O<sup>+</sup> is the second example we have found for remarkably efficient conversion of translational energy into electronic excitation in the low and medium collision energy ranges and inefficient randomization of internal energy. Like acetone, it is an example of a polyatomic ion exhibiting the "weak assumption" QET behavior<sup>44</sup> in which a bottleneck to energy randomization prevents intramolecular relaxation to the ground state prior to dissociation. Nitromethane ion is a rather special case in other ways that suggest it exhibits several bottlenecks at branch points for nuclear rearrangements leading to specific dissociation mechanisms which are a strong function of internal energy.

**Acknowledgment.** Support of this research by the National Science Foundation, Grant No. CHE-8312069, is gratefully acknowledged. We thank Professor Tomas Baer for a very useful discussion of nuances in the unimolecular decay of the methyl nitrite cation and for sharing his data prior to publication.

Registry No. CH<sub>3</sub>NO<sub>2</sub><sup>+</sup>, 74868-54-5; CH<sub>3</sub>ONO<sup>+</sup>, 85232-65-1.

(42) Fenstein, S.; Futrell, J.; Heninger, M.; Marx, R.; Mauclaire, G.; Yang, Y. *Chem. Phys. Lett.* **1991**, *179*, 125.

(43) Qian, K.; Shukla, A.; Futrell, J. *Chem. Phys. Lett.* **1990**, *175*, 51.

(44) Lorquet, J. C. *Org. Mass Spectrom.* **1981**, *16*, 469.

(45) Kandel, R. J. *J. Chem. Phys.* **1955**, *23*, 84.

(46) Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. *J. Phys. Chem. Ref. Data*, **Suppl.** **1988**, *17*.

## Ab Initio SCF Calculations on Low-Energy Conformers of N-Acetyl-N'-methylalaninamide and N-Acetyl-N'-methylglycinamide

Hans-Joachim Böhm\* and Stefan Brode

Contribution from BASF AG, Central Research and Informatics, 6700 Ludwigshafen, Germany.  
Received March 27, 1990

**Abstract:** Results from ab initio SCF calculations with a double- $\zeta$  plus polarization (DZP) basis set on five low-energy conformations of N-acetyl-N'-methylalaninamide ("dialanine") and three conformations of N-acetyl-N'-methylglycinamide ("diglycine") are reported. In addition, data from calculations with a triple- $\zeta$  plus polarization (TZP) basis set on two conformations of diglycine are given. The results are used to assess the quality of current force fields aimed at peptides and proteins. The present ab initio data point to some important deficiencies of current protein force fields with respect to relative energies of dipeptide conformers.

### 1. Introduction

Accurate force fields are one of the most important tools in molecular modeling. Especially force fields for proteins and peptides are heavily needed in biochemical and drug design applications. Thus, several force fields such as AMBER,<sup>1</sup> ECEPP,<sup>2,3</sup> CVFF,<sup>4,5</sup> or CHARMM<sup>6</sup> have been proposed and are now widely

used in force field calculations. However, the validation of a force field by comparing calculated energies and geometries with experimental data or high-quality quantum chemical ab initio calculations is difficult for several reasons.

Experimental structural information on peptides and proteins is only available for the condensed phase. Such a system is

(1) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghlo, C.; Alagona, G.; Profeta, S.; Welner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.

(2) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.

(3) Zimmerman, S. S.; Pottle, M. S.; Nemethy, G.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1.

(4) Stern, P. S.; Chorev, M.; Goodman, M.; Hagler, A. T. *Biopolymers* **1983**, *22*, 1885.

(5) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Proteins* **1988**, *4*, 31.

(6) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(7) Pettitt, B. M.; Karplus, M. *Chem. Phys. Lett.* **1985**, *121*, 194.

characterized by a large number of different bonded and nonbonded interactions that all contribute significantly to the stabilization or destabilization of the structure. The corresponding potential surface will therefore show a large number of local minima. If an energy minimization is performed by starting from the experimentally observed structure, the system will adopt the conformation of the nearest local minimum, which is not necessarily the global minimum. The deviation between calculated and experimental structures is therefore difficult to interpret, because good agreement may be incidental and/or result from a cancellation of errors. Moreover, experimental information on relative stabilities of different peptide conformations is sparse. In principle, high-quality ab initio calculations can be used to provide this information. However, due to the very high computational requirements both in cpu time and intermediate storage, such calculations were not feasible for most laboratories in the past.

*N*-Acetyl-*N'*-methylalaninamide ("dialanine") and *N*-acetyl-*N'*-methylglycinamide ("diglycine") have attracted considerable interest in their conformational properties because they can be viewed as model dipeptides. These molecules can serve as test cases for force fields derived for peptides and proteins: Several studies have been performed to determine geometries and energies of several conformations of these molecules using force field methods<sup>1-4,6</sup> or quantum chemical calculations.<sup>8-12</sup> Perhaps most closely related to the present study are ab initio SCF calculations at the 4-21G level on dialanine and on diglycine by L. Schäfer et al.<sup>9-12</sup> Although this basis set yields generally good results for bond lengths and angles of organic molecules, there are well-known deficiencies for systems containing hydrogen bonds.<sup>13</sup>

Recent progress in computer hardware and the development of the direct SCF method<sup>14-16</sup> now allow for the use of larger basis sets. In order to assess the accuracy of the current force fields, we have carried out a series of ab initio SCF calculations employing a double- $\zeta$  plus polarization basis set (DZP) for five conformations of dialanine and three conformations of diglycine. The inclusion of polarization functions is known to give a better description of the charge distribution in the molecule, thus improving the accuracy of the electrostatic and polarization contributions to the nonbonded interactions. In order to assess the basis set dependence of the present results, also a calculation using a triple- $\zeta$  plus polarization basis set (TZP) was carried out on two conformers of diglycine.

The next section describes the details of the ab initio calculations. Section 3 compares the present ab initio results with the data obtained from force field calculations. Summary and conclusions are given in section 4.

## 2. Computational Methods

The program TURBOMOLE developed by Häser and Ahlrichs<sup>15-16</sup> running on the computers Convex C210 and Silicon Graphics 4D20 was used in all calculations. The program employs the direct SCF method, allowing large calculations to be carried out with only modest disk space requirements. In the present calculations, we employed the following DZP and TZP basis sets (GTO primitive sets taken from Huzinagas tables<sup>17</sup>): DZP, C (8,4,1)/[4,2,1], N (8,4,1)/[4,2,1], O (8,4,1)/[4,2,1], H (4,1)/[2,1]; TZP, C (9,5,1)/[5,3,1], N (9,5,1)/[5,3,1], O (9,5,1)/[5,3,1], H (5,1)/[3,1].

