ELSEVIER

Journal of Molecular Structure 346 (1995) 141–152

Journal of
MOLECULAR
STRUCTURE

On the structures of free glycine and α-alanine☆

Attila G. Császár

Department of Theoretical Chemistry, Eötvös University, H-1518 Budapest 112, P.O. Box 32, Hungary

Received 26 July 1994

Abstract

Correlated-level ab initio calculations, including large basis set MP2, have been performed for several conformers of the neutral forms of the amino acids glycine and α-alanine. These calculations resulted in accurate geometric structures and relative energies for the conformers considered. The structural results obtained support the rotational constants measured for the lowest-lying two conformers of both glycine and α-alanine. Energetic and structural results, however, indicate necessary model improvements for existing gas-phase electron diffraction studies of these simplest amino acids. The calculations performed also reveal that, in contrast to what has recently been suggested for this class of compound (R.F. Frey, J. Coffin, S.Q. Newton, M. Ramek, V.K.W. Cheng, F.A. Momany and L. Schäfer, J. Am. Chem. Soc., 114 (1992) 5369), correlated-level geometry optimizations can usually be avoided even if nearly quantitative accuracy is sought in relative energy predictions for the conformers.

1. Preview

Owing to advances in relevant experimental and computational methodologies and facilities, there has recently been a growing interest in the shapes and spectra of biomolecules in the gas phase. This interest stems from several sources, three of them being (a) the search for the origin and signs of life in cool interstellar space, which can be aided by careful investigation of the structures of biomolecules and the related signals in the laboratory; (b) the desire to determine the intrinsic tautomeric and conformational energetics and the underlying potential energy surfaces and hypersurfaces of these species; (c) the stimulation and provision of vital data for the development of better, more efficient and/or reliable computational methods, whether they are non-empirical (like correlated-level ab initio techniques) or empirical (like molecular mechanics) in nature.

Shapes and spectral data are usually determined experimentally by the methods of gas-phase electron diffraction (GED), and high-resolution microwave (MW), millimetrewave (MMW), (matrix) infrared (IR) and/or Raman spectroscopy. It is widely accepted now that most data these experimental techniques provide can also be obtained with considerable accuracy from ab initio computations. Of particular interest for the present investigation are those studies which deal with simple amino acids, the basic building blocks of peptides, the backbone of proteins. Amino acids are known to exist as zwitterions in the crystalline state and in solution. However, in the gas phase, they are present in their neutral forms. Note that this was confirmed, in the case of glycine, quite some time ago by sublimation pressure [1] and

☆ This article is dedicated to Professor Emeritus Kenneth W. Hedberg on the occasion of his 75th birthday.

Elsevier Science B.V.
SSDI 0022-2860(94)09017-3

mass spectral [2] measurements. Perhaps the greatest challenge to structural chemists studying the shapes and spectra of neutral amino acids then lies in the conformational freedom of these molecules, a property mostly retained when amino acids form peptides.

The first successful gas-phase structural measurements for the simplest amino acid, glycine, H_2NCH_2COOH, were reported at about the same time in 1978 by the group of Brown [3] and by Suenram and Lovas [4] following the studies of glycine's MW and MMW spectra, respectively. These pioneering studies were followed by more MW investigations [5,6], resulting in accurate rotational constants for two forms of glycine of lowest energy amongst the various conformers. It is of interest to note that the second most stable conformation of glycine was observed first, as that has a larger dipole moment. The molecular structure of the lowest energy form of glycine in the gas phase was determined from a joint analysis of electron diffraction data and rotational constants by Iijima, Tanaka and Onuma [7] (hereafter ITO). X-ray and neutron diffraction structural studies of the glycine crystal have also been reported [8] but are considered not to be relevant for the present study. Measurement of the IR spectrum of matrix-isolated glycine [9], however, is very relevant but has failed to provide convincing data about the conformations of free glycine.

Following earlier neutron and X-ray diffraction studies [10], the molecular structure of the most stable conformer of α-alanine has been determined by Iijima and Beagley [11] (hereafter IB) by GED following ITO's study on glycine [7]. The MMW spectrum of α-alanine has recently been identified and analysed by Godfrey et al. [12], who determined rotational constants and dipole moments for two conformers corresponding to those previously identified for glycine.

There have been numerous ab initio calculations performed on glycine [13–30] and some on alanine [10b,12,16c,21,31]. Still, reliable ab initio computations using high-level correlated methods with relatively large basis sets are rare and have become available even for glycine only fairly recently [18,30].

No experimental studies, but a matrix IR study of proline [32], on the shapes and spectra of free amino acids other than glycine and α-alanine have ben found, even after an extensive search of the literature. Moreover, theoretical studies on larger amino acids are also extremely scarce [15,19,24,33–37]. Some notable calculations have been performed by Schäfer and co-workers on selected conformers of serine [36a,b], cysteine [36c], valine [36d], and threonine [36d].

2. Introduction

The conformational flexibility of neutral amino acids is now well established both experimentally [3,4,12,16] and theoretically [18,30]. This freedom in the case of glycine led, some 15 years ago, to contradiction between MW and MMW experiments [3,4] and entry-level ab initio Hartree–Fock (HF) calculations [13a] performed prior to those experiments, concerning the lowest-energy form of neutral glycine. The emerging dispute [16a] was soon settled [6,16b] in favour of theory, providing a powerful early example of the utility of ab initio quantum chemical calculations in structural research.

