

## Conformers of Gaseous $\alpha$ -Alanine

Attila G. Császár

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

Received: November 15, 1995<sup>⊗</sup>

Accurate geometries, relative energies, rotational and quartic centrifugal distortion constants, dipole moments, harmonic vibrational frequencies, and infrared intensities have been determined from ab initio calculations for 13 conformers, corresponding to minima on the potential energy surface, of the neutral form of the amino acid  $\alpha$ -alanine. The theoretical relative energy predictions, even after zero-point vibrational energy corrections, differ significantly from the experimental lower limits deduced from millimeter wave spectra. Energetic and structural results indicate necessary model improvements for a gas-phase electron diffraction study of  $\alpha$ -alanine. The structural results obtained support the molecular constants measured for two conformers of low relative energy. The accurate calculated rotational and quartic centrifugal distortion constants and vibrational frequency data for the other conformers should aid in identification of the conformers by rotational and vibrational spectroscopies, respectively.

### Introduction

Amino acids exist as zwitterions in the crystalline state and in solution, stabilized by electrostatic, polarization, and H-bonding interactions with their environment.<sup>1–4</sup> In the gas phase, where these intermolecular interactions have no effect, amino acids exist in their neutral form.<sup>1,2</sup> The conformational flexibility of free, neutral amino acids is now well established both experimentally<sup>5–14</sup> and theoretically.<sup>15–39</sup> This freedom is due to the fact that rotamers of amino acids may form intramolecular H-bonds of different strength, while steric strain and repulsion of lone electron pairs have a destabilizing effect. The conformational changes, resulting from the balance of all these effects, are accompanied by rather small overall changes in the total energy of the system.

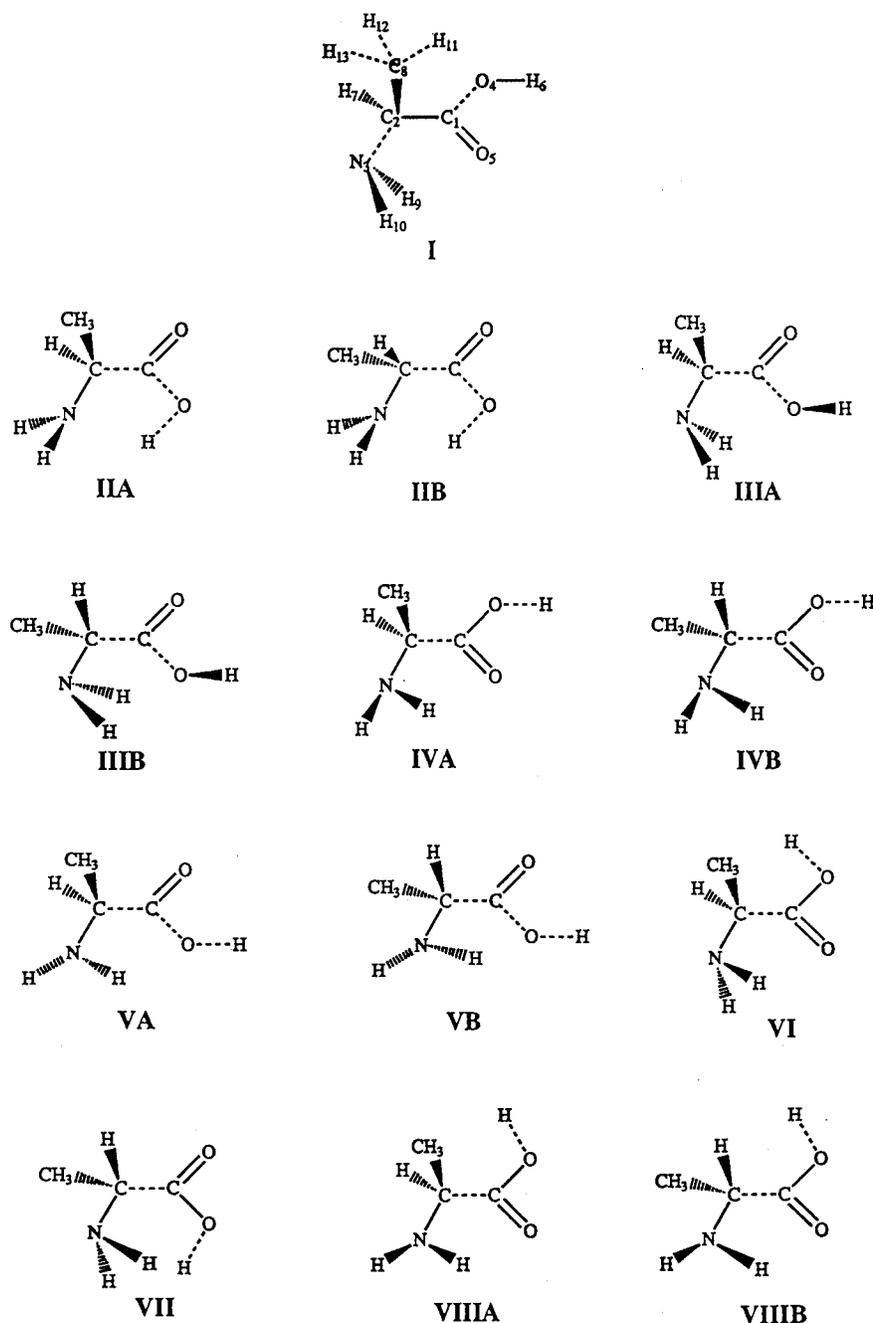
The interest in the shapes and spectra of the conformationally flexible free amino acids, the building blocks of peptides, the backbones of proteins, stems from at least three reasons: (a) the search for the origin and signs of life in cool interstellar space, which can be aided by careful laboratory investigations of the structures and the related signals of these biomolecules; (b) the desire to establish the intrinsic tautomeric and conformational energetics and the underlying potential energy surfaces and hypersurfaces of these species which probably determine these characteristics of polypeptides and proteins; and (c) to stimulate and to provide vital data for the development of better, more efficient and/or reliable computational methods, whether they are nonempirical (like correlated-level ab initio and density functional<sup>39</sup> techniques) or empirical (like molecular mechanics) in nature.