The exponents of the polarization functions were chosen as 0.8 (d on C), 1.0 (d on N), 1.2 (d on O), and 0.8 (p on H). The calculation comprised a total number of 210 basis functions for dialanine and 185

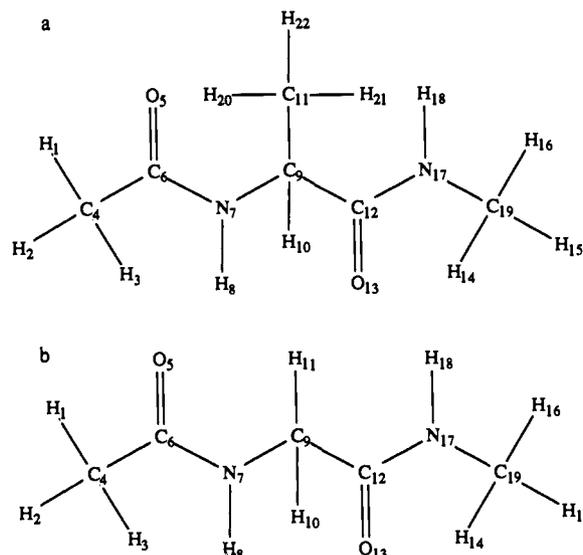


Figure 1. Atom numbering scheme for the molecules *N*-acetyl-*N'*-methylalaninamide and -glycinamide. The torsion angles discussed in the text are defined as follows:  $\omega_1 = \angle C4C6N7C9$ ,  $\omega_2 = \angle C9C12N17C19$ ,  $\phi = \angle C6N7C9C12$ ,  $\psi = \angle N7C9C12N17$ .

(DZP) or 231 (TZP) basis functions for diglycine.

The calculations were carried out as follows. The various conformers were initially generated by using interactive computer graphics.<sup>18</sup> The torsion angles  $\phi$  and  $\psi$  (see Figure 1 for definition of torsion angles), which essentially define the conformation of the dipeptides,<sup>19</sup> were set to the values of the local minima obtained from the force field AMBER.<sup>1</sup> For dialanine, there are five local minima on the AMBER potential surface denoted as  $C_{7,eq}$ ,  $C_{7,ax}$ ,  $\alpha_R$ ,  $\alpha_L$ , and  $C_5$ . Two of them contain an intramolecular hydrogen bond and are denoted  $C_{7,eq}$  and  $C_{7,ax}$  (with the  $CH_3$  side chain oriented equatorially or axially with respect to the seven-membered ring formed by the hydrogen bond). The  $C_5$  conformer is fully extended, and the  $\alpha_R$  and  $\alpha_L$  conformers adopt the right- and left-handed helical conformation. For diglycine, there exist three local minima on the AMBER potential surface denoted as  $C_7$ ,  $\alpha$ , and  $C_5$ . A semiempirical energy minimization of the bond lengths and angles was then carried out with the MNDO method using the AM1 parametrization.<sup>20</sup> The torsion angles were kept fixed during this calculation. (Note that the helical conformations are no local minima within the MNDO-AM1 approximation. The energy minimization starting at the  $\alpha$  conformations yields the  $C_7$  conformation.) The AM1-minimized structures were used as input to the ab initio SCF calculations. A full optimization of all degrees of freedom was carried out in TURBOMOLE. The internal coordinates were defined as suggested by Pulay et al.<sup>22</sup> The minimization was continued until the maximum energy gradient dropped below  $0.1 \text{ kJ mol}^{-1} \text{ \AA}^{-1}$  and the energy change over the last 20 iterations was less than  $0.04 \text{ kJ/mol}$ . Between 50 and 100 SCF + gradient steps were necessary to achieve convergence. The fully optimized structures at the DZP level were used as input to the TZP calculations. Apart from the five conformations of dialanine and three conformations of diglycine minimized in the present study, no attempt was made to find further local minima of the potential surface.

The wave functions calculated for the fully minimized structures of diglycine were used to calculate the electrostatic potential at grid points surrounding the molecule. The set of points was a cubic grid of  $0.75 \text{ \AA}$  spacing in a  $1\text{-\AA}$ -thick shell around the molecule. The inner surface of this shell was  $4 \text{ \AA}$  away from the molecule. The electrostatic potential was calculated at 1278 points for the  $C_5$  conformer, at 1289 points for the  $C_7$  conformer, and at 1342 points for the  $\alpha$  conformer. A set of fractional charges positioned at the nuclei was then fitted to the potential.

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

(9) Klimkowski, V. J.; Schäfer, L.; Momany, F. A.; van Alsenoy, C. *J. Mol. Struct.* **1985**, *124*, 143.

(10) Schäfer, L.; van Alsenoy, C.; Klimkowski, V. J.; Scarsdale, J. N. *J. Chem. Phys.* **1982**, *76*, 1439.

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

(12) Schäfer, L.; Klimkowski, V. J.; Momany, F. A.; Chuman, H.; van Alsenoy, C. *Biopolymers* **1984**, *23*, 2335.

(13) Frisch, M. J.; DelBene, J. E.; Binkley, J. S.; Schaefer, H. F., III. *J. Chem. Phys.* **1986**, *84*, 2279.

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

(15) Häser, M.; Ahlrichs, R. *J. Comput. Chem.* **1989**, *10*, 104.

(16) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.

(17) Huzinaga, S. *J. Chem. Phys.* **1965**, *42*, 1293.

(18) Program INSIGHT; Biosym Technologies, 10065 Barnes Canyon Road, San Diego, CA 92121.

(19) Ramachandram, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 195.

(20) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

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

(22) Chirlian, L. E.; Miller Francl, M. *J. Comput. Chem.* **1987**, *8*, 894.

(23) As implemented in the program DISCOVER; Biosym Technologies, San Diego, CA 92121.

Table I. Ab Initio DZP Results on Dialanine<sup>a</sup>

	conformation			
	C <sub>7,eq</sub>	C <sub>7,ax</sub>	C <sub>5</sub>	α <sub>L</sub>
rel energy (kJ/mol)	0.0	12.6	2.1	18.5
φ (deg)	-85.9	75.8	-156.0	65.9
ψ (deg)	79.1	-58.9	161.0	33.5
ω <sub>1</sub> (deg)	-178.3	174.1	177.6	166.0
ω <sub>2</sub> (deg)	-174.1	-178.8	180.0	-176.1
R <sub>C4-C6</sub> (Å)	1.509	1.510	1.510	1.510
R <sub>C6-N7</sub> (Å)	1.348	1.348	1.347	1.364
R <sub>C6-O5</sub> (Å)	1.201	1.202	1.199	1.192
R <sub>N7-C9</sub> (Å)	1.455	1.461	1.440	1.457
R <sub>C9-C11</sub> (Å)	1.519	1.528	1.532	1.524
R <sub>C9-C12</sub> (Å)	1.534	1.534	1.526	1.533
R <sub>C12-CO13</sub> (Å)	1.198	1.199	1.194	1.198
R <sub>C12-N17</sub> (Å)	1.345	1.340	1.346	1.348
R <sub>N17-C19</sub> (Å)	1.443	1.443	1.444	1.443
R <sub>N7-H8</sub> (Å)	0.997	0.996	0.997	0.996
R <sub>N17-H18</sub> (Å)	1.000	1.000	0.997	0.995
∠C4C6N7 (deg)	116.2	115.2	115.8	115.6
∠C6N7C9 (deg)	123.1	122.5	122.3	127.1
∠N7C9C12 (deg)	109.9	113.1	107.5	113.9
∠C9C12N17 (deg)	114.4	116.0	115.1	117.0
∠C12N17C19 (deg)	121.9	121.3	122.6	121.3

<sup>a</sup>The total energy of the lowest energy conformation C<sub>7,eq</sub> is -492.8309987 au. Note that the α<sub>R</sub> conformer is not a local minimum.