As mentioned in the Preview, not only the experiments but also the theoretical calculations have since been extended to other amino acids, most notably to α-alanine, and also to simple model peptides [21,38]. Among the few related ab initio studies, including at least part of the effect of electron correlation, an important work is due to Frey et al. [21] who investigated the importance of correlation-gradient geometry optimization for molecular conformational analyses in the case of amino acids and peptide analogues. Specifically, they studied three conformations of the amino acid glycine and two conformations of the diamide *N*-formalalanineamide using ab initio theoretical methods at the HF and second-order perturbation (Møller–Plesset, MP2) levels employing basis sets ranging in quality from 3-21G to 6-311G**. The 6-311G** MP2 method was employed as a representative of the correlated methods of ab initio electronic structure theory to obtain optimized geometries of the above conformers beyond the

HF level. With this limited number of examples, Frey et al. [21] wanted to investigate how the geometry of molecules is modified by inclusion of electron correlation during geometry optimization knowing that the "ubiquitous H···N and H···O interactions [present in peptides] may be particularly sensitive to dispersion effects". The principal results emphasized by Frey et al. were

(a) HF "conformational energy maps are intrinsically inaccurate: the nonbonded interactions are incorrectly evaluated not only because dispersion effects are neglected but also because nonbonded distances are wrong due to errors in torsions";

(b) single-point MP2 energies obtained at HF geometries are "potentially highly inaccurate", the "results are of the same limited utility as HF energies at "standard" geometries";

(c) "consequences of optimization [at the HF level] and are not systematic and impossible to predict a priori".

Since some conclusions of Frey et al. [21] might have a substantial effect on future theoretical calculations performed in the field of peptide chemistry, a somewhat more detailed exploration of the problems presented by Frey et al. is of substantial interest and is presented below.

As re-emphasized by Frey et al. [21], in the context of calculations in amino acids and peptides, the ability to judge judiciously detailed ab initio conformational investigations is of extreme importance. To underline this argument, attention is called to the recent GED studies of Iijima and co-workers glycine [7] and alanine [11]. In an extensive theoretical study of the conformers of glycine by Császár [18], resulting in accurate geometries, relative energies, dipole moments and harmonic vibrational frequencies for the most important conformers of this simple amino acid, it was observed that the GED study of ITO on glycine suffered from an unfortunate choice for the model of structural refinement: ITO relied too heavily on some low-level, inaccurate HF conformational energy results. Notably, although they obtained the lowest-energy conformer correctly, their choice for inclusion of a second conformer in the model for structural refinement was not justified by recent high-quality ab initio calculations. Using the notation of Ref. [18] and Fig. 1, i.e. numbering the stable conformers of glycine in the order of their relative energy, the rotation about the C–C bond, introduced by ITO to include a second conformer of glycine in the refinement, leads to the less stable conformer III and not to conformer II. As far as the structure of α-alanine is concerned (for the relevant conformers, see Fig. 2), IB concluded, based on model refinements assuming internal rotation around the C–C bond, that "the vapour of α-alanine consists of one conformer with a high potential barrier around the C–C bond", and consequently carried out their structural refinement for only one conformer. This finding is in clear contrast to indications of a recent MW study [12] and of high-quality ab initio results (see below) which suggest the co-existence of several conformers especially at the temperature of the GED experiment. These and related structural issues are addressed below, based on high-level theoretical calculations, after a short summary of details of the theoretical calculations employed.

3. Computational details

Two basis sets have been selected for this study. The smaller one (B1) is the 6-311++G** basis [39]; it contains 145 contracted Gaussian functions (CGFs) for glycine and 181 for α-alanine. The core part of the larger basis set (B2) was constructed from the 13s8p primitives of Partridge [40] according to (6,3,1,1,1,1) and (4,1,1,1,1) schemes for the s and p functions, respectively, of the C, N, and O atoms and by a (6s/4s) contraction of the unscaled exponents of Huzinaga [41a] for hydrogen and was augmented by three sets of d and two sets of f functions (3d2f) on C, N, and O atoms, by two sets of p and one set of d functions (2p1d) on hydrogens, and by one set of diffuse functions on each atom, resulting in 350 and 436 CGFs for glycine and α-alanine, respectively. All polarization exponents were taken from Dunning [42], and all diffuse function exponents were chosen to be one-third of the lowest related exponents. All d and f sets of both basis sets

