# Ab Initio SCF Calculations on Low-Energy Conformers of *N*-Acetylglycylglycine *N'*-Methylamide

**Hans-Joachim Böhm**

*Contribution from BASF AG, Central Research, 6700 Ludwigshafen, West Germany*

*Received February 14, 1992*

**Abstract:** Results from ab initio SCF calculations with a 3-21G and double-$\zeta$ plus polarization (DZP) basis set on six low-energy conformations of *N*-acetylglycylglycine *N'*-methylamide (diglycine diamide) are reported. In addition, MP2 energies for DZP-optimized geometries are also given. The lowest energy conformation found at the DZP level is a type II $\beta$-turn with dihedral angles close to the reported "ideal" values. This contrasts with the 3-21G results which yield a conformation with two $C_7$-turns as the lowest energy structure. The geometric differences between the 3-21G- and the DZP-minimized conformers are very small except for the type II $\beta$-turn conformer. At the 3-21G level, this conformer is distorted towards a $C_7$-turn so that a bifurcated hydrogen bond is formed. The helical conformer is not a local minimum for diglycine diamide at the 3-21G level or at the DZP level. The present ab initio results are compared with data from semiempirical and molecular mechanics force field calculations. Significant differences are found for the geometry of type I $\beta$-turn and for the conformational energies.

## 1. Introduction

In the present paper, we report results from ab initio calculations of several conformers of *N*-acetylglycylglycine *N'*-methylamide (Ac-Gly-Gly-$NHCH_3$: "diglycine diamide"). The purpose of the work was 2-fold. First, as an extension of the previous work on dipeptide conformers,[1] we wanted to use the data to assess current protein force fields with the further goal of deriving improved parameter sets. Second, we hoped to get a better understanding of the structural motifs found in proteins by studying the intrinsic stability of relevant small peptide conformers.

Diglycine diamide is the smallest peptide capable of forming $\beta$-turns. This turn is characterized by a hydrogen bond formed between the backbone carbonyl group of the amino acid at the chain position *i* and the backbone amide proton of the amino acid at position $i+3$. This important structural motif occurs frequently in proteins.[2] $\beta$-Turns have been classified by Richardson and others.[2-6] The most abundant structures in proteins are type I and type II $\beta$-turns.[5] A statistical analysis of $\beta$-turns in proteins revealed a preference for glycine in position $i+2$ for type II $\beta$-turns.[5] Several consecutive type I turns are also referred to as $3_{10}$-helix.[6] $3_{10}$-Helices are found in some proteins although much less frequently than $\alpha$-helices.[6] There is also experimental evidence that some small peptides adopt a $\beta$-turn conformation in aqueous solution.[7,8]

Quantum mechanical ab initio investigations on peptides have so far mainly focussed on structures with only two amide groups.[1,9-11] A recent investigation on the conformational properties of tripeptides[12] reported a geometry optimization of the double-$C_7$ conformer of Ac-Ala-Ala-$NHCH_3$ at the 3-21G level. Chesnut et al.[13] carried out calculations on a planar conformer of H-Gly-Gly-OH. Sapse et al.[14] studied several conformers of Ac-Ala-Gly-$NHCH_3$ at the STO-3G level. In view of the limited accuracy of the STO-3G basis set and the importance of conformational properties of short peptides, we decided to carry out an investigation of the conformational properties of Ac-Gly-Gly-$NHCH_3$ using the 3-21G and double-$\zeta$ plus polarization (DZP) basis set.

The next section describes the technical details of the ab initio calculations. The results are presented in section 3. Summary and conclusions are given in section 4.

## 2. Computational Methods

The program TURBOMOLE developed by Häser and Ahlrichs[15,16] running on a Silicon Graphics 4D35 computer was used for all calculations. In the present calculations, we employed the 3-21G basis set[17] and the following DZP basis set (GTO primitive sets taken from Huzinagas tables:[18] C (8,4,1)/[4,2,1]; N (8,4,1)/[4,2,1]; O (8,4,1)/[4,2,1]; and H (4,1)/[2,1].

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 143 basis functions at the 3-21G level and 260 basis functions at the DZP level.

Calculations were carried out as follows. The various conformers were initially generated using interactive computer graphics.[19] The initial values of torsion angles $\phi$ and $\psi$ (see Figure 1 for definition of torsion angles), which essentially define the conformation of peptides,[20] were chosen as follows. For the extended, double-$C_7$, and helical conformers, the dihedral angles were set to values corresponding to the local minima

(1) Böhm, H. J.; Brode, S. *J. Am. Chem. Soc.* **1991**, *113*, 7129.
(2) Richardson, J. *Adv. Protein Chem.* **1981**, *34*, 167.
(3) Smith, J. A.; Pease, L. G. *Crit. Rev. Biochem.* **1980**, 315.
(4) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221.
(5) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1.
(6) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97.
(7) Montelione, G. T.; Arnold, E.; Meinwald, Y. C.; Stimson, E. R.; Denton, J. B.; Huang, S. G.; Clardy, J.; Scheraga, H. A. *J. Am. Chem. Soc.* **1984**, *106*, 7946.
(8) Wright, P. E.; Dyson, H. J.; Lerner, R. A. *Biochemistry* **1988**, *27*, 7167.
(9) Klimkowski, V. J.; Schäfer, L.; Momany, F. A.; van Alsenoy, C. *THEOCHEM* **1985**, *124*, 143.
(10) Perczel, A.; Angyan, J. G.; Kajtar, M.; Viviani, W.; Rivail, J.; Marcoccia, J. F.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1991**, *113*, 6256.
(11) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989.
(12) Perczel, A.; Kajtar, M.; Marcoccia, J. F.; Csizmadia, I. G. *THEOCHEM* **1991**, *232*, 291.
(13) Chesnut, D. B.; Phung, C. G. *Chem. Phys. Lett.* **1991**, *183*, 505.
(14) Sapse, A. M.; Daniels, S. B.; Erickson, B. W. *Tetrahedron* **1988**, *44*, 999.
(15) Häser, M.; Ahlrichs, R. *J. Comp. Chem.* **1989**, *10*, 104.
(16) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.
(17) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.
(18) Huzinaga, S. *J. Chem. Phys.* **1965**, *42*, 1293.
(19) INSIGHT; Biosym Technologies, 10065 Barnes Canyon Road, San Diego, CA 92121.
(20) Ramachandram, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 195.