In a first series of least-square fits, all charges were allowed to vary. In a second set of fits, all charges on nonpolar hydrogen (those bound to carbon) were set to zero and charges at similar atoms were kept at identical values. For example, in this second approach, five groups of atoms ((C<sub>4</sub>, C<sub>9</sub>, C<sub>19</sub>), (O<sub>5</sub>, O<sub>13</sub>), (C<sub>6</sub>, C<sub>12</sub>), (N<sub>7</sub>, N<sub>17</sub>), (H<sub>8</sub>, H<sub>18</sub>)) are forced to carry the same partial charges, thus leading to four independent adjustable parameters for the fit. Furthermore, an attempt was made to fit the potentials of all three diglycine conformers simultaneously by a single set of charges. The present approach is similar to the work of Chirlian et al.<sup>22</sup>

### 3. Results

**3.1. Dialanine.** The results from the ab initio SCF calculations on five conformations of dialanine using a DZP basis set are shown in Table I. Plots of the minimized structures of dialanine are given in Figure 2. The lowest energy conformation found is the conformation C<sub>7,eq</sub> with φ = -86° and ψ = 79°. It is characterized by a strained intramolecular hydrogen bond with ∠N<sub>17</sub>H<sub>18</sub>...O<sub>5</sub> = 139° and R<sub>H<sub>18</sub>...O<sub>5</sub></sub> = 2.22 Å. Both peptide groups are slightly tilted, with ω<sub>1</sub> = -178° and ω<sub>2</sub> = -174°. The C<sub>5</sub> conformation (φ = -156°, ψ = 161°) is 2.1 kJ/mol less stable than C<sub>7,eq</sub>. The C<sub>5</sub> intramolecular hydrogen bond geometry is characterized by R<sub>H<sub>18</sub>...O<sub>5</sub></sub> = 2.21 Å and ∠N<sub>17</sub>H<sub>18</sub>...O<sub>5</sub> = 105°. C<sub>7,ax</sub> (φ = 76°, ψ = -59°) is 12.6 kJ/mol less stable than C<sub>7,eq</sub>. Interestingly the intramolecular hydrogen bond in C<sub>7,ax</sub> is shorter than in C<sub>7,eq</sub>, with R<sub>H<sub>18</sub>...O<sub>5</sub></sub> = 2.03 Å and ∠N<sub>17</sub>H<sub>18</sub>...O<sub>5</sub> = 146°.

The helical conformation α<sub>R</sub> is not a local minimum on the DZP ab initio potential surface of dialanine. In the course of the minimization, the energy decreases rapidly to 13.5 kJ/mol above the C<sub>7,eq</sub> energy while the right-handed helical conformation (φ = -63°, ψ = -34°) is maintained. Then there is a slow change of the conformation toward the C<sub>7,eq</sub> conformation. The maximum residual energy gradient along this α<sub>R</sub> → C<sub>7,eq</sub> pathway is 5 kJ mol<sup>-1</sup> Å<sup>-1</sup>. In contrast to the right-handed helical conformation, the left-handed conformation α<sub>L</sub> (φ = 66°, ψ = 33°) is a local minimum on the DZP ab initio potential surface with an energy of 18.5 kJ/mol above the global minimum.

The dependence of bond lengths and bond angles on the conformation is very small. For example, R<sub>N<sub>17</sub>H<sub>18</sub></sub> varies from 0.995 to 1.000 Å depending on whether it is involved in a hydrogen bond or not. The largest change of a bond length occurs for R<sub>N<sub>7</sub>C<sub>9</sub></sub>, which varies from 1.440 (C<sub>5</sub>) to 1.461 Å (C<sub>7,ax</sub>). The largest change of a bond angle is found for ∠N<sub>7</sub>C<sub>9</sub>C<sub>12</sub>, which varies from 107.5° (C<sub>5</sub>) to 113.9° (C<sub>7,ax</sub>).

Additional structural details may be derived from the atomic Cartesian coordinates of the four local minima of dialanine as given in Table II.

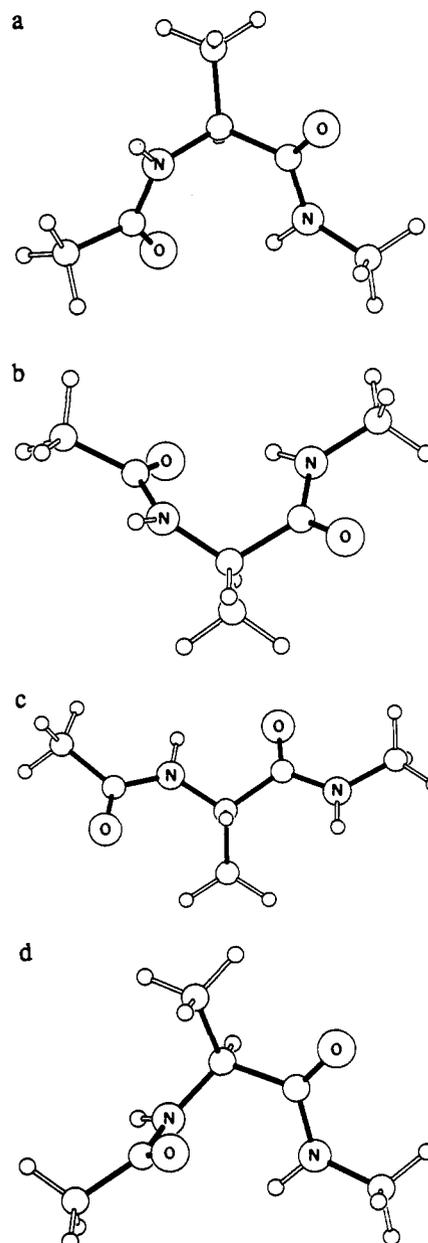


Figure 2. Molecular structure of the four low-energy conformations C<sub>7,eq</sub> (a), C<sub>7,ax</sub> (b), C<sub>5</sub> (c), and α<sub>L</sub> (d) of dialanine.

**3.2. Diglycine.** The ab initio DZP results on the three conformations C<sub>7</sub>, C<sub>5</sub>, and α of diglycine are shown in Table III. Plots of the minimized structures are given in Figure 3. Note that, for diglycine, C<sub>7,eq</sub> and C<sub>7,ax</sub> as well as α<sub>R</sub> and α<sub>L</sub> are identical. The lowest energy conformation found is the C<sub>5</sub> conformation with φ = -180° and ψ = 180°. The C<sub>7</sub> conformation (φ = -85°, ψ = 75°) is slightly less stable than C<sub>5</sub> by 1.3 kJ/mol. The C<sub>7</sub> strained intramolecular hydrogen bond is characterized by ∠N<sub>17</sub>H<sub>18</sub>...O<sub>5</sub> = 140° and R<sub>H<sub>18</sub>...O<sub>5</sub></sub> = 2.20 Å. The α conformation is not a local minimum on the DZP ab initio potential surface. We have carried out a constrained geometry optimization of the conformer with φ and ψ kept fixed at the values of the corresponding AMBER structure<sup>1</sup> (φ = -60°, ψ = -39°). With use of these values, the α conformation is 16.4 kJ/mol less stable than C<sub>5</sub>. The largest deviation from planarity of the peptide groups occurs for the helical conformation with ω<sub>1</sub> = -170° and ω<sub>2</sub> = 175°.