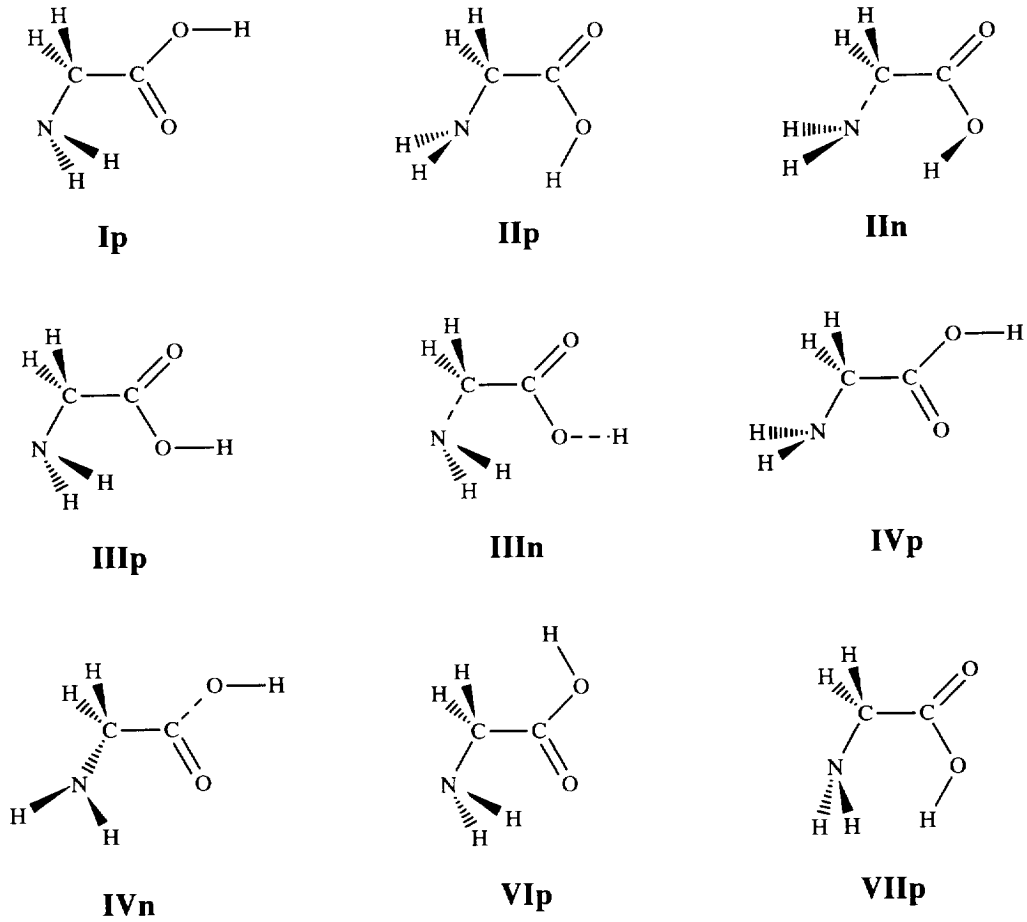

Fig. 1. Conformers of glycine considered in this study.

included only the five and seven pure spherical harmonics, respectively.

Electronic wave functions were determined by the single-configuration, self-consistent-field (SCF), restricted HF method [43–46], and by MP2 [45-48] theory for the incorporation of electron correlation. The 1s core and 1s* virtual orbitals of all second-row elements were kept frozen in all MP2 calculations except geometry optimizations, where all orbitals were correlated.

The geometrical structures of the conformers of glycine and alanine were optimized at the 6-311++G** SCF and MP2(full) levels of theory. The residual Cartesian gradients were in all cases less than 3×10^{-4} hartree bohr^{-1}.

All electronic structure computations were performed with the program packages GAUSSIAN90 [49] and PSI [50].

4. Results and discussion

Tables 1 and 2 contain relative (in cm^{-1}) and total energies (in hartrees) obtained for selected conformers of glycine and α-alanine, respectively. Important geometry parameters of the conformers of glycine and alanine considered are given in Table 3. In Table 4, some theoretical, and the available experimental, rotational constants and dipole moments are collected.

As indicated above, in a recent study [18], the present author found and detailed several

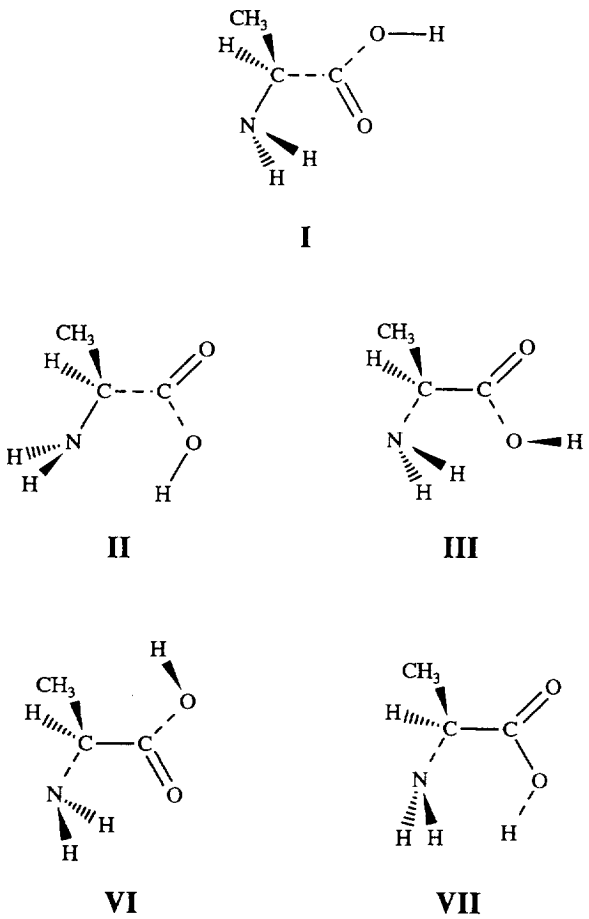

Fig. 2. Conformers of α-alanine considered in this study.

problems with the model selection in the GED study of glycine by ITO. As a result of ITO's unfortunate model assumptions, several discrepancies between theory and experiment exist for the structure of free glycine. One of them concerns the length of the C–C and C–N bonds. The r_α^0 values of ITO are much closer to the values obtained for conformer II than to those of conformer I, whilst according to the calculations, the differences between the C–C and C–N bond lengths of conformers Ip and IIp (IIn) are sizeable, 0.015 (0.013) and 0.019 (0.018) Å, respectively. (Naturally, too much emphasis should not be placed on the absolute values of the calculated bond lengths, as they might be, even at the level of theory employed in this study, somewhat less accurate than the calculated bond angles for which remaining errors of less than 1° can be assumed. Still, in a high-quality structural model the accurately calculated differences between the C–C and C–N bond lengths of conformers I and II should probably be included as constraints.) Another important discrepancy concerns the CCN angle, since its measured value is 2.6° smaller than the probably highly accurate calculated value. Moreover, the measured value is right in between the calculated CCN angles of conformers I and II. The large difference between the calculated CCN bond angle values for these two conformers can easily be rationalized by the trans-angle rule [51]. Note also that, as a less important constraint, the COH angle was fixed during the structure refinements to a 4-21G SCF optimized value, which turns out to be too large by as much as 4.2° if compared with the B1 MP2 value. One plausible explanation for most observed discrepancies is that during recording of the GED data ITO measured scattering from at least two low-energy conformers (those of I and II), while in the process of structural refinement, they basically included parameters of only one of the conformers, those of I. Consequently, the geometry parameters included in their carefully executed fit became (weighted) averages of structures I and II, making some of the structural results, especially those of certain bond angles, rather questionable.