0002-7863/93/1515-6152$04.00/0 

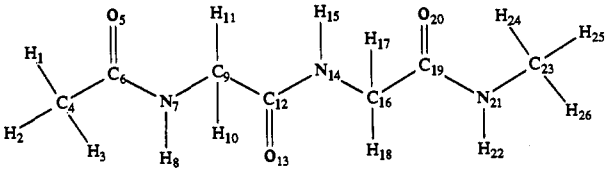

**Figure 1.** Atom numbering scheme for the molecule *N*-acetylglycylglycine *N'*-methylamide. The torsion angles discussed in the text are defined as follows: $\omega_1$ = ∠C4C6N7C9; $\omega_2$ = ∠C9C12N14C16; $\omega_3$ = ∠C16C19N21C23; $\phi_1$ = ∠C6N7C9C12; $\psi_1$ = ∠N7C9C12N17; $\phi_2$ = ∠C12N14C16C19; $\psi_2$ = ∠N14C16C19N21.

of Ac-Gly-NHCH$_3$ obtained from the AMBER force field.[21] The torsional angles of the type I and II β-turn were set to the "ideal" values as given by:[2]

extended: $\phi_1 = 180°, \psi_1 = 180°, \phi_2 = 180°, \psi_2 = 180°$

double-$C_7$ (boat): $\phi_1 = 76°, \psi_1 = -64°, \phi_2 = -76°, \psi_2 = 64°$

double-$C_7$ (chair): $\phi_1 = -76°, \psi_1 = 64°, \phi_2 = -76°, \psi_2 = 64°$

type I β-turn: $\phi_1 = -60°, \psi_1 = -30°, \phi_2 = -90°, \psi_2 = 0°$

type II β-turn: $\phi_1 = -60°, \psi_1 = 120°, \phi_2 = 80°, \psi_2 = 0°$

α-helix: $\phi_1 = -66°, \psi_1 = -35°, \phi_2 = -66°, \psi_2 = -35°$

The double-$C_7$ conformers contain two intramolecular hydrogen bonds, each forming a seven-membered ring. The two different double-$C_7$ conformers were labeled as "boat" (sign of the dihedrals $\phi_1,\psi_1\phi_2,\psi_2$: +,−,−,+) and "chair" (sign of the dihedrals $\phi_1,\psi_1\phi_2,\psi_2$:−,+,−,+) due to the resemblence of the $C_\alpha$ chain trace with the corresponding objects. The type I β-turn and type II β-turn each contain one intramolecular hydrogen bond. Note that the dihedral angles $\phi_1$ and $\psi_1$ correspond to amino acid $i + 1$ in the usual nomenclature of β-turns,[2] and $\phi_2$ and $\psi_2$ correspond to amino acid $i + 2$. Diglycine diamide clearly has a considerable degree of conformational flexibility even if only trans amide bonds are taken into account. We did not attempt to cover all possible conformers of diglycine diamide in the present study. Instead, we focussed on the six conformers listed above.

The manually built structures were used as input for the ab initio SCF calculations. A full optimization of all degrees of freedom was carried out in TURBOMOLE. All calculations were performed in $C_1$ symmetry. The minimization was continued until the maximum energy gradient dropped below 0.1 kJ mol$^{-1}$ Å$^{-1}$ and the energy change over the last 20 iterations was less than 0.04 kJ/mol. Between 30 and 70 SCF+ gradient steps were necessary to achieve convergence. The fully optimized structures at the 3-21G level were used as input for the DZP calculations. Single-point MP2 calculations were then performed on the DZP-optimized structures. Apart from the six conformations of diglycine diamide, no attempt was made to find further local minima on the potential surface. For comparison purposes, the five conformers found to be local minima in the ab initio calculations were also optimized using the semiempirical AM1 method[22] and the force fields CVFF[23] and XPLOR[24].

A very severe problem for ab initio calculations on large molecules (such as diglycine diamide) is the sparsity of experimental data with which to compare the results. Experimental gas-phase data on the structure of diglycine diamide are not available. In order to assess the accuracy of the present method used to calculate conformational energies, we carried out a series of calculations on small molecules where experimental data for conformational energies are available. The small molecules considered were butane, ethylamine, glycine, *N*-methylacetamide, ethyl methyl ether, propanol, and *N*-methylpiperidine. These examples (except glycine) were selected from a recent publication by Gundertofte et al.,[25] who calculated conformational energies of small molecules using various force fields and semiempirical methods. The conformational energies at the 3-21G and the DZP level for these molecules are compared with experimental data[26-32] in Table I. The maximum deviation for the DZP results is 1.5 kJ/mol. This good agreement with available experimental data, although far from being conclusive, gives confidence over the accuracy of the DZP basis set used in the present study to calculate conformational energies of organic molecules.

**Table I.** Comparison of Calculated Conformational Energies of Small Molecules at the 3-21G and DZP Level with Experimental Data

| | | ΔE | | |
|---|---|---|---|---|
| molecule | conformers | 3-21G | DZP | exp |
| butane | g–a | 3.2 | 4.1 | 4.1[a] |
| ethylamine | g–a | 1.2 | 0.4 | 1.3[b] |
| glycine | b–a | −0.2 | 7.8 | 7.1[c] |
| *N*-methylacetamide | *E–Z* | 12.9 | 10.9 | 9.6[d] |
| ethyl methyl ether | g–a | 4.2 | 7.1 | 6.3[e] |
| propanol | g–a | −2.6 | −1.0 | −1.2[f] |
| *N*-methylpiperidine | ax–eq | 8.4 | 15.8 | 14.6[g] |

[a] Reference 26. [b] Reference 27. [c] Reference 28; see Figure 1 in that paper for definition of conformers a and b. [d] Reference 29. [e] Reference 30. [f] Reference 31. [g] Reference 32.

The SCF wave functions calculated for the fully minimized structures of diglycine diamide were used to calculate the electrostatic potential at grid points surrounding the molecule. Three sets of points were used. The first set (no. 1) was a cubic grid of 0.75-Å spacing in a 1-Å thick shell around the molecule. The inner surface of this shell was 4 Å away from the molecule. The number of points was roughly 1500 for each conformer. The second set (no. 2) of points was again a cubic grid of 1-Å spacing in a 2-Å-thick shell 5 Å away from the molecule. This amounted to roughly 2000 points per conformer. The third set (no. 3) of points was also a cubic grid of 1-Å spacing in a 1-Å-thick shell 10 Å away from the molecule, which amounted to roughly 2200 points per conformer. A set of fractional charges positioned at the nuclei was then fitted to the potential at the points from sets 1 and 2. Furthermore, an attempt was made to fit the potentials of the extended, both β-turn, and double-$C_7$ (boat) conformers of diglycine diamide simultaneously by a single set of charges. The present approach is similar to the work of Chirlian et al.[33] and of Williams.[34]

The charges obtained from this fit were then used to calculate the electrostatic potential at the points from set 3. The comparison of the potential calculated from the point charges with the potential obtained directly from the ab initio wave function was carried out to assess the ability of the point charge model to describe electrostatic long range interactions.

## 3. Results

One conformer out of the six structures considered in this paper, the helical conformer, turned out to be neither a local minimum at the 3-21G level nor at the DZP level. The helical conformer converted into a type I β-turn. The results from the ab initio SCF calculations on the remaining five conformations of diglycine diamide using a 3-21G basis set are shown in Table II.