The variation of bond lengths and angles upon conformational change is very similar to that of dialanine. The largest change of a bond length occurs for R<sub>N<sub>7</sub>C<sub>9</sub></sub>, which varies from 1.431 (C<sub>5</sub>) to 1.448 Å (C<sub>7</sub>). The largest change of a bond angle is found for ∠N<sub>7</sub>C<sub>9</sub>C<sub>12</sub>, which varies from 109.4° (C<sub>5</sub>) to 115.7° (α).

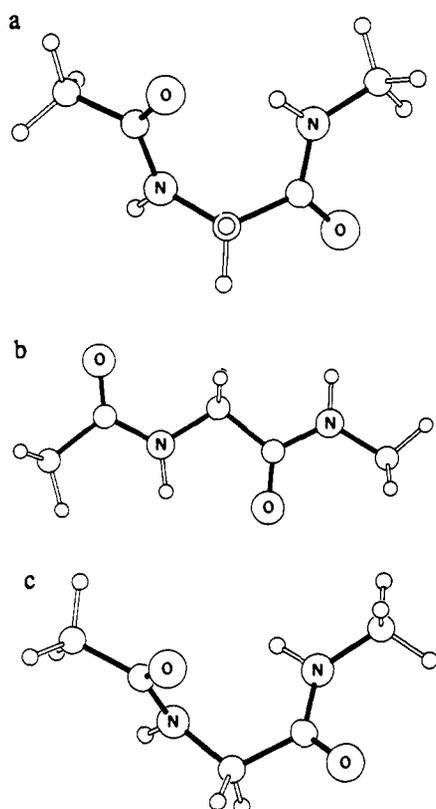


Figure 3. Molecular structure of the three low-energy conformations  $C_7$  (a),  $C_5$  (b), and  $\alpha$  (c) of diglycine.

The ab initio TZP results on the  $C_7$  and  $C_5$  conformer of diglycine are also given in Table III.  $C_5$  is 1.7 kJ/mol more stable than  $C_7$  at the TZP level. The geometric differences between the DZP and the TZP minimized structures are very small. The intramolecular hydrogen bond is slightly longer, with  $R_{\text{H18} \cdots \text{O5}} = 2.21 \text{ \AA}$  ( $\angle_{\text{N17H18} \cdots \text{O5}} = 105^\circ$ ) for the  $C_5$  conformer and  $R_{\text{H18} \cdots \text{O5}} = 2.24 \text{ \AA}$  ( $\angle_{\text{N17H18} \cdots \text{O5}} = 139^\circ$ ) for the  $C_7$  conformer, as compared to the DZP results. On going from the DZP level to the TZP level, C-N, C-C, and C-O bond lengths are increased by 0.001–0.004 Å, whereas C-H and N-H bond lengths are decreased by 0.003–0.005 Å. The change in bond angles is less than  $1^\circ$  on going from DZP to TZP.

The atomic Cartesian coordinates corresponding to the three conformers of diglycine and the final energy gradients for the conformers  $C_7$  and  $C_5$  (DZP basis set) are given in Tables IV and V.

The results from the least-squares fit of partial charges to the molecular electrostatic potential of diglycine are shown in Table VI. The "all-atom" model yields a good representation of the electrostatic potential. If a single set of charges is used to describe the potential of all three conformers, the rms deviation of the fit is increased from 0.4 to 1.05 kJ/mol. The "united-atom" representation is able to fit the electrostatic potential with rms values of 1–1.9 kJ/mol for individual conformers and with a rms value of 2.3 kJ/mol for the simultaneous fit of all three conformers. In view of the small number of only four adjustable parameters, these fits can also be regarded as reasonable. However, they are achieved at the expense of partial charges deviating from expectations based on electronegativity: The fit using the united-atom representation yields a very small partial charge on the nitrogen atoms.

#### 4. Discussion

In judging the present results, there are two points one should bear in mind: (1) At the DZP and the TZP levels, the electrostatic interactions and multipole moments are usually overestimated by typically 5–20%.<sup>22</sup> (2) Correlation effects are not included at the Hartree-Fock level.

Table II. Atomic Cartesian Coordinates for Dialanine Conformers  $C_{7,\text{eq}}$ ,  $C_{7,\text{ax}}$ ,  $C_5$ , and  $\alpha_L$  (DZP Basis Set)