Only relatively few experimental studies have been published on the shapes and spectra of free amino acids. Investigations<sup>5–8</sup> of the microwave (MW) and millimeter wave (MMW) spectra of glycine, H<sub>2</sub>NCH<sub>2</sub>COOH, the simplest amino acid, resulted in accurate rotational constants for the two conformers of lowest energy. The MMW spectrum of  $\alpha$ -alanine has been identified and analyzed by Godfrey et al.,<sup>9</sup> who determined rotational and quartic centrifugal distortion constants and dipole moments for two conformers corresponding to those identified previously for glycine. Godfrey et al.<sup>9</sup> also estimated relative abundances of certain  $\alpha$ -alanine conformers at the temperature of their experi-

ment, about 530 K, and determined approximate lower limits for conformational energies relative to that of  $\alpha$ -alanine **I** of  $\sim 800$  cm<sup>-1</sup> for **IIA**,  $\sim 700$  cm<sup>-1</sup> for **IIIA**, and  $\sim 1200$  cm<sup>-1</sup> for **VI** (for numbering of the conformers adopted in this study see Figure 1). The heavy-atom positions of the most stable conformers of glycine and  $\alpha$ -alanine have been determined by Iijima, Tanaka, and Onuma<sup>10</sup> and by Iijima and Beagley,<sup>11</sup> respectively, using the technique of gas-phase electron diffraction (GED). As far as the structure of  $\alpha$ -alanine is concerned, Iijima and Beagley<sup>11</sup> concluded, on the basis of model refinements assuming internal rotation only around the C–C bond, that “the vapour of  $\alpha$ -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 contrast to indications of recent MMW<sup>9</sup> and ab initio<sup>16,34b</sup> studies, which suggest the coexistence of several conformers at the temperature of the GED experiment (about 500 K). Further experimental studies by low-resolution photoelectron spectroscopy<sup>13</sup> did not yield any conformational information about  $\alpha$ -alanine. Even after an extensive search of the literature no experimental studies, but a matrix infrared study of proline,<sup>12a</sup> on the shapes and spectra of free amino acids other than glycine and  $\alpha$ -alanine have been found.

Although there have been a number of nonempirical theoretical studies<sup>15–39</sup> published on free amino acids, only a few of them are exhaustive and dependable. This is due partly to the size of these systems and partly to demonstrated problems<sup>16,18,19,27,31</sup> with entry-level theoretical calculations on these species, including, most importantly, the inability of the restricted Hartree–Fock (RHF) level of theory to predict the relative energies of the N–H $\cdots$ O and O–H $\cdots$ N H-bonded conformers in these systems. Most theoretical studies have been published on glycine (for a list of studies published prior to 1992 see ref 18), and only a few studies are available for  $\alpha$ -alanine.<sup>9,16,24,25,34</sup> Furthermore, none of the latter studies, except the most recent ones,<sup>16,25,34</sup> included results from correlated-level (post-Hartree–Fock) calculations. In entry-level (4-21G RHF) ab initio calculations Schäfer and co-workers examined two  $\alpha$ -alanine conformers of low relative energy<sup>24a,b</sup> and the N–C–C=O torsional potential energy curve.<sup>24c</sup> In an extensive study at the RHF level Godfrey et al.<sup>9</sup> reported 6-31G\*\* RHF relative energies and several molecular parameters of six conformers of  $\alpha$ -alanine (conformers **I**, **II**, **III**, **IV**,

<sup>⊗</sup> Abstract published in *Advance ACS Abstracts*, February 1, 1996.