At the 3-21G level, the lowest energy conformation found is the double-$C_7$ (boat) conformer with $\phi_1 = 83.7°, \psi_1 = -66.2, \phi_2 = -81.6, \psi_2 = 66.4°$. The structure contains two intramolecular hydrogen bonds with $R_{O5\text{-}H15} = 1.98$ Å, $R_{O5\text{-}N14} = 2.86$ Å, $R_{O13\text{-}H22} = 2.00$ Å, and $R_{O13\text{-}N21} = 2.87$ Å. The second double-$C_7$ conformer is marginally less stable. The other structures considered all have very similar energies 7.5-8 kJ/mol above the double-$C_7$ conformers. The hydrogen-bond lengths in the β-turn structures at the 3-21G level are $R_{O5\text{-}H22} = 2.09$ Å and $R_{O5\text{-}N21}$

(21) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
(22) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(23) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Proteins* **1988**, *4*, 31.
(24) (a) Brünger, A. T.; Kuriyan, J.; Karplus, M. *Science* **1987**, *235*, 458. (b) XPLOR, Version 2.1; distributed by Molecular Simulations, Burlington, MA 01803-5297. The parameter file param19x was used in the calculations.
(25) Gundertofte, K.; Palm, J.; Pettersson, I.; Stamvik, A. *J. Comp. Chem.* **1991**, *12*, 200.
(26) Verma, A. L.; Murphy, W. F.; Bernstein, H. J. *J. Chem. Phys.* **1974**, *60*, 1540.
(27) Fischer, E.; Botskor, I. *J. Mol. Spectros.* **1984**, *104*, 226.
(28) Iijima, K.; Tanaka, K.; Onuma, S. *J. Mol. Struct.* **1991**, *246*, 257.
(29) Kitano, M.; Fukuyama, T.; Kuchitsu, K. *Bull. Chem. Soc. Jpn.* **1973**, *46*, 384.
(30) Kitagawa, T.; Miyazawa, T. *Bull. Chem. Soc. Jpn.* **1968**, *41*, 1976.
(31) Fang, H. L.; Swofford, R. L. *Chem. Phys. Lett.* **1984**, *105*, 5.
(32) Crowley, P. J.; Robinson, M. J. T.; Ward, M. G. *Tetrahedron* **1977**, *33*, 915.
(33) Chirlian, L. E.; Miller Francl, M. *J. Comp. Chem.* **1987**, *8*, 894.
(34) Williams, D. E. *Biopolymers* **1990**, *29*, 1367.

**Table II.** Ab Initio 3-21G Results on Diglycine Diamide[a]

| angle [deg] or bond [Å] | conformation, rel energy [kJ/mol] type I B-turn, 7.5 | type II β-turn, 8.0 | extd, 7.5 | double-$C_7$ (bt), 0.0 | double-$C_7$ (ch), 0.1 |
|---|---|---|---|---|---|
| $\phi_1$ | −67.2 | −75.5 | −179.8 | 83.7 | −83.5 |
| $\psi_1$ | −20.9 | 93.6 | 180.0 | −66.2 | 63.6 |
| $\phi_2$ | −108.3 | 117.4 | 179.9 | −81.6 | −84.1 |
| $\psi_2$ | 20.0 | −3.5 | 180.0 | 66.4 | 64.4 |
| $\omega_1$ | −171.5 | −179.7 | 180.0 | 175.8 | −175.0 |
| $\omega_2$ | 175.9 | −175.4 | 180.0 | −177.3 | −175.3 |
| $\omega_3$ | 176.9 | 179.4 | 180.0 | 179.9 | −179.3 |
| $R_{C4-C6}$ | 1.513 | 1.513 | 1.517 | 1.513 | 1.513 |
| $R_{C6-N7}$ | 1.361 | 1.349 | 1.352 | 1.355 | 1.354 |
| $R_{C6-O5}$ | 1.223 | 1.231 | 1.221 | 1.227 | 1.227 |
| $R_{N7-C9}$ | 1.459 | 1.467 | 1.447 | 1.465 | 1.464 |
| $R_{C9-C12}$ | 1.527 | 1.531 | 1.521 | 1.529 | 1.528 |
| $R_{C12-O13}$ | 1.218 | 1.220 | 1.225 | 1.230 | 1.230 |
| $R_{C12-N14}$ | 1.351 | 1.353 | 1.342 | 1.341 | 1.341 |
| $R_{N14-C16}$ | 1.453 | 1.452 | 1.449 | 1.465 | 1.465 |
| $R_{N7-H8}$ | 0.998 | 0.998 | 1.000 | 0.997 | 0.997 |
| $R_{N14-H15}$ | 0.996 | 0.998 | 1.002 | 1.004 | 1.004 |
| $R_{N21-H22}$ | 0.999 | 0.999 | 0.997 | 1.002 | 1.002 |
| $R_{C16-C19}$ | 1.528 | 1.529 | 1.524 | 1.531 | 1.532 |
| $R_{C19-N21}$ | 1.341 | 1.345 | 1.345 | 1.347 | 1.348 |
| $R_{C19-O20}$ | 1.224 | 1.224 | 1.223 | 1.222 | 1.222 |
| $R_{N21-C23}$ | 1.463 | 1.463 | 1.467 | 1.461 | 1.461 |
| $\angle_{C4C6N7}$ | 115.1 | 115.3 | 114.4 | 115.2 | 115.2 |
| $\angle_{C6N7C9}$ | 120.2 | 121.0 | 120.8 | 121.3 | 121.5 |
| $\angle_{N7C9C12}$ | 114.8 | 110.1 | 107.6 | 112.0 | 112.7 |
| $\angle_{C9C12N14}$ | 116.1 | 114.1 | 115.2 | 114.7 | 114.8 |
| $\angle_{C12N14C16}$ | 121.6 | 121.9 | 121.0 | 121.8 | 121.7 |
| $\angle_{N14C16C19}$ | 114.1 | 115.0 | 107.3 | 112.6 | 112.8 |
| $\angle_{C16C19N21}$ | 116.6 | 117.0 | 115.3 | 114.1 | 113.8 |
| $\angle_{C19N21C23}$ | 119.8 | 120.8 | 120.4 | 120.4 | 121.6 |

[a] The total energy of the lowest energy conformation double-$C_7$ is −656.95813 au. The $\alpha_R$ conformer is not a local minimum and is therefore not included in this table.

**Table III.** Ab Initio DZP Results on Diglycine Diamide[a]

| angle [deg] or bond [Å] | conformation; rel energy SCF, SCF-MP2 [kJ/mol] type I β-turn; 4.6, 4.4 | type II β-turn; 0.0, 0.0 | extd, 1.3, 16.4 | double-$C_7$ (bt); 3.7, 2.9 | double-$C_7$ (ch); 3.7, 3.2 |
|---|---|---|---|---|---|
| $\phi_1$ | −72.5 | −60.8 | −179.8 | 85.1 | −85.7 |
| $\psi_1$ | −18.1 | 135.7 | 180.0 | −67.4 | 66.2 |
| $\phi_2$ | −104.0 | 96.6 | 180.0 | −83.8 | −86.2 |
| $\psi_2$ | 13.4 | −11.4 | 180.0 | 74.6 | 70.8 |
| $\omega_1$ | −166.1 | 170.6 | 180.0 | 178.7 | −174.4 |
| $\omega_2$ | 175.8 | −174.2 | 180.0 | −178.1 | −175.4 |
| $\omega_3$ | 175.6 | −176.1 | 180.0 | −176.3 | −176.2 |
| $R_{C4-C6}$ | 1.509 | 1.509 | 1.512 | 1.509 | 1.509 |
| $R_{C6-N7}$ | 1.361 | 1.355 | 1.352 | 1.352 | 1.352 |
| $R_{C6-O5}$ | 1.198 | 1.202 | 1.198 | 1.202 | 1.201 |
| $R_{N7-C9}$ | 1.448 | 1.440 | 1.433 | 1.448 | 1.447 |
| $R_{C9-C12}$ | 1.524 | 1.527 | 1.518 | 1.526 | 1.526 |
| $R_{C12-O13}$ | 1.195 | 1.195 | 1.201 | 1.204 | 1.204 |
| $R_{C12-N14}$ | 1.350 | 1.350 | 1.343 | 1.343 | 1.342 |
| $R_{N14-C16}$ | 1.441 | 1.442 | 1.434 | 1.449 | 1.449 |
| $R_{N7-H8}$ | 0.998 | 0.998 | 1.000 | 0.997 | 0.997 |
| $R_{N14-H15}$ | 0.997 | 0.998 | 1.001 | 1.003 | 1.003 |
| $R_{N21-H22}$ | 1.000 | 1.000 | 0.997 | 1.001 | 1.001 |
| $R_{C16-C19}$ | 1.526 | 1.527 | 1.520 | 1.528 | 1.528 |
| $R_{C19-N21}$ | 1.343 | 1.342 | 1.344 | 1.347 | 1.347 |
| $R_{C19-O20}$ | 1.199 | 1.200 | 1.198 | 1.198 | 1.198 |
| $R_{N21-C23}$ | 1.445 | 1.445 | 1.448 | 1.443 | 1.443 |
| $\angle_{C4C6N7}$ | 116.3 | 116.0 | 115.0 | 116.3 | 116.2 |
| $\angle_{C6N7C9}$ | 119.2 | 120.8 | 121.7 | 122.5 | 122.6 |
| $\angle_{N7C9C12}$ | 111.3 | 116.4 | 109.3 | 113.5 | 113.7 |
| $\angle_{C9C12N14}$ | 115.2 | 116.8 | 114.9 | 115.5 | 115.6 |
| $\angle_{C12N14C16}$ | 122.8 | 123.1 | 121.8 | 122.7 | 122.8 |
| $\angle_{N14C16C19}$ | 116.4 | 116.2 | 109.1 | 113.2 | 113.5 |
| $\angle_{C16C19N21}$ | 117.6 | 117.3 | 115.0 | 114.9 | 114.9 |
| $\angle_{C19N21C23}$ | 121.1 | 121.4 | 121.8 | 122.1 | 122.1 |