As Table 3 reveals, one of the most characteristic differences between the structure of the conformers of α-alanine and those of glycine is probably the nonplanarity of the heavy-atom framework. The deviations from planarity can be as large as 20° for τ(NCCO). Other structural changes, especially in the bond lengths, seem to be rather small, never larger than about 0.005 Å. These latter differences should hardly be detectable by GED for molecules of this size and of low symmetry. (This is in clear contrast to some changes detailed by IB in Table 5 of Ref. [11]. Note also that while the average experimental C–C bond length, r_α = 1.522(0.004) Å, is in excellent agreement with the BI MP2(full) result of 1.525 Å, the experimental splitting in the C–C distances, 0.038(0.025) Å, is some four times larger than the probably highly accurate theoretical value of 0.009 Å.) Some angles change more noticably; especially

Table 1
Relative and total energies of selected glycine conformers[a]

Method	Ip	IIp	IIn	IIIp	IIIn	IVp	IVn	VIp	VIIp
Geometry optimization									
B1 SCF	0.0	1040.9	988.4	668.5	649.5	1726.1	560.6	2303.4	2936.7
	(−282.925015)	(−282.920272)	(−282.920512)	(−282.921969)	(−282.922055)	(−282.917150)	(−282.922460)	(−282.914520)	(−282.911634)
B1 MP2(full)	0.0	200.2	179.7	557.2	508.7	1637.1	444.1	1997.9	2430.5
	(−283.883499)	(−283.882587)	(−283.882680)	(−283.880960)	(−283.881182)	(−283.876040)	(−283.881489)	(−283.874396)	(−283.872427)
Single-point energy									
B1 MP2//B1 SCF	0.0	342.9	319.5	585.5	540.0	1649.8	465.5	1993.7	2475.0
	(−283.783835)	(−283.782273)	(−283.782380)	(−283.781166)	(−283.781375)	(−283.776318)	(−283.781714)	(−283.774751)	(−283.772558)
B1 SCF//B1 MP2(full)	0.0	1228.4	1183.3	710.2	697.3	1748.1	592.8	2306.9	2995.6
	(−282.921719)	(−282.916122)	(−282.916325)	(−282.918483)	(−282.918542)	(−282.913754)	(−282.919018)	(−282.911208)	(−282.908070)
B2 SCF//B1 MP2(full)	0.0	1185.8	1140.0	754.3	751.0	1682.7	554.4	1989.5	2658.5
	(−282.952342)	(−282.946939)	(−282.947148)	(−282.948905)	(−282.948920)	(−282.944675)	(−282.949816)	(−282.943277)	(−282.940229)
B2 MP2//B1 MP2(full)	0.0	154.7	131.9	583.1	586.2	1589.4	413.7	1730.1	2036.9
	(−284.009528)	(−284.008823)	(−284.008927)	(−284.006871)	(−284.006857)	(−284.002286)	(−284.007643)	(−284.001645)	(−284.000247)
6-311G**MP2// 6-311G**SCF[b]	0.0	452	401						

[a] All total energies in hartrees, all relative energies in cm^{-1}. The B1 (6-311++G**) and B2 basis sets consist of 145 and 350 CGFs, respectively (for details, see text). For notation and pictorial description of the conformers see Ref. [18].
[b] Ref. [17b].

Table 2
Relative and total energies of selected α-alanine conformers[a]

Method	I	II	III	VI	VII
Geometry optimization					
B1 SCF	0.0	847.8	420.5	2350.1	2815.0
	(−321.970926)	(−321.967063)	(−321.969010)	(−321.960218)	(−321.958100)
B1 MP2(full)	0.0	50.5	334.5	2024.9	2324
	(−323.103013)	(−323.102783)	(−323.101489)	(−323.093787)	(−323.092424)
Single-point energy					
B1 MP2//B1 SCF	0.0	163.1	358.0	2023.6	2352.8
	(−322.983731)	(−322.982988)	(−322.982100)	(−322.974511)	(−322.973011)
B1 SCF//B1 MP2(full)	0.0	998.8	442.5	2356.7	2844.6
	(−321.967238)	(−321.962687)	(−321.965222)	(−321.956500)	(−321.954277)
B2 SCF//B1 MP2(full)	0.0	914.5	486.1	2073.8	2523.1
	(−322.001915)	(−321.997748)	(−321.999700)	(−321.992466)	(−321.990419)
B2 MP2//B1 MP2(full)	0.0	−41.0	401.6	1791.8	1966.1
	(−323.242056)	(−323.242243)	(−323.240226)	(−323.233892)	(−323.233098)

[a] All total energies in hartrees, all relative energies in cm^{-1}. The B1 and B2 basis sets consist of 181 and 436 CGFs, respectively (for details, see text). Notation of the conformers follows that of glycine (see also Fig. 2).

pronounced is a decrease of 1.9° in the CCN angle of alanine I compared with gylcine I. None of these structural trends, except the one for the CCN angle, seem to be present in the GED structure [11] of α-alanine. (Note, however, that the CCN angles of the conformers of alanine seem to be smaller by about 2–3° than the CCN angles of the respective glycine conformers; thus, although the experimental decrease of about 3° of the CCN angle of alanine I compared with glycine I is consistent with this general trend, this does not mean that the GED CCN bond angle determined [11] for alanine I is correct. In fact, it may be off by as much as 3°.) Furthermore, during determination of the structure of the lowest-energy conformer of α-alanine IB concluded, based on model refinements assuming internal rotation only around the C–C bond, that "the vapour of α-alanine consists of one conformer with a high potential barrier around the C–C bond" and, consequently, they carried out their structural refinements for only one conformer. The energy results of Table 2 suggest a considerably different picture in contrast to the conclusion of IB, the vapour of α-alanine should contain several conformers. It is of interest to note the result of the B2 MP2 single-point energy calculation (Table 2), which places the energy of conformer II of alanine some 40 cm^{-1} below that of conformer I. While this result cannot be taken literally (inaccuracies of 150–200 cm^{-1} are inherent at this level of theory, especially so since the important [52, 53] effect of core–core and core–valence correlation has not yet been established for this class of compounds), it serves well the point about the small energy difference between the two conformers. In summary, all the new and previous theoretical results seem to indicate that there is no large change between the conformational behaviour of glycine (an amino acid with R = H) and alanine (R = CH_3). This conclusion is further supported by the recent MMW spectroscopy measurements of free α-alanine [12], as Godfrey et al. could identify and analyse spectra belonging to both conformers I and II. Moreover, the accuracy of the theoretical geometry parameters determined for the lowest-lying conformers of alanine in this study is supported by the good agreement between the calculated and experimental [12] rotational constants, albeit the agreement between theory and experiment is not as impressive as for glycine (correction of the experimental rotational constants for effects of vibrations might improve this agreement as it did for conformer I of glycine). It can thus be further concluded that the unfounded assumption about the presence of only one conformer during the course of the GED experiments