**Figure 1.** Conformers of  $\alpha$ -alanine considered in this study and their atom numbering.

V, and VI of Godfrey et al.<sup>9</sup> correspond to conformers I, IIA, IIB, IIIA, IIIB, and VI, respectively, of Figure 1). Császár<sup>16</sup> reported limited 6-311++G\*\* MP2 results for  $\alpha$ -alanine I, IIA, IIIA, VI, and VII. In their detailed study Gronert and O'Hair<sup>34b</sup> located 10 conformers of  $\alpha$ -alanine and characterized them by 6-31G\* RHF and MP2 calculations. Their numbering scheme is somewhat different from the one employed in this study; their conformers 1–10 correspond to I, IIA, IIIA, IIB, IVA, IVB, IIIB, VA, VB, and VI, respectively, of this study. Cao et al.<sup>25</sup> determined all 13 conformers investigated in this study; their conformers 1–13 correspond to I, IIB, IIA, IIIA, IIIB, IVA, IVB, VA, VB, VI, VII, VIIIA, and VIIIB. Some notable small basis ab initio RHF calculations have been performed by Schäfer and co-workers on larger amino acids, namely, on selected conformers of serine,<sup>20,21</sup> cysteine,<sup>22</sup> valine,<sup>23</sup> and threonine,<sup>23</sup> and by Sapse et al.<sup>35</sup> on proline. Recently, Gronert and O'Hair<sup>34</sup> reported results of medium-level RHF and MP2 calculations for 51 conformers of serine and 42 conformers of cysteine.

### Computational Details

Theoretical computations performed in this study follow a strategy based on high-quality ab initio results obtained for glycine.<sup>18</sup> This strategy takes advantage of the observation<sup>18</sup> that in most cases higher order (MP3, MP4, MP $\infty$ , CCSD, and CCSD(T); definitions of these correlated-level ab initio methods are given below) contributions to the energy differences of the conformers of glycine either are small or seem to cancel out; thus, large basis set MP2 calculations should result in highly accurate relative energies.

Several basis sets have been selected for this study. Most of them contain both polarization and diffuse functions, as the need for including these functions in the basis set for calculations on hydrogen-bonded systems has long been recognized.<sup>40</sup> The smallest basis set employed is the 6-31G\*\* basis.<sup>41</sup> The 6-311++G\*\* basis,<sup>42</sup> used most often in this study, contains 181 contracted Gaussian functions (CGFs) for  $\alpha$ -alanine. The

correlation-consistent, polarized-valence cc-pVTZ, aug-cc-pVTZ, and cc-pVQZ basis sets of Dunning and co-workers<sup>43,44</sup> have also been used in some calculations on  $\alpha$ -alanine and *cis*- and *trans*-formic acid. (Note that only the augmented (aug) basis set contains diffuse functions.) The cc-pVTZ and aug-cc-pVTZ basis sets contain 278 and 437 CGFs for  $\alpha$ -alanine, respectively. The same extended basis set as used in a previous study on glycine<sup>18</sup> (denoted here as 7s6p3d2f) has been employed in some single-point RHF and MP2 calculations. Its core part was constructed from the 13s8p primitives of Partridge<sup>45</sup> 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<sup>46</sup> 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 s and p functions on each atom, resulting in 436 CGFs for  $\alpha$ -alanine. All polarization function exponents were taken from Dunning;<sup>43</sup> all diffuse function exponents were chosen to be one-third of the lowest related s and p exponents. All d and f sets of all basis sets, except 6-31G\*\*, where the six Cartesian d functions have been employed, included only the five and seven pure spherical harmonics, respectively.

Electronic wave functions were determined by the single-configuration, self-consistent-field, restricted Hartree–Fock (RHF) method,<sup>47–49</sup> by second-, third-, and fourth-order Møller–Plesset<sup>50</sup> theory, i.e., MP2, MP3, and MP4(SDTQ), by coupled cluster (CC) methods including all single and double excitations (CCSD) and in cases, additionally, a perturbative correction for contributions from connected triple excitations (CCSD(T)),<sup>51–54</sup> and by a density functional (DFT) approach usually abbreviated as B3LYP.<sup>55</sup> Extrapolation of the perturbation series to estimate the infinite-order energy (MP $\infty$ , obtained by estimating the exact correlation energy within a given one-particle basis set) was performed using a formula suggested by Pople and co-workers.<sup>56,57</sup> The  $\mathcal{T}_1$  diagnostic values of coupled cluster theory<sup>54c</sup> are around 0.013 and 0.017 for the different conformers of  $\alpha$ -alanine and formic acid, respectively, suggesting that both molecules can adequately be described by single-reference-based electron correlation methods. For  $\alpha$ -alanine(formic acid) the six(three) lowest 1s core orbitals and the six(three) highest 1s\* virtual orbitals were kept frozen in all MP $n$  and CC treatments except 6-311++G\*\* MP2(full) geometry optimizations, where all orbitals were correlated.

The geometrical structures of the conformers of  $\alpha$ -alanine were optimized at the 6-31G\*\* RHF and at the 6-311++G\*\* RHF, DFT(B3LYP), and MP2(full) levels of theory. The residual Cartesian gradients were in all cases less than  $1 \times 10^{-4}$  hartrees/bohr. To determine whether the optimized structures correspond to minima on the potential energy surface of  $\alpha$ -alanine and to avoid problems arising from the nonzero force dilemma,<sup>58</sup> the 6-311++G\*\* RHF Cartesian quadratic force constants were determined at the respective optimized geometries employing analytic second derivatives.<sup>59,60</sup>

Electronic structure computations were performed with the program packages ACES II,<sup>61</sup> GAUSSIAN94,<sup>62</sup> PSI,<sup>63</sup> and TURBOMOLE.<sup>64</sup> For the calculation of the quartic centrifugal distortion constants the package ASYM20 was employed.<sup>65</sup>