[a] The total SCF energy of the lowest energy conformation type II β-turn is −660.59785 au. The lowest MP2 energy was obtained for the double-$C_7$ (boat) conformer with −1.98924 au. Note that the $\alpha_R$ conformer is not a local minimum.

= 3.05 Å for the type I β-turn and 2.12 and 2.90 Å for the type II β-turn. The type II β-turn structure is slightly distorted at the 3-21G level toward a $C_7$-turn with an additional intramolecular hydrogen bond between $O_5$ and $H_{15}$ with $R_{O5-H15}$ = 2.33 Å.

Results from the ab initio SCF calculations using a DZP basis set on the five conformations of diglycine diamide are shown in Table III. Plots of the minimized structures of diglycine diamide using the DZP basis set are given in Figure 2. The agreement between the 3-21G-optimized and DZP-optimized structures is very good for double-$C_7$ boat and chair, the extended, and type I β-turn conformations. The superposition of the heavy atoms of the two sets of optimized structures yields root mean square (rms) deviations of 0.07, 0.07, 0.02, and 0.12 Å, respectively. The agreement is, however, notably worse for the type II β-turn, where the rms deviation between the 3-21G and DZP structure is 0.41 Å.

The 3-21G results are, however, in disagreement with the DZP data with respect to the energetic ordering of the diglycine diamide conformers. At the DZP level, both the fully extended and type II β-turn conformations have a lower energy than the double-$C_7$ conformers with the type II β-turn being lowest in energy. The extended conformer is 1.3 kJ/mol less stable than the type II β-turn, and both double-$C_7$ structures and the type I β-turn are less stable by 3.7 and 4.7 kJ/mol, respectively. The type II β-turn is characterized by an intramolecular hydrogen bond with ∠N21H22··O5 = 164.9° and ∠C6O5··H22 = 136.9° and $R_{O5-H22}$ = 2.22 Å and $R_{O5-N21}$ = 3.20 Å. The dihedral angles of the DZP-minimized structure are $\phi_1$ = −72.5°, $\psi_1$ = 18.1°, $\phi_2$ = −104.0°, and $\psi_2$ = 13.4°. These values differ by less than 17° from the "ideal" values quoted in the literature.[2] A distortion of the structure noted at the 3-21G level is not longer found at the DZP level.

The type I β-turn contains a hydrogen bond with $R_{O5-H22}$ = 2.30 Å and $R_{O5-N21}$ = 3.25 Å and ∠N21H22··O5 = 158.7° and ∠C6O5··H22 = 121.8°. The first amide group deviates significantly from planarity with $\omega_1$ = −166.1°. This is due to an additional weak hydrogen bond formed between the amide proton of the central amide group and the nitrogen of the first amide group with $R_{H15-N7}$ = 2.37 Å. The hydrogen-bond geometries in the double-$C_7$ (boat) structure are, at the DZP level, $R_{O5-H15}$ = 2.15 Å, $R_{O5-N14}$ = 3.00 Å, and ∠N14H15··O5 = 142.2° and $R_{O13-H22}$ = 2.18 Å, $R_{O13-N21}$ = 3.02 Å, and ∠N21H22··O13 = 140.7°. The corresponding values for the double-$C_7$ (chair) structure are $R_{O5-H15}$ = 2.15 Å, $R_{O5-N14}$ = 3.00 Å, and ∠N14H15··O5 = 140.4° and $R_{O13-H22}$ = 2.18 Å, $R_{O13-N21}$ = 3.02 Å, and ∠N21H22··O13 = 142.7°. Thus, the deviation of the hydrogen-bond angle ∠NH··O from linearity is significantly larger in double-$C_7$ structures than in β-turn structures.

Both double-$C_7$ structures have very similar hydrogen-bond geometries. Also, the bond lengths of the double-$C_7$ boat and chair conformers do not differ by more than 0.001 Å. The maximum difference for the bond angles is 0.2°.

The dependence of bond lengths and bond angles on the conformation at the DZP level is very small and similar to that found for Ac-Gly-$NHCH_3$.[1] The strongest variation of bond lengths is found for $R_{N7C9}$ which varies from 1.433 Å (extended) to 1.448 Å (double-$C_7$ (boat)). The largest change in bond angles is found for ∠N14C16C19 which varies from 109.1° (extended) to 116.2° (type II β-turn).

It is interesting to compare the present tripeptide bond lengths and bond angles with the data from the corresponding dipeptide conformers.[1] The bond lengths of Ac-Gly-Gly-$NHCH_3$ are all slightly longer than their Ac-Gly-$NHCH_3$ counterparts by 0.001–