Table 3
Selected theoretical and experimental geometry parameters for the most stable conformers of glycine and α-alanine[a]

Parameter	Glycine							α-Alanine						
	Ip			IIp		IIIp		I			II		III	
	Theory		Expt.[b]	Theory		Theory		Theory		Expt.[c]	Theory		Theory	
	SCF	MP2		SCF	MP2	SCF	MP2	SCF	MP2		SCF	MP2	SCF	MP2
r(CC)	1.515	1.519	1.529	1.527	1.533	1.520	1.522	1.523	1.521	1.507	1.531	1.533	1.523	1.520
r(CN)	1.438	1.447	1.466	1.455	1.466	1.440	1.449	1.443	1.452	1.471	1.458	1.468	1.448	1.457
r(CO)	1.329	1.356	1.354	1.317	1.340	1.327	1.356	1.329	1.356	1.347	1.319	1.342	1.331	1.359
r(C=O)	1.183	1.209	1.204	1.180	1.207	1.184	1.209	1.184	1.211	1.192	1.180	1.208	1.184	1.211
r(OH)	0.946	0.968	0.966	0.950	0.981	0.946	0.968	0.946	0.968	[0.977]	0.949	0.980	0.946	0.968
r(CX)	1.085	1.094	1.081	1.084	1.093	1.084	1.094	1.532	1.530	1.544	1.530	1.527	1.530	1.528
$\langle r(\mathrm{NH})\rangle$	1.000	1.014	1.001	0.997	1.012	0.999	1.014	1.001	1.016	[1.014]	1.000	1.015	1.000	1.015
$\angle$(CCN)	115.5	115.6	113.0	112.8	111.3	118.6	119.0	113.3	113.7	110.1	110.3	109.4	114.6	115.2
$\angle$(CCO)	111.5	110.9	111.5	115.6	113.9	114.2	113.0	112.2	111.4	110.3	115.5	114.0	112.9	111.7
$\angle$(CC=O)	125.7	125.7	125.0	121.6	122.3	123.3	123.9	125.4	125.4	125.6	122.0	122.6	124.8	125.5
$\angle$(COH)	109.1	106.3	110.5	108.5	103.9	108.6	105.7	108.9	106.2	[112.3]	108.8	104.0	108.9	106.0
τ(NCCO)	180.0	180.0	180.0	0.0	0.0	0.0	0.0	166.0	161.3		20.4	18.2	45.4	44.5
τ(NCC=O)	0.0	0.0	0.0	180.0	180.0	180.0	180.0	14.5	20.5		161.4	164.4	135.9	
τ(XCCO)		56.9			122.9		123.9	70.3	40.9		101.0	101.7	78.7	78.7
τ(CCOH)	180.0	180.0	180.0	0.0	0.0	180.0	180.0	178.7	176.9	[180.0]	4.4	5.1	178.3	177.4

Quantities in brackets were assumed.

[a] Distances (r) in angstroms, angles ($\angle$ and τ) in degrees. All theoretical values, if not noted otherwise, were obtained at the B1 (6-311++G**) MP2(full) level. Parameter X denotes H for glycine and C for alanine.

[b] Ref. [7]. For uncertainties of the parameters determined see the original publication.

[c] Ref. [11]. For uncertainties of the parameters determined see the original publication.

Table 4
Theoretical and experimental rotational constants and dipole moments of the lowest-energy conformers of glycine and α-alanine[a]

	Glycine					α-Alanine				
	Ip		IIp		IIIp Theory	I		II		III Theory
	Theory	Expt.[b]	Theory	Expt.[b]		Theory	Expt.[c]	Theory	Expt.[c]	
A	10279.0	10341.7 (10297.9)[e]	10175.1	10130.5	9975.0	5084.8	5066.1	5003.8	4973.1	5078.4
B	3877.0	3876.2 (3867.5)[e]	4076.3	4071.5	3989.2	3053.9	3100.9	3202.7	3228.3	2839.4
C	2908.1	2912.4 (2911.0)[e]	3010.9	3007.5	2944.7	2305.6	2264.0	2351.0	2307.8	2503.8
μ[d]	1.3	1.1	6.3	4.6	2.0	1.4	1.8	6.0	5.1	1.8

[a] Rotational constants (*A*, *B*, and *C*) in MHz, and dipole moments (μ) in debyes. All theoretical values, if not noted otherwise, were obtained at the B1 (6-311++G**) MP2(full) level. Note that while theoretical rotational constants refer to equilibrium A_e, B_e, and C_e values, available experimental constants do not.
[b] Refs. [3] and [4].
[c] Ref. [12].
[d] The theoretical dipole moments were obtained at the B1 (6-311++G**) SCF//B1 MP2(full) level.
[e] B_z values, i.e. rotational constants corrected for vibrational effects [7].

resulted in considerable flaws of IB's structural results, including perhaps their conclusion that "the methyl group strongly hinders the rotation of the acid group". Thus, based on all the evidence available, it is suggested that a new joint structural refinement of GED and microwave data, biased toward correlated-level ab initio structural parameters, quadratic force fields, and relative energies, be performed both tor glycine and α-alanine.