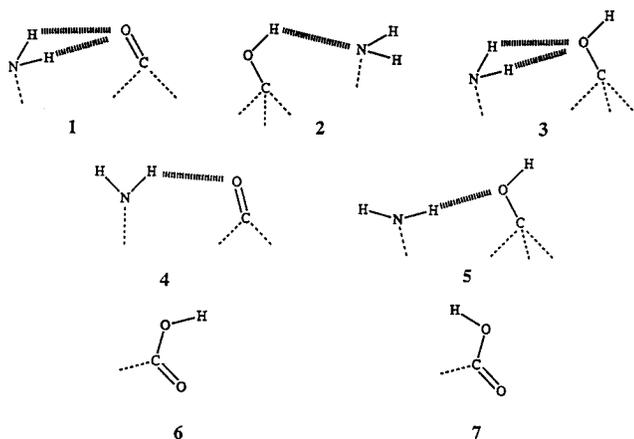
## Results and Discussion

The stable conformers of  $\alpha$ -alanine can be built up perhaps most straightforwardly from the eight conformers of glycine (an amino acid with R = H) corresponding to minima on its potential energy surface (PES).<sup>18</sup> Three of the glycine conform-

ers belong to the  $C_s$  point group, while five have  $C_1$  symmetry. The  $C_s$  symmetry conformers (**I**, **VI**, and **VII** in the notation of ref 18) have equivalent  $C_\alpha$  hydrogens, and as suggested by the systematic calculations of this study, upon substitution of one of them by a methyl group this equivalence is maintained, resulting in only one corresponding stationary point on the PES of  $\alpha$ -alanine. On the other hand, the  $C_1$  symmetry conformers have nonequivalent  $C_\alpha$  hydrogens, and according to detailed calculations presented below, upon the R = CH<sub>3</sub> substitution this nonequivalence seems to be maintained, resulting in two different  $\alpha$ -alanine conformers, corresponding to conformers **II<sub>n</sub>**, **III<sub>n</sub>**, **IV<sub>n</sub>**, **V<sub>n</sub>**, and **VIII<sub>n</sub>** of glycine. Thus, the close correspondence between the conformational behavior of  $\alpha$ -alanine and glycine results in 13 relevant conformers for  $\alpha$ -alanine. It is perhaps of general interest to note at this point that during their conformational search Godfrey et al.<sup>9</sup> missed more than half of the stable conformers of  $\alpha$ -alanine; Gronert and O'Hair,<sup>34b</sup> despite their own warning about the "need to undertake a systematic survey of all the reasonable rotamers in amino acid studies" and their careful search for conformers employing semiempirical (AM1 and PM3) and entry-level ab initio (3-21G RHF) calculations, failed to identify three stable conformers on the PES of  $\alpha$ -alanine; and only Schäfer and co-workers<sup>25</sup> located all 13 conformers. The likely existence of all 13 conformers but **IIIB** is established in the present study and in ref 25.

Table 1 contains relative (in cm<sup>-1</sup>) and total (in hartrees) energies obtained for all the conformers. (See Figure 1 for the notation applied. The numbering scheme follows that adopted for glycine;<sup>18</sup> that is, numbering of the conformers reflects the relative energies of the distinct  $\alpha$ -alanine conformers obtained from 6-311++G\*\* MP2(full) geometry optimizations, while in those cases where two conformers belong to the same basic structure, which formally appear as the D and L enantiomer pairs on Figure 1, the letters **A** and **B** following the Roman numbers are used to distinguish them.) Although  $\alpha$ -alanine has 33 independent geometry parameters, structural results are given in Table 2 only for the 21 most important ones. Table 3 contains some theoretical and the available experimental rotational and quartic centrifugal distortion constants and dipole moments. Harmonic vibrational frequencies, infrared intensities, and (relative) zero-point vibrational energies (ZPVEs) of all  $\alpha$ -alanine conformers are presented in Table 4. The high-quality ab initio data presented in Table 5 for the energy difference between *cis*- and *trans*-formic acid clearly support the computational strategy employed in this study, as the post-MP2 energy corrections seem to be always smaller than  $\sim 50$  cm<sup>-1</sup>.

**Energies.** Similarly to glycine,<sup>18</sup> the relative energies of the  $\alpha$ -alanine conformers **I–VIII<sub>B</sub>** are determined by the interplay of the five different types of H-bonds (**1–5** of Figure 2) that can be formed in these compounds, the *cis* vs *trans* arrangement of the carboxylic functional group (**6** and **7** of Figure 2), the steric strain, and the repulsion of lone electron pairs on the N and O atoms. A rather interesting approximate, qualitative picture emerges if one completely neglects the effects of steric strain and lone electron pair repulsion and assigns independent energy contributions to the H-bonds and the carboxylic functional group. Naturally, H-bonding properties cannot be ascribed, in the traditional sense, to the *cis* COOH arrangement. However, as can be seen from results of high-level theoretical calculations performed on *cis*- and *trans*-formic acid (see Table 5), the *cis* form has a considerably lower total energy. The best estimate of the corresponding energy difference is 1550-(100) cm<sup>-1</sup>, where the error attached to the estimate accounts for the neglect of higher level geometry optimizations, the