atom	x	y	z	atom	x	y	z
$C_{7,\text{eq}}$							
H1	-0.3415	1.1219	-0.0294	C12	4.2016	1.2102	1.4428
H2	-0.3265	-0.4533	-0.8475	O13	4.5420	2.1955	2.0322
H3	-0.3545	-0.3870	0.9068	H14	5.4415	0.3889	3.6148
C4	0.0337	0.1021	0.0158	H15	3.7902	-0.0720	4.0268
O5	2.1021	-1.0475	0.0644	H16	4.9316	-1.2971	3.4516
C6	1.5396	0.0136	0.0396	N17	4.2530	-0.0284	1.9640
N7	2.2064	1.1851	0.0254	H18	3.8072	-0.7627	1.4512
H8	1.6923	2.0365	0.0962	C19	4.6176	-0.2620	3.3401
C9	3.6590	1.2728	0.0093	H20	5.1866	2.5902	-0.7436
H10	4.0160	0.4009	-0.5317	H21	3.7074	2.5560	-1.7149
C11	4.1009	2.5412	-0.6991	H22	3.7587	3.4249	-0.1640
$C_{7,\text{ax}}$							
H1	-0.4217	1.0312	-0.1994	C12	4.2168	1.1632	-1.4958
H2	-0.4750	-0.7164	-0.5254	O13	4.8459	2.0522	-1.9974
H3	-0.2400	-0.1343	1.1199	H14	4.3072	0.3797	-4.1083
C4	-0.0099	0.0614	0.0726	H15	5.7209	-0.3170	-3.3240
O5	2.0389	-1.0786	-0.2439	H16	4.3093	-1.3316	-3.6555
C6	1.4869	-0.0196	-0.1094	N17	4.0576	-0.0382	-2.0677
N7	2.1413	1.1585	-0.0823	H18	3.5179	-0.7308	-1.5894
H8	1.5880	1.9845	-0.0250	C19	4.6321	-0.3377	-3.3568
C9	3.5875	1.3648	-0.1113	H20	5.3905	0.8607	0.9425
H10	3.7245	2.4248	0.0736	H21	3.9427	0.8582	1.9543
C11	4.3353	0.5908	0.9735	H22	4.2429	-0.4829	0.8463
$C_5$							
H1	-0.3060	1.0903	-0.2192	C12	4.0358	2.7240	0.1498
H2	-0.3172	-0.6499	1.0912	O13	3.2648	3.6175	-0.5888
H3	-0.2934	-0.1059	1.0912	H14	5.3700	4.9610	0.2313
C4	0.0691	0.1171	0.0890	H15	5.5699	4.5345	1.9316
O5	2.1468	-1.0000	0.3196	H16	6.8858	4.2346	0.7836
C6	1.5767	0.0307	0.0969	N17	5.2899	2.9223	0.5956
N7	2.2318	1.1814	-0.1525	H18	5.8303	2.1272	0.8578
H8	1.7393	2.0371	-0.2994	C19	5.8054	4.2350	0.9093
C9	3.6686	1.2695	-0.1291	H20	5.3893	0.8471	-1.3979
H10	4.0359	0.6437	0.6830	H21	4.0139	-0.2390	-1.6224
C11	4.3015	0.7935	-1.4402	H22	3.9574	1.4088	-2.2714
$\alpha_L$							
H1	-0.3849	1.1011	-0.2492	C12	4.3217	1.2787	-1.3877
H2	-0.3758	-0.6261	-0.6791	O13	5.4446	0.8794	-1.4675
H3	-0.0396	-0.1205	0.9797	H14	5.1527	1.7550	-3.8417
C4	0.0946	0.1397	-0.0711	H15	3.9532	0.4970	-4.1440
O5	2.1458	-0.7448	-0.8785	H16	3.5568	2.1956	-4.4628
C6	1.5716	0.1757	-0.3841	N17	3.6139	1.7117	-2.4501
N7	2.2042	1.3424	-0.0684	H18	2.6405	1.8678	-2.3157
H8	1.6904	2.0016	0.4759	C19	4.0939	1.5223	-3.7981
C9	3.6556	1.4579	-0.0184	H20	5.3693	0.7560	1.0852
H10	3.8487	2.5011	0.2365	H21	3.8625	0.7893	2.0163
C11	4.3016	0.5729	1.0414	H22	4.1535	-0.4783	0.8104

The first error will yield an artificial stabilization of the  $C_7$  conformation, because this structure contains an intramolecular hydrogen bond. On the other hand, the neglect of the correlation effects such as the dispersion interaction will artificially destabilize the  $C_7$  conformation relative to the  $C_5$  conformation because it is more compact with more comparatively short nonbonded atom-atom contacts. One may hope that these two errors cancel at least in part. A rough estimate for both contributions may be obtained by analyzing the results from force field calculations and considering the contributions from the attractive part of the Lennard-Jones potential and the point-charge contributions. By making the assumption that the  $C_6/R^6$  term of the force field gives a measure of the correlation effects  $E_{\text{corr}}$  (or at least describes the relative importance of the correlation effects for different conformations) and the  $q_i q_j / R_{ij}$  term of the force field gives a measure of the electrostatic interaction  $E_{\text{es}}$ , we can calculate the differences in the nonbonded correlation interaction  $\Delta E_{\text{corr}}$  and the electrostatic interaction  $\Delta E_{\text{es}}$  between the  $C_5$  and  $C_7$  conformations. With the CVFF force field,<sup>5</sup> we obtain the following values for dialanine:

$$\Delta E_{\text{es}}(C_5-C_7) \cong 11.6 \text{ kJ/mol} \quad \Delta E_{\text{corr}}(C_5-C_7) \cong 1.3 \text{ kJ/mol}$$

Thus, if  $\Delta E_{\text{es}}$  is overestimated by 10%, this error ( $0.1 \times 11.6$  kJ/mol) roughly cancels the neglect of the correlations effects (1.3 kJ/mol). However, as both  $E_{\text{es}}$  and  $E_{\text{corr}}$  are highly dependent on conformational changes, this result may be fortuitous. It is

Table III. Ab Initio DZP and TZP Results on Diglycine<sup>a</sup>

	conformation (basis set)				
	C <sub>7</sub> (DZP)	C <sub>5</sub> (DZP)	α (DZP)	C <sub>7</sub> (TZP)	C <sub>5</sub> (TZP)
rel energy (kJ/mol)	1.3	0.0	16.4	1.7	0.0
φ (deg)	-85.4	-179.9	-60.0	-85.6	-179.8
ψ (deg)	75.5	-179.8	-39.0	77.1	-179.8
ω <sub>1</sub> (deg)	-177.2	-179.9	-170.0	-178.0	-179.9
ω <sub>2</sub> (deg)	-175.4	-178.4	174.8	-175.7	-178.9
R <sub>C4-C6</sub> (Å)	1.508	1.509	1.510	1.511	1.512
R <sub>C6-N7</sub> (Å)	1.348	1.347	1.362	1.349	1.347
R <sub>C6-O5</sub> (Å)	1.200	1.198	1.192	1.201	1.199
R <sub>N7-C9</sub> (Å)	1.448	1.431	1.446	1.450	1.434
R <sub>C9-Cl2</sub> (Å)	1.526	1.517	1.524	1.529	1.521
R <sub>Cl2-O13</sub> (Å)	1.197	1.197	1.193	1.198	1.197
R <sub>Cl2-N17</sub> (Å)	1.344	1.343	1.350	1.343	1.344
R <sub>N17-C19</sub> (Å)	1.442	1.445	1.444	1.447	1.449
R <sub>N7-H8</sub> (Å)	0.996	0.998	0.997	0.991	0.993
R <sub>N17-H18</sub> (Å)	1.000	0.996	0.996	0.995	0.992
∠C4C6N7 (deg)	116.3	115.9	115.3	116.3	116.0
∠C6N7C9 (deg)	122.6	121.5	121.6	122.5	121.4
∠N7C9C1 (deg)	112.8	109.4	115.7	112.5	109.4
∠C9C12N17 (deg)	114.8	114.9	115.8	115.1	114.8
∠C12N17C19 (deg)	122.0	121.8	121.4	121.6	121.8

<sup>a</sup> The total energy of the lowest energy conformation is -453.794 877 8 au at the DZP level and -453.944 236 1 au at the TZP level. Note that, for the helical conformation, φ and ψ were kept fixed during the calculation.