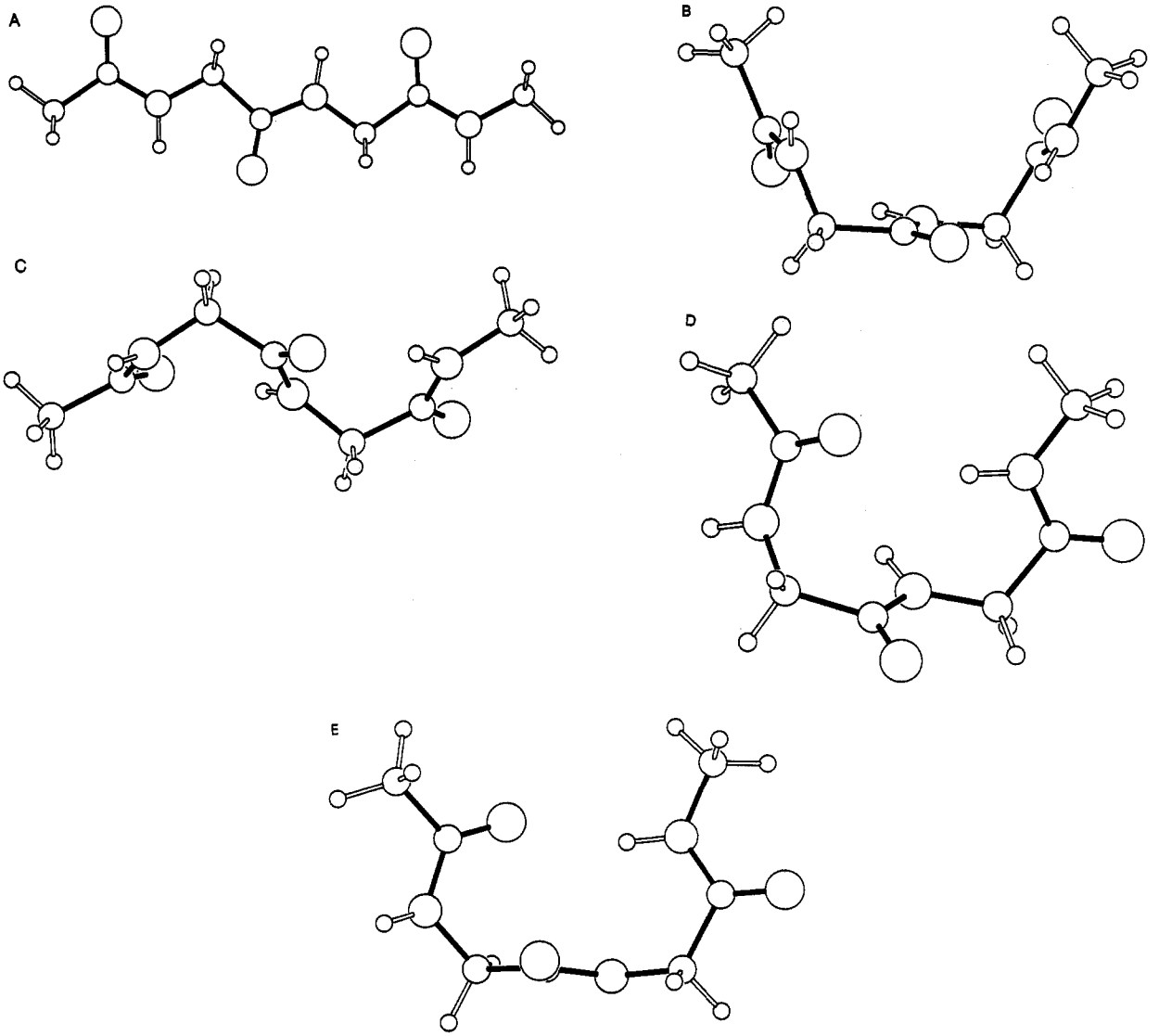


**Figure 2.** Molecular structures of the five low-energy conformations of diglycine diamide: (A) extended, (B) double-$C_7$ (boat), (C) double-$C_7$ (chair), (D) type I $\beta$-turn, and (E) type II $\beta$-turn.

0.005 Å. The most pronounced elongation of a bond is found for the N-terminal amide bond which is stretched by 0.005 Å. The hydrogen-bond lengths $R_{O\cdots H}$ of the diglycine diamide double-$C_7$ conformer are 2.15 and 2.18 Å whereas in Ac-Gly-$NHCH_3$, the corresponding bond length is 2.20 Å. This shortening reflects the strengthening of the hydrogen bonds due to polarization from the adjacent hydrogen bond.

Table III also contains the conformational energies obtained from the single-point SCF-MP2 calculations using the DZP basis set. This calculation again gives a type II $\beta$-turn as the most stable structure of Ac-Gly-Gly-$NHCH_3$. The effect of the correlation energy is most pronounced for the extended conformer which is destabilized by 15 kJ/mol as compared to the other conformers. The correlation energies for the other four conformers do not differ more than 0.8 kJ/mol.

Additional structural details may be derived from the atomic cartesian coordinates of the five local minima of the DZP potential surface of diglycine diamide, given in Table IV.

The partial charges obtained from a fit to the molecular electrostatic potential (MEP) yielded a good representation of the MEP. The rms deviation between the potential calculated from the partial charges and the MEP calculated directly from the wave function was less than 0.8 kJ/mol in all cases. The partial charges for the extended conformer obtained from a fit to the MEP at the point sets 1 and 2 are listed in Table V. It can be seen that the dependence of the partial charges from the particular choice of points is very small. However, it should be noted that the charges are not in line with expectations based on electronegativities. For example, the nitrogen atoms carry a positive charge, and the charges on the polar hydrogen atoms are less positive than those on the hydrogen atoms bound to carbon. A similar situation was encountered for the other conformers of diglycine diamide. Moreover, we obtained completely different charges for the other conformers. For example, $q(N_7)$ varies from –0.002 (type II $\beta$-turn) to 0.350 (type I $\beta$-turn). The result of the fit of a single set of charges to the MEP of all five conformers is also given in Table V. This approach resulted in very large charges on some atoms: $q(C_{16}) = -1.177$, $q(C_{23}) = -1.219$, and etc. Again, the rms deviation of fit (1.6 kJ/mol) was quite good. However, the charges deviate significantly from the corresponding values obtained for Ac-Gly-$NHCH_3$ (see Table VI in ref 1). Therefore, the point charge model given in Table V is not transferable to other molecules.

The point charge model obtained from a fit to the MEP of all five conformers was then used to calculate the MEP at a distance of 10–11 Å away from the molecules (point set 3). The rms deviation between the potential calculated from the partial charges and the ab initio calculated MEP was found to be less than 0.4 kJ/mol. The minima and maxima of the MEP are reproduced with errors of less than 5% in all cases. These results indicate

Table IV. Atomic Cartesian Coordinates of the Five Fully DZP-Optimized Conformers of Diglycine Diamide