**Figure 2.** Simple structural units in  $\alpha$ -alanine conformers.

incompleteness of the basis sets employed, and remaining errors in the treatment of electron correlation. The core–core and core–valence effects, which were shown<sup>66</sup> to be important in several extensive calculations attempting to produce quantitative energy and property estimates, are predicted<sup>67</sup> to have a negligible effect,  $0.7\text{ cm}^{-1}$ , on the energy difference between *cis*- and *trans*-formic acid, favoring the *trans* form. (Due to the large size of the basis sets employed, calculation of core–core and core–valence effects is rather expensive; thus, they have not been determined explicitly for  $\alpha$ -alanine but were assumed to be similarly negligible.) The results of Table II of ref 18 and Table 1 of this paper suggest that the following energy values can be assigned to the five different H-bonds of Figure 2: **1** =  $-1500\text{ cm}^{-1}$ , **2** =  $-3000\text{ cm}^{-1}$ , **3** =  $-1000\text{ cm}^{-1}$ , **4** =  $-1100\text{ cm}^{-1}$ , and **5** =  $-650\text{ cm}^{-1}$ . (These energy values are in accord, when appropriate, with the qualitative picture presented by Hu et al.<sup>15</sup> and with the linear regression analysis estimates of Gronert and O’Hair.<sup>34b</sup>) These data clearly assess the importance of the related bonding effects as follows: (a) the O–H $\cdots$ N H-bond, **2**, has the largest stabilization effect; (b) the stabilization contribution of **1**, the N–H $\cdots$ O H-bond, is only about half as large as that of **2**, even though it’s made up by two bifurcated H-bonds; (c) the N–H $\cdots$ OH bifurcated H-bond, **3**, seems to be even weaker. With the numbers given above the relative energies of all the glycine and  $\alpha$ -alanine conformers investigated can be determined. Appropriate sums of these simple estimates agree within about  $250\text{ cm}^{-1}$  with the relative energy values obtained from high-level theoretical computations.

Another important observation is that the energy order of the conformers of  $\alpha$ -alanine coincides with the order observed for the eight stable conformers of glycine.<sup>18</sup> This means, among other things, that the global minima on the glycine and  $\alpha$ -alanine potential energy surfaces have a *cis* COOH arrangement and two N–H $\cdots$ O H-bonds, while the next lowest energy conformer has a *trans* COOH arrangement and an O–H $\cdots$ N H-bond. The near equivalence of the total energies of conformers **I**, **IIA**, and **IIB** can now be understood as a result of replacing the strong O–H $\cdots$ N H-bond of **II** with considerably less stable N–H $\cdots$ O H-bonds, which, on the other hand, allow for the more stable *cis* COOH arrangement, making up for the difference. It is of interest to note at this point that at several correlated levels of theory  $\alpha$ -alanine **IIA** is calculated to be more stable than  $\alpha$ -alanine **I**, at the cc-pVTZ MP2 level by almost  $100\text{ cm}^{-1}$ . All available experimental data (the GED study of Iijima and Beagley<sup>11</sup> and the MMW spectrum analyzed by Godfrey et al.<sup>9</sup>) suggest that  $\alpha$ -alanine **I** is the true global minimum on the potential energy surface of  $\alpha$ -alanine. Due to remaining errors in the relative energy predictions at the MP2 level (neglect of

higher level geometry optimizations, incompleteness of the basis sets employed, remaining errors in the treatment of electron correlation, and neglect of core–core and core–valence correlation effects), the only certain statement about the calculated energy difference between  $\alpha$ -alanine **I** and **II** is that it must be small. Note, furthermore, that the difference in the zero-point vibrational energy (ZPVE) corrections (see Table 4) of  $\alpha$ -alanine **I** and **II**,  $110\text{ cm}^{-1}$  after correction for the systematic overestimation of the frequencies at the 6-311++G\*\* RHF level, tends to stabilize  $\alpha$ -alanine **I** over **II**. Thus, this correction puts the theoretical predictions more in line with the experimental observations.

A further similarity between the potential energy surfaces of glycine and  $\alpha$ -alanine is that both amino acids have five distinct conformers below a relative energy value of  $1000\text{ cm}^{-1}$ , while the remaining “high-energy” conformers corresponding to minima on the appropriate potential energy surfaces have relative energies between  $1500$  and  $2400\text{ cm}^{-1}$ . Note also that while ZPVE corrections increase the effective energy difference of conformers **IIA** and **IIB** with respect to **I**, the vibrational effects tend to stabilize conformers **VI**–**VIII**.

It can also be noticed that electron correlation stabilizes all conformers relative to **I**; that is, the RHF level of theory overestimates the relative energies of the  $\alpha$ -alanine conformers in all cases. This overestimation is, nevertheless, not completely systematic. Most importantly, the RHF level fails completely during prediction of the relative energy of **IIA** and **IIB**. While in general the difference between 6-311++G\*\* SCF and MP2 energy predictions is only about  $300\text{ cm}^{-1}$ , for **II** it is above  $800\text{ cm}^{-1}$ .