Table IV. Atomic Cartesian Coordinates (Å) and Final Energy Gradients (kJ mol<sup>-1</sup> Å<sup>-1</sup>) for Diglycine Conformers C<sub>7</sub> and C<sub>5</sub> (DZP Basis Set)

atom	x	y	z	dE/dx	dE/dy	dE/dz
C <sub>7</sub>						
H1	-1.5851	1.3541	-3.3979	-0.0155	0.0076	0.0033
H2	-2.4824	2.5768	-2.4807	0.0000	-0.0065	-0.0002
H3	-0.9353	2.9968	-3.2003	0.0336	0.0038	-0.0146
C4	-1.4896	2.1966	-2.7167	-0.0155	-0.0070	0.0231
O5	-0.4623	2.6939	-0.6410	-0.0176	-0.0012	0.0103
C6	-0.7987	1.8467	-1.4226	-0.0176	-0.0012	0.0103
N7	-0.6049	0.5336	-1.1853	-0.0332	0.0044	0.0366
H8	-0.8210	-0.1274	-1.8984	-0.0094	-0.0081	-0.0152
C9	-0.0112	0.0413	0.0402	0.0470	0.0595	0.0449
H10	-0.3356	-0.9818	0.1961	0.0076	-0.0071	-0.0105
H11	-0.3595	0.6568	0.8651	-0.0341	-0.0079	-0.0383
C12	1.5136	0.0287	-0.0038	-0.0336	-0.0380	-0.0057
O13	2.1203	-0.9943	-0.1418	0.0117	-0.0043	0.0008
H14	3.8001	2.3400	0.5115	-0.0098	0.0028	0.0053
H15	3.8457	1.4719	-1.0305	-0.0091	0.0024	0.0095
H16	4.0322	0.5864	0.4830	-0.0015	-0.0063	0.0062
N17	2.0925	1.2351	0.1262	0.0058	-0.0041	0.0242
H18	1.4963	2.0384	0.1230	0.0159	0.0001	-0.0040
C19	3.5185	1.4172	0.0098	-0.0058	0.0397	-0.0345
C <sub>5</sub>						
H1	-1.2961	-3.7571	0.0580	-0.0008	-0.0007	0.0047
H2	-2.7633	-3.2618	0.9254	0.0000	0.0002	-0.0024
H3	-2.7753	-3.3327	-0.8290	0.0035	0.0006	-0.0023
C4	-2.1590	-3.0954	0.0355	0.0021	0.0009	0.0127
O5	-2.6211	-0.7729	-0.1045	0.0030	-0.0008	-0.0489
C6	-1.7907	-1.6326	-0.0256	0.0030	-0.0008	-0.0489
N7	-0.4738	-1.3540	0.0182	0.0017	0.0006	-0.0231
H8	0.2174	-2.0711	0.0836	0.0003	0.0004	-0.0044
C9	0.0071	-0.0070	-0.0242	0.0145	-0.0053	0.0256
H10	-0.3789	0.5785	0.8115	0.0006	0.0036	-0.0007
H11	-0.3187	0.4988	-0.9346	0.0000	-0.0026	-0.0007
C12	1.5242	-0.0101	0.0296	-0.0165	0.0073	-0.0627
O13	2.1536	-1.0263	0.1001	0.0083	-0.0005	0.0101
H14	3.7517	2.3675	0.4526	0.0194	-0.0162	-0.0440
H15	3.9259	1.3971	-1.0191	-0.0042	0.0027	0.0025
H16	3.9936	0.6190	0.5635	0.0135	-0.0043	-0.0069
N17	2.0878	1.2116	0.0061	-0.0093	0.0284	0.0425
H18	1.4983	2.0035	-0.1325	-0.0008	-0.0012	-0.0005
C19	3.5189	1.4098	-0.0069	0.0046	-0.0423	0.0524

clear that more accurate calculations including correlation effects are needed to provide a more reliable estimate for the relative energies.

Schäfer et al. have reported 4-21G calculations on dialanine<sup>10,11</sup> and on diglycine.<sup>9</sup> In agreement with our results, they have found

Table V. Atomic Cartesian Coordinates for Dialanine α-Helix Conformer and Diglycine Conformers C<sub>7</sub> and C<sub>5</sub>

atom	x	y	z	atom	x	y	z
Dialanine α-Helix Conformer (DZP Basis Set)							
H1	-1.4892	2.8667	2.4151	H11	-0.2607	-1.0518	0.3393
H2	-2.1887	1.4555	3.2158	C12	1.5341	-0.0192	-0.0087
H3	-0.7623	2.2320	3.9102	O13	2.1425	-1.0360	-0.1435
C4	-1.2542	1.9660	2.9803	H14	3.8014	2.3515	-0.3411
O5	0.4984	0.3810	2.7185	H15	4.0126	0.6096	-0.5576
C6	-0.3617	1.0269	2.2052	H16	3.9072	1.2825	1.0704
N7	-0.6095	0.9585	0.8675	N17	2.1021	1.2056	-0.0249
H8	-1.4205	1.4182	0.5152	N18	1.5321	1.9686	0.2671
C9	0.0110	-0.0445	0.0317	C19	3.5345	1.3713	0.0486
H10	-0.3401	0.1065	-0.9873				
Diglycine Conformer C <sub>7</sub> (TZP Basis Set)							
H1	-1.6828	1.3588	-3.3528	H11	-0.3556	0.6191	0.8521
H2	-2.4444	2.6859	-2.4727	C12	1.5234	0.0314	-0.0339
H3	-0.8854	2.9313	-3.2287	O13	2.1399	-0.9784	-0.2210
C4	-1.5010	2.2093	-2.7086	H14	3.7693	2.4202	0.3523
O5	-0.4987	2.6809	-0.6132	H15	3.8667	1.2858	-0.9937
C6	-0.8198	1.8429	-1.4108	H16	4.0454	0.7077	0.6532
N7	-0.6077	0.5283	-1.1971	N17	2.0938	1.2342	0.1432
H8	-0.8149	-0.1200	-1.9180	H18	1.4996	2.0314	0.1766
C9	-0.0044	0.0238	0.0213	C19	3.5242	1.4183	0.0272
H10	-0.3139	-1.0002	0.1606				
Diglycine Conformer C <sub>5</sub> (TZP Basis Set)							
H1	-1.2975	-3.7563	-0.0121	H11	-0.3517	0.5343	-0.8741
H2	-2.7464	-3.3016	0.8901	C12	1.5229	0.0000	0.0089
H3	-2.7804	-3.3029	-0.8594	O13	2.1599	-1.0140	0.0151
C4	-2.1585	-3.1006	0.0033	H14	3.7395	2.4539	0.0908
O5	-2.6266	-0.7740	-0.0153	H15	3.9559	1.0319	-0.9272
C6	-1.7914	-1.6335	-0.0045	H16	3.9824	0.8704	0.8211
N7	-0.4741	-1.3534	0.0040	N17	2.0860	1.2204	0.0126
H8	0.2143	-2.0695	0.0119	H18	1.5011	2.0204	-0.0245
C9	0.0019	-0.0004	0.0016	C19	3.5243	1.3985	-0.0040
H10	-0.3599	0.5412	0.8694				

the C<sub>7,eq</sub> conformation as the lowest energy structure for dialanine. Also the bond lengths and bond angles obtained from the 4-21G calculation are very similar to the present DZP geometries. Significant differences are, however, found for the relative energies of the conformers and for the lowest energy structure of diglycine. For dialanine, the C<sub>5</sub> conformation is 5.8 kJ/mol less stable than C<sub>7,eq</sub> with the 4-21G basis set<sup>12</sup> compared with the energy difference of 2.1 kJ/mol at the DZP level. The difference between the 4-21G and the DZP basis sets is more striking for diglycine. The 4-21G calculation<sup>9</sup> yields an energy difference between C<sub>5</sub> and C<sub>7</sub> of 3.3 kJ/mol (C<sub>7</sub> more stable than C<sub>5</sub>), whereas the present DZP calculation predicts the C<sub>5</sub> conformer of diglycine to be more stable than C<sub>7</sub> by 1.3 kJ/mol. At the TZP level, this