Finally, let us investigate the validity of the statements of Frey et al. [21] detailed in the Introduction. It is clear from the data of Tables 1 and 2 that once a sufficiently flexible basis set (such as 6-311++G**, applied throughout this study and in Ref. [18] as the "smaller" basis set) is used[1] (see, for example Ref. [54]) for optimization of the different conformers of glycine and alanine, the choice of reference geometry selected for the subsequent single-point energy calculations is rather irrelevant. For example, Table 1 shows that for conformers IIIp, IIIn, IVp, IVn, VIp and VIIp of glycine, the differences between the relative energies at the B1 SCF//B1 SCF and B1 SCF//B1 MP2(full) results are a mere 32, 48, 22, 32, 4 and 59 cm^{-1}, respectively. Similarly, for conformers III, VI and VII of alanine, the differences are only 22, 7 and 30 cm^{-1}, respectively (see Table 2). Will this situation change by inclusion of electron correlation during the single-point energy calculations? If the B1 MP2//B1 MP2 and B1 MP2//B1 SCF results are compared, it turns out that the differences are small, spreading from 4 (VIp of glycine) to 45 cm^{-1} (VIIp of glycine). All these differences are certainly much smaller than the expected error limits of these calculations. Thus, they can be regarded as insignificant. However, to be fair to Frey et al. [21], it must be noted that they based their conclusion about the "potentially highly inaccurate" single-point MP2 energies obtained at HF geometries on a small set of calculated results, notably on energy difference changes between conformers I and II of glycine (i.e. not among those of the less stable conformers selected above). The B1 SCF//B1 SCF and B1 SCF//B1 MP2, and B1 MP2//B1 MP2 and B1 MP2//B1 SCF energy differences between conformers I and II of glycine (alanine) are 188(151) cm^{-1} and 142(112) cm^{-1}, respectively. These differences, though still rather small do appear less tolerable.

[1] Note that although the necessity to include diffuse functions in basis sets designed for calculations on H-bonded systems has long been recognized. Frey et al. [21] have not used them in their calculations.

However, as larger basis set (B2 MP2) results show, they are again within the expected accuracy of the computations. Thus, it can be concluded that single-point MP2 energies obtained at HF geometries for amino acids are rather accurate as conventional wisdom would imply. (Note also the general tendency of HF theory to overestimate the relative energies of all conformers.) As far as the other two warnings of Frey et al. [21] about inaccurate HF conformational energy maps and the impossibility to predict consequences of HF optimizations are concerned, they seem to be valid though slightly exaggerated. Comparison of geometry parameters obtained from B1 SCF and B1 MP2 optimizations (see Table 3) mostly reveals only rather small changes even in the torsion angles. This suggests that once a reasonably extended basis set has been selected for the computations the geometry parameters are sufficiently well determined even at the HF level of theory, again in accordance with conventional wisdom. Furthermore, it is now obvious that although HF theory completely fails at predicting the energy difference between N–H···O and O–H···N H-bonded conformers (unfortunately, these are just the type of interactions responsible for stabilization of the lowest-energy, and therefore most important, conformers of glycine, alanine, and, indeed, most amino acids), for the other, less important, conformers not only the predicted geometries but also most of the predicted relative energies seem to be reasonably accurate. (See, furthermore, the modest differences between B1 SCF and B1 CCSD energies for conformers IV–VIII of glycine, as reported in Table 2 of Ref. [18].)

In summary, the data presented here send mixed signals to all those trying to perform high-quality ab initio calculations on amino acids and peptide models. While it seems that the HF level of theory with basis sets containing both polarization and diffuse functions is capable of predicting reasonable geometries for these compounds, the relative energies of the most stable conformers stabilized by H-bonds are subject to considerable inaccuracies. Energy differences for the less stable, non-H-bonded forms seem to be reasonably accurate at the HF level, while the still relatively inexpensive MP2 level seems to be able to provide relative energies of quantitative accuracy for even the strongly H-bonded species regardless of whether the underlying geometries were optimized at the HF or correlated levels of theory.

5. Conclusions

Based on the large amount of high-level ab initio data presented in this and a previous [18] study the following conclusions can be drawn about the structures of free glycine and α-alanine and about theoretical calculations performed on neutral amino acids (and possibly on simple peptide analogues).

(1) The larger set of data presented in this paper does not seem to support the earlier conclusions of Frey et al. [21] about the imprecision of HF geometry optimizations of amino acids and their claim that single-point MP2 energies obtained at the HF-optimized geometries are "potentially highly inaccurate". Just the contrary, it seems that MP2 energies are accurate even if they are obtained at HF-optimized geometries.

(2) Although the HF level of theory completely fails at predicting the energy difference between N–H···O and O–H···N H-bonds (unfortunately, these are just the type of interactions responsible for stabilization of the lowest-energy conformers of glycine, alanine, and indeed most amino acids) for the other, less important, conformers not only the predicted geometries but also the predicted relative energies seem to be reasonably accurate. Still, the use of correlated levels of theory in the calculation of most of the relative energies of amino acids is strongly advocated as interest will probably be focused on the low-energy, H-bond-stabilized conformers. Note also the general tendency of the HF level of theory to overestimate the relative energies of all conformers.

(3) It seems that introduction of the methyl group for the amino acid alanine with R = CH_3 to replace one of the hydrogen atoms of glycine (R = H) has a rather small effect on either the geometry or the conformational preferences.

(4) Within the limits of their chosen model, Iijima and co-workers [7,11] performed careful

GED analyses that have resulted in heavy-atom structures of the lowest-energy conformers of glycine and α-alanine. Their models for structure refinement, however, should be revised based on the newly available theoretical data (and, in the case of α-alanine, on the recently measured rotational constants), and the analysis should be repeated since considerable changes in some structural parameters are expected as the result of the suggested model revisions.