| atom | x | y | z | atom | x | y | z | atom | x | y | z |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | (a) type I β-turn | | | | | | |
| H1 | −1.794 | 2.236 | 2.456 | H10 | 1.207 | −1.321 | 2.454 | C19 | −0.328 | 0.020 | −2.328 |
| H2 | −0.613 | 2.291 | 3.760 | H11 | −0.308 | −2.188 | 2.634 | O20 | −0.621 | 0.137 | −3.509 |
| H3 | −0.306 | 3.168 | 2.260 | C12 | 0.220 | −1.682 | 0.641 | N21 | 0.372 | 0.935 | −1.643 |
| C4 | −0.734 | 2.273 | 2.682 | O13 | 1.052 | −2.459 | 0.212 | H22 | 0.620 | 0.802 | −0.685 |
| O5 | 0.933 | 1.151 | 1.346 | N14 | −0.721 | −1.087 | −0.121 | C23 | 0.857 | 2.134 | −2.322 |
| C6 | −0.023 | 1.071 | 2.102 | H15 | −1.394 | −0.496 | 0.312 | H24 | 1.598 | 1.887 | −3.073 |
| N7 | −0.529 | −0.135 | 2.474 | C16 | −0.737 | −1.238 | −1.565 | H25 | 0.041 | 2.647 | −2.814 |
| H8 | −1.276 | −0.177 | 3.133 | H17 | −0.035 | −2.022 | −1.812 | H26 | 1.302 | 2.785 | −1.583 |
| C9 | 0.177 | −1.362 | 2.132 | H18 | −1.715 | −1.525 | −1.923 | | | | |
| | | | | | (b) Type II β-Turn | | | | | | |
| H1 | −1.744 | 1.754 | −2.994 | H10 | 0.726 | −2.371 | −2.290 | C19 | −0.058 | −0.015 | 2.542 |
| H2 | −0.318 | 2.760 | −3.259 | H11 | 1.777 | −1.126 | −1.627 | O20 | −0.358 | 0.062 | 3.699 |
| H3 | −0.671 | 1.405 | −4.359 | C12 | −0.082 | −1.452 | −0.550 | N21 | 0.220 | 1.047 | 1.772 |
| C4 | −0.706 | 1.748 | −3.327 | O13 | −1.274 | −1.507 | −0.594 | H22 | 0.359 | 0.936 | 0.789 |
| O5 | 0.593 | 1.267 | −1.395 | N14 | 0.633 | −1.458 | 0.593 | C23 | 0.088 | 2.391 | 2.282 |
| C6 | 0.083 | 0.859 | −2.404 | H15 | 1.612 | −1.276 | 0.532 | H24 | 0.610 | 2.489 | 3.230 |
| N7 | 0.182 | −0.442 | −2.763 | C16 | 0.022 | −1.394 | 1.896 | H25 | 0.524 | 3.076 | 1.561 |
| H8 | −0.446 | −0.788 | −3.455 | H17 | −0.992 | −1.771 | 1.817 | H26 | −0.956 | 2.665 | 2.443 |
| C9 | 0.742 | −1.385 | −1.832 | H18 | 0.557 | −2.033 | 2.594 | | | | |
| | | | | | (c) Extended | | | | | | |
| H1 | −5.237 | −1.214 | −0.877 | H10 | −1.542 | 1.097 | −0.873 | C19 | 3.154 | 0.183 | 0.000 |
| H2 | −5.245 | −1.201 | 0.886 | H11 | −1.541 | 1.096 | 0.876 | O20 | 3.033 | 1.374 | 0.002 |
| H3 | −6.155 | 0.035 | −0.008 | C12 | −0.462 | −0.528 | 0.000 | N21 | 4.348 | −0.433 | −0.002 |
| C4 | −5.252 | −0.566 | 0.000 | O13 | −0.625 | −1.717 | 0.000 | H22 | 4.380 | −1.428 | −0.002 |
| O5 | −4.141 | 1.542 | −0.002 | N14 | 0.754 | 0.039 | 0.000 | C23 | 5.589 | 0.310 | 0.000 |
| C6 | −4.050 | 0.349 | −0.001 | H15 | 0.872 | 1.032 | 0.001 | H24 | 5.665 | 0.947 | −0.879 |
| N7 | −2.855 | −0.279 | 0.002 | C16 | 1.953 | −0.745 | −0.001 | H25 | 5.669 | 0.939 | 0.884 |
| H8 | −2.784 | −1.275 | 0.002 | H17 | 2.000 | −1.395 | 0.874 | H26 | 6.413 | −0.398 | −0.005 |
| C9 | −1.622 | 0.449 | 0.001 | H18 | 2.000 | −1.393 | −0.877 | | | | |
| | | | | | (d) Double-$C_7$ (Boat) | | | | | | |
| H1 | −1.588 | −1.877 | −3.599 | H10 | 2.387 | −0.853 | −1.919 | C19 | −0.538 | 0.510 | 2.261 |
| H2 | −2.643 | −0.966 | −2.503 | H11 | 1.724 | 0.743 | −2.297 | O20 | −1.651 | 0.847 | 2.546 |
| H3 | −2.293 | −0.323 | −4.103 | C12 | 1.463 | 0.087 | −0.268 | N21 | −0.005 | −0.677 | 2.604 |
| C4 | −1.873 | −0.880 | −3.270 | O13 | 2.165 | −0.503 | 0.509 | H22 | 0.895 | −0.907 | 2.235 |
| O5 | −0.765 | 1.100 | −2.582 | N14 | 0.571 | 1.019 | 0.100 | C23 | −0.773 | −1.695 | 3.278 |
| C6 | −0.721 | −0.093 | −2.700 | H15 | 0.001 | 1.436 | −0.609 | H24 | −1.423 | −1.235 | 4.015 |
| N7 | 0.355 | −0.808 | −2.305 | C16 | 0.393 | 1.429 | 1.477 | H25 | −1.395 | −2.270 | 2.590 |
| H8 | 0.325 | −1.803 | −2.357 | H17 | 1.366 | 1.473 | 1.962 | H26 | −0.092 | −2.376 | 3.783 |
| C9 | 1.545 | −0.188 | −1.766 | H18 | −0.053 | 2.417 | 1.484 | | | | |
| | | | | | (e) Double-$C_7$ (Chair) | | | | | | |
| H1 | −4.785 | 0.035 | −1.212 | H10 | −1.429 | −0.890 | 1.647 | C19 | 2.520 | 0.786 | −0.241 |
| H2 | −5.269 | 0.690 | 0.360 | H11 | −1.555 | −2.398 | 0.732 | O20 | 3.047 | 1.807 | 0.094 |
| H3 | −4.486 | 1.742 | −0.813 | C12 | −0.230 | −0.998 | −0.129 | N21 | 3.034 | −0.434 | −0.006 |
| C4 | −4.510 | 0.729 | −0.421 | O13 | 0.436 | −1.873 | −0.615 | H22 | 2.483 | −1.230 | −0.257 |
| O5 | −2.566 | 1.233 | 0.840 | N14 | 0.081 | 0.304 | −0.218 | C23 | 4.235 | −0.623 | 0.769 |
| C6 | −3.172 | 0.419 | 0.200 | H15 | −0.508 | 0.964 | 0.251 | H24 | 4.954 | 0.154 | 0.528 |
| N7 | −2.708 | −0.836 | 0.018 | C16 | 1.196 | 0.792 | −1.000 | H25 | 4.045 | −0.587 | 1.844 |
| H8 | −3.214 | −1.464 | −0.566 | H17 | 1.289 | 0.187 | −1.900 | H26 | 4.665 | −1.591 | 0.525 |
| C9 | −1.500 | −1.318 | 0.650 | H18 | 0.997 | 1.820 | −1.281 | | | | |

that the averaged charges obtained from a fit to the MEP of all conformers are also useful in representing the long range part of the electrostatic potential.

## 4. Summary and Discussion

We have carried out ab initio SCF calculations using 3-21G and DZP basis sets on six conformers of *N*-acetylglycylglycine *N'*-methylamide (Ac-Gly-Gly-$NHCH_3$: diglycine diamide). In addition, MP2 energies for the DZP-optimized geometries were also calculated. The lowest energy conformation found at the DZP level is a type II β-turn with dihedral angles close to the "ideal" values reported in the literature. The type I β-turn is 4.6 kJ/mol (4.4 kJ/mol at the SCF-MP2 level) less stable. The double-$C_7$ (boat) conformer is calculated to be 3.7 kJ/mol (2.9 kJ/mol at the SCF-MP2 level) less stable than the type II β-turn. The calculations show that both type I and type II β-turns have a high intrinsic energetic stability. In our view, this is additional support for the hypothesis that β-turns play an active role in protein folding.[5,35]

The helical conformation is not a local minimum on the potential energy surface of diglycine diamide. The structure transforms into a type I β-turn. This is not unexpected because diglycine diamide has only three amide groups and can therefore not form an intramolecular hydrogen bond in the α-helical conformation. However, conversion into the β-turn structure involves rather small adjustments in the torsional angles. The rms deviation between the helix and type I β-turn conformers of diglycine diamide is only 0.54 Å at the DZP level. This calculation possibly points to a potential pathway for a conformational change in the very early stage of the protein folding process. As the conversion from a type I β-turn into an α-helix is possible through a low-energy pathway, one may speculate that the formation of a type I β-turn is the first step in forming an α-helix. If this is so, then amino acids known to favor the β-turn structure (e.g., proline) could by this reasoning act as a helix starter. This hypothesis could easily be verified experimentally by comparing the helix contents of two peptides—one with proline as the first amino acid and the other peptide without proline. A similar conclusion,

(35) Ptitsyn, O. B. *FEBS Lett.* **1981**, *131*, 197.