**Table VI.** Partial Charges Positioned at the Nuclei Calculated from a Least-Squares Fit of the Molecular Electrostatic Potential and the Root Mean Square Deviation of the Fit (kJ/mol)

"All-Atom" Fit						
	conformer (basis set)					
	C <sub>7</sub> (DZP)	C <sub>5</sub> (DZP)	α (DZP)	all (DZP)	C <sub>7</sub> (TZP)	C <sub>5</sub> (TZP)
H <sub>1</sub>	-0.036	-0.038	-0.027	-0.034	-0.032	-0.027
H <sub>2</sub>	0.020	0.013	0.017	0.020	0.017	0.009
H <sub>3</sub>	0.015	0.013	0.020	-0.004	0.015	0.009
C <sub>4</sub>	0.026	0.029	0.000	0.033	0.033	0.036
O <sub>5</sub>	-0.637	-0.636	-0.584	-0.619	-0.644	-0.647
C <sub>6</sub>	0.718	0.727	0.706	0.747	0.714	0.736
N <sub>7</sub>	-0.698	-0.710	-0.686	-0.732	-0.698	-0.718
H <sub>8</sub>	0.404	0.395	0.378	0.389	0.402	0.388
C <sub>9</sub>	-0.049	-0.048	-0.114	-0.080	-0.046	-0.052
H <sub>10</sub>	0.076	0.094	0.082	0.072	0.079	0.096
H <sub>11</sub>	0.099	0.097	0.124	0.108	0.105	0.099
C <sub>12</sub>	0.747	0.761	0.774	0.847	0.746	0.775
O <sub>13</sub>	-0.634	-0.636	-0.584	-0.619	-0.629	-0.627
H <sub>14</sub>	0.043	0.042	0.030	0.047	0.050	0.044
H <sub>15</sub>	0.044	0.044	0.086	0.054	0.049	0.049
H <sub>16</sub>	0.066	0.072	0.055	0.065	0.057	0.055
N <sub>17</sub>	-0.642	-0.620	-0.613	-0.672	-0.641	-0.616
H <sub>18</sub>	0.370	0.324	0.343	0.350	0.367	0.326
C <sub>19</sub>	0.051	0.058	0.026	0.042	0.060	0.067
rmsd (kJ/mol)	0.42	0.38	0.44	1.05	0.46	0.44

"United-Atom" Fit				
	conformer (basis set)			
	C <sub>7</sub> (DZP)	C <sub>5</sub> (DZP)	α (DZP)	all (DZP)
H <sub>1</sub>	0.0	0.0	0.0	0.0
H <sub>2</sub>	0.0	0.0	0.0	0.0
H <sub>3</sub>	0.0	0.0	0.0	0.0
C <sub>4</sub>	0.083	0.112	0.074	0.095
O <sub>5</sub>	-0.520	-0.477	-0.519	-0.486
C <sub>6</sub>	0.351	0.211	0.433	0.291
N <sub>7</sub>	-0.084	-0.038	-0.125	-0.079
H <sub>8</sub>	0.128	0.134	0.098	0.131
C <sub>9</sub>	0.083	0.112	0.074	0.095
H <sub>10</sub>	0.0	0.0	0.0	0.0
H <sub>11</sub>	0.0	0.0	0.0	0.0
C <sub>12</sub>	0.351	0.211	0.433	0.291
O <sub>13</sub>	-0.520	-0.477	-0.519	-0.486
H <sub>14</sub>	0.0	0.0	0.0	0.0
H <sub>15</sub>	0.0	0.0	0.0	0.0
H <sub>16</sub>	0.0	0.0	0.0	0.0
N <sub>17</sub>	-0.084	-0.038	-0.125	-0.079
H <sub>18</sub>	0.128	0.134	0.098	0.131
C <sub>19</sub>	0.083	0.112	0.074	0.095
rmsd (kJ/mol)	1.18	1.00	1.89	2.34

energy difference amounts to 1.7 kJ/mol. In view of the small differences on geometric properties and conformational energies between the DZP and the TZP basis set it, appears that the DZP basis set is a reliable choice at least at the Hartree-Fock level for calculations on peptide conformers.

Table VII, the present results on the relative energies of the different conformations of dialanine and diglycine are compared with the 4-21G data from Schäfer et al.<sup>8-11</sup> and results from force field calculations using the force fields AMBER,<sup>1</sup> ECEPP,<sup>3</sup> and CVFF<sup>4</sup> and data reported by Pettit et al.<sup>7</sup> A recent reparameterization of the CHARMM force field by Momany et al.<sup>24</sup> using the 4-21G data from Schäfer et al. yields conformational energies close to original ab initio data. All force fields give in agreement with the present ab initio calculations the C<sub>7</sub> conformation as the global minimum for dialanine. However, large discrepancies occur for the relative conformational energies and for the lowest energy conformer of diglycine. Most important, the energy difference between the C<sub>5</sub> conformation and the C<sub>7</sub> conformation is strongly overestimated in the AMBER and the CVFF force field and in the force field used by Pettit et al.<sup>7</sup> The ECEPP force field gives the best agreement with the present ab initio results on the C<sub>5</sub>-C<sub>7,eq</sub>

**Table VII.** Comparison of Conformational Energies of Dialanine and Diglycine Obtained from the Present ab Initio DZP and TZP Calculations with 4-21G Data from Schäfer et al.<sup>8-11</sup> and Results from the Force Fields AMBER,<sup>1</sup> CVFF,<sup>4</sup> and ECEPP<sup>3</sup> and Results by Pettit et al.<sup>7</sup> (kJ/mol)<sup>a</sup>

Dialanine						
conformer	DZP	4-21G	AMBER	CVFF	ECEPP	Pettit
C <sub>7,eq</sub>	0.0	0.0	0.0	0.0	0.0	0.0
C <sub>7,ax</sub>	12.5	10.9	2.5	5.4	37.2	0.4
α <sub>R</sub>	(13.5)	(24)	15.1	17.6	4.7	n.d.
α <sub>L</sub>	18.5	(28)	18.0	(30.5)	9.8	n.d.
C <sub>5</sub>	2.1	5.8	13.4	(10)	1.6	15.1
Diglycine						
conformer	DZP	TZP	4-21G	AMBER	CVFF	ECEPP
C <sub>7</sub>	1.3	1.7	0.0	0.0	0.0	0.0
α	(16.4)	n.d.	(20)	17.2	(26)	4.9
C <sub>5</sub>	0.0	0.0	3.3	13.8	(12)	3.4

<sup>a</sup> Values in parentheses indicate that this conformation is not a local minimum within the given method or force field.

energy difference for dialanine; however, agreement with the present DZP results with respect to the other conformations, most notably for C<sub>7,ax</sub>, is less satisfactory. The present ab initio data

(24) Momany, F. A.; Klimkowski, V. J.; Schäfer, L. *J. Comput. Chem.* 1990, 11, 654.

predict a significantly larger energy difference for  $E(C_{7,eq}) - E(C_{7,ax})$  than most of the force fields (except ECEPP; see Table VII). The hydrogen bond in  $C_{7,ax}$  is calculated to be shorter than in  $C_{7,eq}$ . This is in agreement at least with the CVFF force field, which also predicts a shorter hydrogen bond for  $C_{7,ax}$ . In the CVFF force field, this is achieved at the expense of unfavorable contributions from angle bending. It appears that, in the  $C_{7,ax}$  conformer, the hydrogen bond is pushed together by intramolecular strain.