Recently, Barone et al.<sup>39</sup> investigated the conformational behavior of gaseous glycine by density functional (DFT) methods. On the basis of the comparison of results from different DFT approaches and the high-quality ab initio results of Császár,<sup>18</sup> they concluded that while standard density functional methods have difficulties in coping with the subtle energy differences in amino acids, hybrid functionals, such as B3LYP,<sup>55</sup> have much promise and approach the quality of MP2 calculations at a fraction of the cost. The present 6-311++G\*\* B3LYP optimizations support this view. Not only the calculated geometry parameters but also the energy differences appear to be of high quality. Thus, use of the B3LYP functional with a sufficiently large basis may open up the way to studies on related medium-sized biomolecules, e.g., oligopeptides and their analogues.

At this point it is worth emphasizing that while Godfrey et al.<sup>9</sup> could observe rotational lines corresponding to conformers **I** and **IIA**, they did not report any experimental measurements for **IIB**, a conformer which, according to the present calculations, should have about the same energy. This is even more surprising if one considers the fact that the dipole moment of **IIB** is as large as that of **IIA**,  $5.6\text{ D}$ . The reliability of all relevant theoretical data determined for **IIB** strongly suggests that it should be possible to observe this conformer by rotational spectroscopy. Failure of the observation of **IIB** should not suggest a lack of reliability in ab initio predictions but might be indicative of the existence of low rotational barriers, resulting in fast relaxation following the rotational and kinetic cooling during the jet expansion experiments. It is noted that in a recent jet expansion MMW study Godfrey and Brown<sup>8</sup> questioned the reliability of recent ab initio relative energy predictions for the conformers of free glycine. This author believes that this scepticism is unfounded and that the lack of observation of certain low-energy conformers in the vapor of glycine and/or