**Table V.** Partial Charges Positioned at the Nuclei Obtained from a Least-Squares Fit of the Molecular Electrostatic Potential for the Extended Conformer and All Conformers[a]

| | extended conformer | | |
|---|---|---|---|
| | set 1 | set 2 | all conformers |
| H1 | 0.131 | 0.136 | 0.241 |
| H2 | 0.133 | 0.139 | 0.257 |
| H3 | 0.184 | 0.180 | 0.261 |
| C4 | −0.424 | −0.424 | −0.956 |
| O5 | −0.432 | −0.429 | −0.628 |
| C6 | 0.070 | 0.060 | 1.008 |
| N7 | 0.343 | 0.343 | −0.628 |
| H8 | 0.058 | 0.040 | 0.406 |
| C9 | 0.103 | 0.128 | −0.701 |
| H10 | −0.024 | −0.031 | 0.283 |
| H11 | −0.026 | −0.035 | 0.259 |
| C12 | 0.051 | 0.092 | 0.825 |
| O13 | −0.524 | −0.533 | −0.599 |
| N14 | 0.329 | 0.307 | −0.232 |
| H15 | 0.127 | 0.128 | 0.321 |
| C16 | −0.288 | −0.302 | −1.177 |
| H17 | 0.160 | 0.166 | 0.371 |
| H18 | 0.160 | 0.166 | 0.335 |
| C19 | −0.160 | −0.169 | 1.204 |
| O20 | −0.352 | −0.348 | −0.672 |
| N21 | 0.165 | 0.169 | −0.575 |
| H22 | 0.125 | 0.123 | 0.451 |
| C23 | −0.004 | −0.002 | −1.219 |
| H24 | 0.042 | 0.041 | 0.392 |
| H25 | 0.041 | 0.042 | 0.386 |
| H26 | 0.010 | 0.011 | 0.388 |
| rmsd | 0.73 | 0.32 | 1.60 |

[a] The root mean square deviation (rmsd) of the fit is given in kJ/mol.

although based on free energy simulations on helices, was drawn recently by Yun et al.[36]

One important aspect of the present work concerns the assessment of the accuracy of the DZP basis set. The inherent limitations of ab initio DZP calculations on peptides have been described previously.[1] First, we note that the calculations on conformational energies of small molecules using the DZP basis set yield results differing from the available experimental data by less than 1.5 kJ/mol. Moreover, in our previous work on dipeptides,[1] we showed that the results from DZP calculations are very close to the data obtained from calculations using a TZP (triple-$\zeta$ plus polarization) basis set. However, one might anticipate that correlation effects will become more important for larger molecules. The present single-point MP2 calculations demonstrate that the effect of the electron correlation on conformational energies of Ac-Gly-Gly-$NHCH_3$ is very small except for the extended conformer. This is not unexpected because differences in correlation energies will be most pronounced between conformers with very different overall shapes. As both the double-$C_7$ and $\beta$-turn conformers have a folded conformation with numerous nonbonded contacts, the correlation effects, such as the dispersion interaction, are roughly the same.

Single-point MP2 calculations on the DZP- and TZP-optimized geometries of the $C_5$ and $C_7$ conformer of Ac-Gly-$NHCH_3$[1,37] gave conformational energy differences $E(C_5) - E(C_7)$ of 10.3 kJ/mol (DZP) and 4.5 kJ/mol (TZP). This has to be compared with the SCF energy differences[1] of −1.3 kJ/mol (DZP) and −1.7 kJ/mol (TZP). Our calculations on Ac-Gly-$NHCH_3$ suggest that the SCF-MP2 calculated conformational energy differences using the DZP basis are slightly overestimated. Therefore, it is expected that the conformational energy of the extended conformer of Ac-Gly-Gly-$NHCH_3$ will be closer to the energies of the other conformers than those calculated by the present SCF-MP2 calculation.

At the DZP level, both $\beta$-turn structures show significant deviations of the amide groups from planarity. The most pronounced deviation is found for type I $\beta$-turn with $\omega_1 = 166°$. In order to elucidate the ability of the DZP basis set to accurately describe the geometry of an amide group under intramolecular strain, we carried out a geometry optimization of cyclohexaglycine with the starting geometry taken from an X-ray crystal structure.[38] In the solid state, cyclohexaglycine forms two type I $\beta$-turns. In the fully DZP-optimized structure of cyclohexaglycine, the largest deviation from 180° of amide dihedral angle is 9.2°. The corresponding value in the experimentally determined structure is 10.0°. Therefore, in the case of cyclohexaglycine, the calculated dihedrals $\omega$ using the DZP basis set are in very good agreement with the experimental solid-state data. A full account of the ab initio calculations on cyclohexaglycine will be given elsewhere.[39] Moreover, a statistical analysis of the observed value for $\omega$ in amides found in the Cambridge structural database[40] carried out by Klebe[41] reveals a distribution of the value of the dihedral $\omega$ between 165° and 180°. This also covers the values calculated for $\omega$ in the present work.

An important contribution to the conformational energy differences of diglycine diamide arises from intramolecular hydrogen bonds. The double-$C_7$ conformer has two intramolecular hydorgen bonds. Both $\beta$-turn conformers contain one hydrogen bond whereas the extended conformer has no hydrogen bonds. Therefore, an underestimation of the hydrogen-bonding interactions will artificially destabilize the double-$C_7$ conformer with respect to the other conformers. At the DZP level, the double-$C_7$ conformer hydrogen-bond lengths $R_{O\cdots N}$ are 3.00 and 3.02 Å, respectively. For the $\beta$-turn conformers, the calculated hydrogen-bond lengths are 0.2 Å longer. The calculated $\beta$-turn hydrogen-bond lengths are also 0.1–0.2 Å longer than what is typically observed in X-ray structures of $\beta$-turns of peptides. However, it is well known that due to crystal field distortions, hydrogen-bond lengths tend to be shorter in crystals than those observed experimentally in hydrogen-bonded dimers.[42] Further support for the accuracy of the DZP basis set in predicting hydrogen-bond lengths comes from calculations on the $H_2O$ dimer[43], which yield a hydrogen-bond length of $R_{O\cdots O}$ = 2.98 Å, in very good agreement with the experimental data.

In summary, it is concluded that SCF calculations using the DZP basis set yield accurate conformational energies with the exception of the comparison of an extended and a folded conformer of a large molecule.

The major error of the 3-21G basis set with respect to peptides appears to be artificial stabilization of the $C_7$-turn. This has two consequences. First, it yields the wrong conformer as the most stable structure. Second, it also perturbs other structures toward the $C_7$-turn motif as compared to those in the DZP-optimized structures. This effect is particularly pronounced for the type II $\beta$-turn, where the rms deviation between the 3-21G- and the DZP-optimized structure is 0.41 Å.