It is difficult to track down the origins of the force field deficiencies. The analysis that we have carried out so far indicates that several small deviations in various terms of the force field add up to the significant energy differences described above. To be more specific, we would like to point to some possible origins for the discrepancies between ab initio data and at least some of the force fields:

(1) The ab initio data predict similar hydrogen bond lengths  $R_{H18-O5}$  for the  $C_5$  and the  $C_7$  conformers. In contrast, both the CVFF and the AMBER force fields give a longer hydrogen bond for  $C_5$  than for  $C_7$ . In the CVFF force field, this is at least in part due to deviations in the bond and angle terms. CVFF calculates similar values of  $R_{N7C9}$  for the  $C_5$  (1.482 Å) and the  $C_7$  conformers (1.485 Å) of diglycine whereas the ab initio data predict a significantly shorter N7-C9 bond length of 1.434 Å for the  $C_5$  conformer than for the  $C_7$  conformer (1.450 Å). CVFF also calculates similar values of  $\angle_{N7C9C12}$  for the  $C_5$  (115.5°) and the  $C_7$  (115.7°) conformers of diglycine whereas the ab initio DZP data predict a  $\angle_{N7C9C12} = 109.4^\circ$  for the  $C_5$  conformer and  $112.5^\circ$  for the  $C_7$  conformer. The same trend is found for dialanine.

Apparently the CVFF force field underestimates the dependence of the bond length N7-C9 and the angle N7-C9-C12 on conformational change. As the values for  $R_{N7C9}$  ( $C_5$ , 1.451 Å;  $C_7$ , 1.457 Å) and  $\angle_{N7C9C12}$  ( $C_5$ , 108.5°;  $C_7$ , 111.6°) calculated with the AMBER force field<sup>23</sup> for the  $C_5$  and  $C_7$  conformers are in better agreement with the present ab initio data, the origin of the discrepancy must be different. We suspect that the hydrogen bond term in the AMBER force field does not describe accurately the difference between these two conformers. An analysis of the various energy terms of the AMBER force field shows indeed that the hydrogen bond term shows the largest energy differences between the  $C_5$  and the  $C_7$  conformers.

(2) All major current force fields use fixed partial charges that do not change upon conformational changes. However, the present data demonstrate that there is a certain change of partial charges on conformational change. If this charge variation is ignored, the error introduced into the force field is around 1-2 kJ/mol.

In summary, we have shown that ab initio SCF calculations at the DZP level on small model dipeptides can be used to assess the quality of current force fields. The ab initio results point to some important deficiencies of the force fields with respect to the relative energies of the  $C_7$  and the  $C_5$  conformations of dialanine and diglycine. We conclude that several of the current force fields overestimate the stability of the  $C_7$  conformation. We are currently using the present results together with additional ab initio calculations to derive an improved parameter set for a protein force field.

Registry No. Dialanine, 19701-83-8; diglycine, 7606-79-3.

## Substituent Effects and the Charge Topology in Nitriles and Cyanides

Yosslen Aray,<sup>†,‡</sup> Juan Murgich,<sup>\*,‡</sup> and Miguel A. Luna<sup>§</sup>

Contribution from the Centro de Química, Instituto Venezolano de Investigaciones Científicas (IVIC), Apartado 21827, Caracas 1020A, Venezuela, and Centro Científico, IBM de Venezuela C.A., Apartado 64778, Caracas 1060, Venezuela. Received June 4, 1990

**Abstract:** In order to study the substituent effects and the bonding of the  $C\equiv N$  group, a SCF 4-31G\*\* ab initio calculation was undertaken for the  $C\equiv N^-$  ion, for nitriles with  $R = H, F, HO, FO, H_3C, HC\equiv C, NH_2, O=N, N_3, C\equiv N$ , the  $E$  and  $Z$  form of iminoacetonitrile and for cyanides with  $R = Li, HBe, and B$ . The electron density  $\rho(r)$  was analyzed by means of the topological theory of atoms in molecules. The surfaces of zero-flux in  $\nabla\rho(r)$  showed that the C valence density of the target fragment absorbed most of the changes, while the N was only slightly perturbed by the R-C bond formation. A correlation was found between the position of the bond critical point and the electronegativity of R for R-C in nitriles but not for the  $C\equiv N$  bond. This difference was explained by the changes in the strength of the substituent fields acting on both bonds. The  $\nabla^2\rho(r)$  distribution in the  $C\equiv N$  group for the different substituents and counterions showed that all the bonded and nonbonded local concentrations present in the C and N valence shells were affected in a complex way by the bond formation.

### Introduction

One of the most important concepts of chemistry is that rates and equilibria are markedly affected by changes in the molecular skeleton or by the introduction of additional substituents.<sup>1,2</sup> Then, a detailed study of the electronic effects of substitution and coordination on a particular molecular fragment will help to understand the effects governing its chemical behavior. Most of the analyses of the electronic substituent effects were based in the use of  $\sigma$  and  $\pi$  molecular orbitals and their respective populations.<sup>3</sup> This approach is not the most appropriate because molecular

orbitals and its population do not constitute physical observables as defined by quantum mechanics.<sup>4,5</sup> An analysis based on a specific orbital description is then flawed by the dependence of the parameters such as the Mulliken population on the model used. In view of these severe limitations, it is more convenient and

(1) Isaacs, N. S. *Physical Organic Chemistry*; Longman Ltd.: Essex, England, 1987; pp 114-170.

(2) Taft, R. W.; Topsom, R. D. *Prog. Phys. Org. Chem.* **1987**, *16*, 1-83.

(3) Topsom, R. D. *Acc. Chem. Res.* **1983**, *16*, 292-298 and *Prog. Phys. Org. Chem.* **1987**, *16*, 85-124.

(4) Bader, R. F. W.; Nguyen-Dang, T. I.; Tai, Y. *Rep. Prog. Phys.* **1981**, *44*, 895-948. Bader, R. F. W. *Atoms in Molecules: a Quantum Theory*; Oxford University Press: Oxford, England, 1990; Chapter 7.

(5) Bader, R. F. W. *Acc. Chem. Res.* **1985**, *18*, 9-15.

<sup>†</sup> On leave from Departamento de Química, Facultad Experimental de Ciencias, Universidad del Zulia, Maracaibo, Venezuela

<sup>‡</sup> Instituto Venezolano de Investigaciones Científicas.

<sup>§</sup> IBM de Venezuela C.A.