**TABLE 1: Relative and Total Energies of  $\alpha$ -Alanine Conformers<sup>a</sup>**

| method                | I              | IIA        | IIIB      | IIIA      | IIIB             | IVA       | IVB       | VA        | VB        | VI        | VII       | VIIIA     | VIIIB     |
|-----------------------|----------------|------------|-----------|-----------|------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 6-31G** RHF           | 0              | 905        | 827       | 524       | 564              | 639       | 662       | 722       | 900       | 2474      | 3039      | 3117      | 3227      |
|                       | (-321.888 180) | (884 056)  | (884 414) | (885 792) | (885 608)        | (885 267) | (885 162) | (884 890) | (884 080) | (876 910) | (874 335) | (873 979) | (873 475) |
| 6-311++G** RHF        | 0              | 848        | 867       | 421       | 515              | 542       | 581       | 646       | 765       | 2350      | 2815      | 2888      | 2987      |
|                       | (-321.970 926) | (967 063)  | (966 976) | (969 010) | (968 578)        | (968 456) | (968 278) | (967 984) | (967 443) | (960 218) | (958 100) | (957 766) | (957 312) |
| 6-311++G** DFT(B3LYP) | 0              | 10         | 5         | 373       | 457              | 457       | 509       | 693       | 787       | 2002      | c         | 2424      | 2428      |
|                       | (-323.856 078) | (856 033)  | (856 054) | (854 378) | c                | (853 994) | (853 760) | (852 920) | (852 492) | (846 956) | c         | (845 032) | (845 014) |
| 6-311++G** MP2(full)  | 0              | 51         | 166       | 335       | 449 <sup>b</sup> | 420       | 514       | 621       | 705       | 2025      | 2324      | 2425      | 2537      |
|                       | (-323.103 013) | (102 783)  | (102 256) | (101 489) | (100 969)        | (101 101) | (100 673) | (100 182) | (099 801) | (093 787) | (092 424) | (091 962) | (091 454) |
| 6-31G** RHF//         | 0              | 1010       | 993       | 519       | 692              | 610       | 627       | 683       | 886       | 2503      | 3078      | 3102      | 3225      |
|                       | (-321.885 019) | (880 418)  | (880 494) | (882 656) | (881 865)        | (882 240) | (882 162) | (881 909) | (880 980) | (873 615) | (870 995) | (870 883) | (870 323) |
| 6-311++G** MP2(full)  | 0              | 999        | 1034      | 443       | 628              | 533       | 569       | 633       | 771       | 2357      | 2845      | 2878      | 2984      |
|                       | (-321.967 238) | (962 687)  | (962 528) | (965 222) | (964 374)        | (964 809) | (964 645) | (964 353) | (963 724) | (956 500) | (954 277) | (954 126) | (953 641) |
| 6-311++G** MP2(full)  | 0              | 932        | 982       | 490       | 602              | 509       | 570       | 668       | 843       | 2138      | 2615      | 2638      | 2781      |
|                       | (-321.995 676) | (0991 431) | (991 201) | (993 444) | (992 931)        | (993 358) | (993 079) | (992 634) | (991 834) | (985 935) | (983 763) | (983 658) | (983 003) |
| cc-pVTZ RHF//         | 0              | 923        | 988       | 481       | 490              | 490       | 530       | 647       | 809       | 2074      | 2523      | 2545      | 2662      |
|                       | (-322.000 405) | (996 201)  | (995 902) | (998 215) | (998 172)        | (997 992) | (997 992) | (997 458) | (998 228) | (992 466) | (990 419) | (990 318) | (989 786) |
| aug-cc-pVTZ RHF//     | 0              | 915        | 977       | 486       | 599              | 488       | 524       | 642       | 797       | 2017      | 2127      | 2438      | 2450      |
|                       | (-322.001 915) | (997 748)  | (997 465) | (999 700) | (999 186)        | (999 691) | (999 529) | (998 991) | (998 228) | (992 466) | (990 419) | (990 318) | (989 786) |
| 7s6p3d2f RHF//        | 0              | 28         | -3        | 383       | 465              | 465       | 519       | 697       | 797       | 2017      | 2127      | 2438      | 2450      |
|                       | (-323.855 752) | (855 626)  | (855 767) | (854 007) | (853 633)        | (853 633) | (853 387) | (852 578) | (852 121) | (846 577) | (846 078) | (844 664) | (844 609) |
| 6-311++G** RHF        | 0              | 128        | 147       | 458       | 444              | 495       | 557       | 664       | 831       | 2240      | 2648      | 2726      | 2861      |
|                       | (-322.810 166) | (809 585)  | (809 496) | (808 078) | (808 141)        | (807 910) | (807 627) | (807 141) | (806 378) | (799 960) | (798 099) | (797 747) | (797 130) |
| 6-311++G** MP2//      | 0              | 163        | 276       | 358       | 506              | 419       | 505       | 612       | 712       | 2024      | 2353      | 2419      | 2533      |
|                       | (-322.983 731) | (982 988)  | (982 473) | (982 100) | (981 427)        | (981 824) | (981 429) | (980 941) | (980 487) | (974 511) | (973 011) | (972 711) | (972 190) |
| cc-pVTZ MP2//         | 0              | -96        | 7         | 379       | 437              | 402       | 502       | 642       | 783       | 1818      | 2025      | 2218      | 2285      |
|                       | (-323.175 162) | (175 601)  | (175 128) | (173 436) | (173 173)        | (173 320) | (172 874) | (172 237) | (171 594) | (166 877) | (165 934) | (165 058) | (164 749) |
| aug-cc-pVTZ MP2//     | 0              | -78        | 393       | 393       | 385              | 385       | 449       | 631       | 787       | 1792      | 1966      | 2174      | 2213      |
|                       | (-323.206 196) | (206 551)  | (204 403) | (204 403) | (204 444)        | (204 444) | (240 010) | (239 183) | (238 468) | (233 892) | (233 098) | (232 149) | (231 974) |
| 7s6p3d2f MP2//        | 0              | -41        | 93        | 402       | 428              | 391       | 449       | 631       | 787       | 1792      | 1966      | 2174      | 2213      |
|                       | (-323.242 056) | (242 243)  | (241 630) | (240 226) | (240 104)        | (240 275) | (240 010) | (239 183) | (238 468) | (233 892) | (233 098) | (232 149) | (231 974) |
| 6-31G** MP3//         | 0              | 296        | 267       | 489       | 496              | 509       | 540       | 617       | 839       | 2192      | 2694      | 2688      | 2780      |
|                       | (-322.840 613) | (839 263)  | (839 398) | (838 383) | (838 352)        | (838 293) | (838 153) | (837 800) | (836 789) | (830 625) | (828 338) | (828 365) | (827 946) |
| 6-31G** MP4(SD/TQ)//  | 0              | 203        | 190       | 467       | 433              | 494       | 549       | 642       | 819       | 2175      | 2652      | 2657      | 2782      |
|                       | (-322.885 392) | (884 469)  | (884 527) | (883 262) | (883 407)        | (883 142) | (882 879) | (882 466) | (881 659) | (875 484) | (873 310) | (873 284) | (872 718) |
| 6-31G** CCSD//        | 0              | 343        | 339       | 475       | 472              | 505       | 537       | 600       | 813       | 2239      | 2743      | 2732      | 2845      |
|                       | (-322.853 279) | (851 717)  | (851 736) | (851 114) | (851 128)        | (850 979) | (850 831) | (850 544) | (849 573) | (843 079) | (840 779) | (840 830) | (840 314) |
| 6-311++G** MP2(full)  | 0              | 223        | 210       | 467       | 438              | 493       | 546       | 619       | 810       | 2196      | 2675      | 2677      | 2794      |
|                       | (-322.882 183) | (881 165)  | (881 226) | (880 053) | (880 186)        | (879 935) | (879 697) | (879 361) | (878 491) | (872 178) | (869 995) | (869 986) | (869 452) |

<sup>a</sup> All total energies in hartrees, all relative energies in  $\text{cm}^{-1}$ . The 6-31G\*\*, 6-311++G\*\*, cc-pVTZ, 7s6p3d2f, and aug-cc-pVTZ basis sets consist of 125, 181, 278, 436, and 437 CGFs, respectively (for details, see text). For the notation of the conformers see Figure 1. // means "at the geometry of". <sup>b</sup> Note that the potential energy surface is extremely flat around the stationary point. <sup>c</sup> No stationary point was found.