Acknowledgement

Preparation of this article and some of the research described were carried out during a stay of the author, generously supported by the PHARE ACCORD program (contract no. H9112-0379), with Professor N.C. Handy (Department of Chemistry, University of Cambridge, Cambridge, UK), whose hospitality is gratefully acknowledged. The research described was also supported by the Scientific Research Foundation of Hungary (OTKA F013962). Some of the calculations reported here have been performed on an IBM RS/6000 m560 workstation whose purchase was made possible by a grant from "Catching up with European Higher Education" (FEFA II/265).

References

[1] S. Takagi, H. Chihara and S. Seki, Bull. Chem. Soc. Jpn., 32 (1959) 84.
[2] G. Junk and H. Sevec, J. Am. Chem. Soc., 85 (1963) 839.
[3] R.D. Brown, P.D. Godfrey, J.W.V. Storey and M.P. Bassez, J. Chem. Soc., Chem. Commun., (1978) 547.
[4] R.D. Suenram and F.J. Lovas, J. Mol. Spectrosc., 72 (1978) 372.
[5] R.D. Brown, P.D. Godfrey, J.W.V. Storey, M.P. Bassez, B.J. Robinson, R.A. Batchelor, M.G. McCulloch, O.E.H. Rydbeck and A.G. Hjalmarson, Mon. Not. R. Astron. Soc., 186 (1979) 5.
[6] R.D. Suenram and F.J. Lovas, J. Am. Chem. Soc., 102 (1980) 7180.
[7] K. Iijima, K. Tanaka and S. Onuma, J. Mol. Struct., 246 (1991) 257.
[8] J. Almlöf, A. Kvick and J.O. Thomas, J. Chem. Phys., 59 (1973) 3901.
[9] (a) Y. Grenie, J.C. Lassegues and C.G. Lagrange, J. Chem. Phys., 53 (1970) 2980.
(b) Y. Grenie and C.G. Lagrange, J. Mol. Spectrosc., 41 (1972) 240.
[10] (a) M.S Lehmann, T.F. Koetzle and W.C. Hamilton, J. Am. Chem. Soc., 94 (1972) 2657.
(b) R. Destro, R.E. Marsh and R. Bianchi, J. Phys. Chem., 92 (1988) 966.
[11] K. Iijima and B. Beagley, J. Mol. Struct., 248 (1991) 133.
[12] P.D. Godfrey, S. Firth, L.D. Hatherley, R.D. Brown and A.P. Pierlot, J. Am. Chem. Soc., 115 (1993) 9687.
[13] (a) S. Vishveshwara and J.A. Pople, J. Am. Chem. Soc., 99 (1977) 2422.
(b) Y.-C. Tse, M.D. Newton, S. Vishveshwara and J.A. Pople, J. Am. Chem. Soc., 100 (1978) 4329.
[14] J.A. Ryan and J.L. Whitten, J. Am. Chem. Soc., 94 (1972) 2396.
[15] O. Kikuchi, T. Natsui and T. Kozaki, J. Mol. Struct. (Theochem), 66 (1990) 103.
[16] (a) H.L. Sellers and L. Schäfer, J. Am. Chem. Soc., 100 (1978) 7728.
(b) L. Schäfer, H.L. Sellers, F.J. Lovas and R.D. Suenram, J. Am. Chem. Soc., 102 (1989) 6566.
(c) K. Siam, V.J. Klimkowski, J.D. Ewbank, C. Van Alsenoy and L.Schäfer, J. Mol. Struct. (Theochem), 110 (1984) 171.
[17] (a) M. Ramek, Int. J. Quantum Chem., Quantum Biol. Symp., 17 (1990) 45.
(b) M. Ramek, V.K.W. Cheng, R.F. Frey, S.Q. Newton and L. Schäfer, J. Mol. Struct. (Theochem), 235 (1991) 1.
(c) M. Ramek and V.K.W. Cheng, Int. J. Quantum Chem., S19 (1992) 15.
[18] A.G. Császár, J. Am. Chem. Soc., 114 (1992) 9569.
[19] L.R. Wright and R.F. Borkman, J. Am. Chem. Soc., 102 (1980) 6207.
[20] (a) P. Palla, C. Petrongolo and J. Tomasi, J. Phys. Chem., 84 (1980) 435.
(b) R. Bonaccorsi, P. Palla and J. Tomasi, J. Am. Chem. Soc., 106 (1984) 1945.
[21] R.F. Frey, J. Coffin, S.Q. Newton, M. Ramek, V.K.W. Cheng, F.A. Momany and L. Schäfer, J. Am. Chem. Soc., 114 (1992) 5369.
[22] C.E. Dykstra, R.A. Chiles and M.D. Garrett, J. Comput. Chem., 2 (1981) 266.
[23] J. Voogd, J.L. Derissen and F.B. van Duijneveldt, J. Am. Chem. Soc., 103 (1981) 7701.
[24] S. Millefiori and A. Millefiori, J. Mol. Struct., 91 (1983) 391.
[25] H. Basch and W.J. Stevens, Chem. Phys. Lett., 169 (1990) 275.
[26] (a) J.H. Jensen and M.S. Gordon, J. Am. Chem. Soc., 113 (1991) 7917.
(b) J.H. Jensen, K.K. Baldridge and M.S. Gordon, J. Phys. Chem., 96 (1992) 8340.
[27] D. Yu, D.A. Armstrong and A. Rauk, Can. J. Chem., 70 (1992) 1762.
[28] Y. Ding and K. Krogh-Jespersen, Chem. Phys. Lett., 199 (1992) 261.