In Table VI, the present ab initio results on the relative energies of the different conformations of diglycine diamide are compared with data obtained from semiempirical calculations employing the AM1 method[22] and from calculations using the force fields CVFF[23] and XPLOR.[24] In disagreement with the present DZP

(36) Yun, R. H.; Anderson, A.; Hermans, J. *Proteins* **1991**, *10*, 219.
(37) Böhm, H. J.; Brode, S. Unpublished results.
(38) Karle, I. L.; Karle, J. *Acta Crystallogr.* **1963**, *16*, 969. One of the cyclohexaglycine molecules not interacting with water molecules was used in the calculation.
(39) Böhm, H. J.; Brode, S. Manuscript in preparation.
(40) (a) Allen, F. H.; Kennard, O.; Taylor, R. *Acc. Chem. Res.* **1983**, *16*, 146. (b) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink-Peters, T.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr. Struct. Crystallogr. Cryst. Chem.* **1979**, *B35*, 2331.
(41) Klebe, G. Habilitationsschrift, Universität Heidelberg, 1990.
(42) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer Verlag: Berlin, 1991.
(43) Frisch, M. J.; DelBene, J. E.; Binkley, J. S.; Schaefer, H. F., III *J. Chem. Phys.* **1986**, *84*, 2279.

**Table VI.** Comparison of Conformational Energies of Diglycine Diamide Obtained from the Present ab Initio Calculations with Results from Semiempirical AM1 Calculations and Force Fields CVFF[23] and XPLOR[24] (All Values Are in kJ/mol)

| | conformer | | | | |
|---|---|---|---|---|---|
| calculation or force field | type I β-turn | type II β-turn | extd | double-$C_7$ (bt) | double-$C_7$ (ch) |
| 3-21G[a] | 7.5 | 8.0 | 7.5 | 0.0 | 0.1 |
| DZP[a] | 4.6 | 0.0 | 1.3 | 3.7 | 3.7 |
| DZP-MP2[a] | 4.4 | 0.0 | 16.4 | 2.9 | 3.2 |
| AM1 | *b* | 11.8 | 25.0 | 0.0 | 0.1 |
| CVFF | 4.3[c] | 3.8[c] | 18.8 | 0.0 | 0.6 |
| XPLOR | *b* | 15.4 | 28.1 | 2.1 | 0.0 |

[a] This work. [b] Not a local minimium, transforms into a double-$C_7$ conformer. [c] Distorted, see text.

ab initio results, all other methods yield a double-$C_7$ conformer as the global energy minimum for diglycine diamide. Another problem occurs, evidently with the representation of the type I β-turn. None of the investigated force fields yields a proper type I β-turn as a local minimum of Ac–Gly–Gly–$NHCH_3$. Also, the AM1 calculation fails to spot the type I β-turn as a local minimum of diglycine diamide.

The CVFF force field yields diffeent local minima for all six conformers of diglycine diamide. A more detailed comparison of the ab initio DZP-calculated tripeptide geometries with those from CVFF force field calculations reveals some important discrepancies. Basically, the examination of the double-$C_7$ and extended conformers yields the same picture as for the corresponding dipeptide.[1] The CVFF calculated geometries of the extended and double-$C_7$ boat and chair conformers can be matched with the corresponding ab initio DZP structures with rms deviations of 0.11, 0.16, and 0.21 Å, respectively. For the bond lengths, the most pronounced differences are observed for N–C bonds. For example, the CVFF force field calculates similar values of $R_{N7-C9}$ for the extended (1.482 Å) and double-$C_7$ conformer (1.485 Å) whereas the ab initio DZP data predict a significantly shorter N7–C9 bond length (1.433 Å) for the extended than that for the double-$C_7$ conformer (1.448 Å). CVFF also underestimates the dependence of ∠N7C9C12 and ∠N14C16C19 on conformational change. For example, the CVFF calculated values for ∠N7C9C12 in the extended and double-$C_7$ conformer are 115.4° and 115.6°, respectively. The corresponding values from the ab initio DZP structures are 109.3° and 113.5°.

The largest deviations between the CVFF calculated and ab initio DZP geometries are found for β-turn conformations. An inspection of the dihedrals in both structures shows that the largest deviations occur for the dihedrals $\psi_1$ and $\psi_2$. For example, a CVFF force field calculation starting from the ab initio DZP-optimized type II β-turn yields a structure with $\psi_1 = -81°$ and $\psi_2 = 76°$. In comparison with the ab initio DZP data, the CVFF force field appears to have a too strong tendency toward $\psi_i = 180°$, 60°, and –60°. β-Turns are characterized by $\psi_2 \approx 0°$ which is disfavored by the CVFF force field.

XPLOR[24] calculated geometries of diglycine diamide are in very good agreement with the present DZP data for four of the five local minima. The rms deviations are 0.04, 0.13, 0.14, and 0.20 Å for the extended, double-$C_7$ boat and chair, and type II β-turn conformations, respectively. The XPLOR potential function does not, however, yield the type I β-turn as a local minimum. Starting from this conformation, the XPLOR force field yields the double-$C_7$ chair conformer as the final structure.

Nishikawa et al.[44] have reported similar problems in their force field calculations on β-turn structures of Ac–Gly–Gly–$NHCH_3$. Nishikawa et al. found that the type I β-turn is not a local minimum of their force field. The calculated type II β-turn also appears to be distorted as compared to that from the present ab initio results ($\psi_1 = 82°$ and $\psi_2 = 46°$).

The present work clearly shows that both type I and II β-turns are low-energy local minima at the potential surface of Ac–Gly–Gly–$NHCH_3$. One requirement for improved force fields for peptides is, therefore, their ability to generate stable β-turn structures for tripeptides.

Another important implication of the present work, together with our previous calculations on dipeptides,[1] concerns the stability of $C_7$-turns. Our results suggest that $C_7$-turns are artificially stabilized in most of the current force fields for peptides and proteins.[21,23,24,44,45] The problem of the $C_7$-turns formed in molecular mechanics force field calculations is not new. This problem was, in the past, attributed to the neglect of solvent molecules in these calculations.[46,47] Moreover, there is evidence that $C_7$-turns are indeed formed in nonpolar solvents such as chloroform.[48] There is also a small number of high-resolution protein structures containing $C_7$-turns.[6] Our results suggest, however, that the stabilization of $C_7$-turns is not just due to the neglect of solvent but also partly a result of using an incorrect force field. This may have serious consequences because force fields are increasingly used in the refinement of X-ray structures of proteins.[24] If the electron density is not well defined, then there is the possibility that the force field will generate $C_7$ conformations which are artifacts of the force field.

The present data, together with our previous results on Ac-Gly-$NHCH_3$ and Ac-Ala-$NHCH_3$[1] and further results on cyclic hexapeptides and larger oligopeptides,[39] form the basis of our current effort to derive an improved parameter set for a peptide force field.

**Acknowledgment.** I would like to thank Stefan Brode for his help with the MP2 calculations.

(44) Nishikawa, K.; Momany, F. A.; Scheraga, H. A. *Macromolecules* **1974**, *7*, 797.
(45) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187.
(46) Mezei, M.; Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1985**, *107*, 2239.
(47) Pettitt, B. M.; Karplus, M. *Chem. Phys. Lett.* **1985**, *121*, 194.
(48) Madison, V.; Kopple, K. D. *J. Am. Chem. Soc.* **1980**, *102*, 4855.