[29] A. Vijay and D.N. Sathyanarayana, J. Phys. Chem., 96 (1992) 10735.
[30] C.-H. Hu, M. Shen and H.F. Schaefer, III, J. Am. Chem. Soc., 115 (1993) 2923.
[31] H.L. Sellers and L. Schäfer, Chem. Phys. Lett., 63 (1979) 609.
[32] I.D. Reva, S.G. Stepanian, A.M. Plokhotnichenko, E.D. Radchenko, G.G. Sheina and Y.P. Blagoi, J. Mol. Struct., 318 (1994) 1.
[33] (a) M. Masamura, J. Mol. Struct. (Theochem), 37 (1987) 293.
(b) M. Masamura, J. Mol. Struct. (Theochem), 45 (1987) 277.
(c) M. Masamura, J. Mol. Struct. (Theochem), 41 (1988) 299.
[34] P. Tarakeshwar and S. Manogaran, J. Mol. Struct. (Theochem), 111 (1994) 205.
[35] D.A. Dixon and W.N. Lipscomb, J. Biol Chem., 251 (1976) 5992.
[36] (a) C. van Alsenoy, J.N. Scarsdale, H.L. Sellers and L. Schäfer, Chem. Phys. Lett., 80 (1981) 124.
(b) C. van Alsenoy, S. Kulp, K. Siam, V.J. Klimkowski, J.D. Ewbank and L. Schäfer, J. Mol. Struct. (Theochem), 50 (1988) 169.
(c) L. Schäfer, K. Siam, V.J. Klimkowski, J.D. Ewbank, and C. van Alsenoy, J. Mol. Struct. (Theochem), 63 (1990) 361.
(d) L. Schäfer, S.Q. Kulp-Newton, K. Siam, V.J. Klimkowski and C. van Alsenoy, J. Mol. Struct., 209 (1990) 373.
[37] E. Clementi, F. Cavallone and R. Scordamaglia, J. Am. Chem. Soc., 99 (1976) 5531.
[38] (a) L. Schäfer, C. van Alsenoy and J. Scarsdale, J. Chem. Phys., 76 (1982) 1439.
(b) V.J. Klimkowski, L. Schäfer, F.A. Momany and C. van Alsenoy, J. Mol. Struct., 124 (1985) 143.
(c) L. Schäfer, V.J. Klimkowski, F.A. Momany, H. Chuman and C. van Alsenoy, Biopolymers, 23 (1984) 2335.
(d) T. Head-Gordon, M. Head-Gordon, M.J. Frisch, C. Brooks and J.A. Pople, J. Am. Chem. Soc., 113 (1991) 5989.
(e) M. Aida, Bull. Chem. Soc. Jpn., 66 (1993) 3423.
(f) C. van Alsenoy, M. Cao, S.Q. Newton, B. Teppen, A. Perczel, I.G. Csizmadia, F.A. Momany and L. Schäfer, J. Mol. Struct. (Theochem), 105 (1993) 149.
(g) W. Viviani, J.L. Rivail, A. Perczel and I.G. Csizmadia, J. Am. Chem. Soc., 115 (1993) 8321.
(h) A. Perczel, M.A. McAllister, P. Császár and I.G. Csizmadia, J. Am. Chem. Soc., 115 (1993) 4849.
(i) W. Viviani, J.L. Rivail and I.G. Csizmadia, Theor. Chim. Acta, 85 (1993) 189.
(j) L. Schäfer, S.Q. Newton, M. Cao, A. Peeters, C. van Alsenoy, K. Wolinski and F.A. Momany, J. Am. Chem. Soc., 115 (1993) 272.
[39] (a) R. Krishnan, J.S. Binkley, R. Seeger and J.A. Pople, J. Chem. Phys., 72 (1980) 650.
(b) T. Clark, J. Chandrasekhar, G.W. Spitznagel and P.v.R. Schleyer, J. Comput. Chem., 4 (1983) 294.
(c) M.J. Frisch, J.A. Pople and J.S. Binkley, J. Chem. Phys., 80 (1984) 3265.
[40] H. Partridge, Near-Hartree–Fock Quality Gaussian Type Orbital Basis Sets for the First- and Third-Row Atoms, NASA Technical Memorandum 101044, NASA, 1989.
[41] (a) S. Huzinaga, J. Chem. Phys., 42 (1965) 1293.
(b) W.D. Allen and H.F. Schaefer, III, J. Chem. Phys., 89 (1988) 329.
[42] T.H. Dunning, J. Chem. Phys., 90 (1989) 1007.
[43] P. Pulay, Mol. Phys., 17 (1969) 197.
[44] P. Pulay, Mol. Phys., 18 (1970) 473.
[45] J.A. Pople, R. Krishnan, H.B. Schlegel and J.S. Binkley, Int. J. Quantum Chem., Symp., 13 (1979) 225.
[46] W.J. Hehre, L. Radom, P.v.R. Schleyer and J.A. Pople, Ab Initio Molecular Orbital Theory, Wiley-Interscience, New York, 1986.
[47] C. Møller and M.S. Plesset, Phys. Rev., 46 (1934) 618.
[48] J.A. Pople, J.S. Binkley and R. Seeger, Int. J. Quantum. Chem., Symp., 10 (1976) 1.
[49] M.J. Frisch et al., GAUSSIAN 90, Gaussian Inc., Pittsburgh, PA, 1990.
[50] PSI Version 2.0, Psitech, Watkinsville, GA, 1991.
[51] M. Räsänen, A. Aspiala, L. Homanen and J. Murto, J. Mol. Struct., 96 (1982) 81.
[52] W.D. Allen, A.L.L. East and A.G. Császár, in J. Laane, M. Dakkouri, B. van der Veken and H. Oberhammer (Eds.), Structures and Conformations of Non-Rigid Molecules, Kluwer, Dordrecht, 1993, p. 343.
[53] A.G. Császár and W.D. Allen, J. Chem. Phys., 102 (1994) in press.
[54] M.J. Frisch, J.A. Pople and J.E. Del Bene, J. Phys. Chem., 89 (1985) 3664.