TABLE 2: Selected Theoretical (6-311++G\*\* RHF and MP2) and Experimental (GED) Geometry Parameters for  $\alpha$ -Alanine<sup>a</sup>

| param   | I     |       | IIA    |        | IIB    |        | IIIA   |        | IIIB   |       | IVA   |       | IVB   |        | VA     |        | VB     |       | VI    |        | VII    |       | VIII  |       | VIII  |       |       |
|---|-------|-------|--------|--------|--------|--------|--------|--------|--------|-------|-------|-------|-------|--------|--------|--------|--------|-------|-------|--------|--------|-------|-------|-------|-------|-------|-------|
|   | RHF   | MP2   | RHF    | MP2    | RHF    | MP2    | RHF    | MP2    | RHF    | MP2   | RHF   | MP2   | RHF   | MP2    | RHF    | MP2    | RHF    | MP2   | RHF   | MP2    | RHF    | MP2   | RHF   | MP2   | RHF   | MP2   |       |
| $r(\text{C}_1\text{C}_2)$                           | 1.523 | 1.521 | 1.507  | 1.531  | 1.534  | 1.539  | 1.523  | 1.520  | 1.525  | 1.514 | 1.512 | 1.516 | 1.515 | 1.518  | 1.517  | 1.516  | 1.516  | 1.535 | 1.531 | 1.531  | 1.535  | 1.535 | 1.528 | 1.538 | 1.537 | 1.532 | 1.530 |
| $r(\text{C}_1\text{N}_3)$                           | 1.443 | 1.452 | 1.471  | 1.458  | 1.460  | 1.470  | 1.448  | 1.457  | 1.454  | 1.447 | 1.447 | 1.447 | 1.456 | 1.455  | 1.465  | 1.461  | 1.452  | 1.445 | 1.441 | 1.450  | 1.445  | 1.452 | 1.445 | 1.445 | 1.454 | 1.445 | 1.454 |
| $r(\text{C}_1\text{O}_4)$                           | 1.329 | 1.356 | 1.347  | 1.319  | 1.319  | 1.340  | 1.331  | 1.359  | 1.333  | 1.327 | 1.353 | 1.328 | 1.354 | 1.331  | 1.358  | 1.330  | 1.357  | 1.335 | 1.335 | 1.361  | 1.329  | 1.356 | 1.332 | 1.357 | 1.331 | 1.356 |       |
| $r(\text{C}_1=\text{O}_5)$                          | 1.184 | 1.211 | 1.192  | 1.180  | 1.180  | 1.208  | 1.184  | 1.211  | 1.183  | 1.183 | 1.210 | 1.183 | 1.210 | 1.182  | 1.208  | 1.183  | 1.209  | 1.177 | 1.205 | 1.177  | 1.205  | 1.179 | 1.205 | 1.177 | 1.204 | 1.177 | 1.205 |
| $r(\text{O}_4\text{H}_6)$                           | 0.946 | 0.968 | 0.977  | 0.949  | 0.980  | 0.949  | 0.946  | 0.968  | 0.946  | 0.946 | 0.968 | 0.946 | 0.968 | 0.946  | 0.968  | 0.946  | 0.968  | 0.942 | 0.964 | 0.942  | 0.966  | 0.942 | 0.966 | 0.942 | 0.963 | 0.941 | 0.963 |
| $r(\text{C}_2\text{C}_8)$                           | 1.532 | 1.530 | 1.545  | 1.530  | 1.527  | 1.524  | 1.530  | 1.528  | 1.521  | 1.538 | 1.537 | 1.529 | 1.526 | 1.521  | 1.518  | 1.535  | 1.535  | 1.533 | 1.531 | 1.531  | 1.530  | 1.528 | 1.538 | 1.537 | 1.532 | 1.530 |       |
| $r(\text{NH})$                                      | 1.000 | 1.016 | 1.014  | 1.000  | 1.014  | 1.000  | 1.015  | 1.000  | 1.015  | 1.000 | 1.016 | 0.999 | 1.014 | 0.999  | 1.015  | 1.000  | 1.016  | 1.000 | 1.016 | 1.000  | 1.016  | 0.997 | 1.012 | 1.000 | 1.016 | 0.998 | 1.014 |
| $r(\text{CH})$                                      | 1.084 | 1.092 | 1.084  | 1.093  | 1.085  | 1.093  | 1.085  | 1.093  | 1.084  | 1.086 | 1.094 | 1.084 | 1.092 | 1.084  | 1.092  | 1.086  | 1.094  | 1.085 | 1.093 | 1.084  | 1.093  | 1.084 | 1.093 | 1.086 | 1.094 | 1.085 | 1.094 |
| $\angle(\text{C}_2\text{N}_3)$                      | 113.3 | 113.7 | 110.1  | 110.3  | 109.4  | 110.1  | 109.4  | 114.6  | 115.2  | 112.1 | 108.5 | 107.9 | 108.8 | 108.8  | 108.4  | 109.3  | 113.3  | 113.7 | 116.8 | 117.5  | 116.8  | 117.5 | 108.4 | 107.7 | 108.8 | 108.0 |       |
| $\angle(\text{C}_2\text{C}_1\text{O}_4)$            | 112.2 | 111.4 | 110.3  | 115.5  | 114.0  | 115.2  | 114.0  | 112.9  | 111.5  | 112.3 | 111.7 | 112.6 | 111.8 | 111.7  | 110.7  | 112.5  | 111.1  | 115.7 | 115.7 | 117.4  | 116.5  | 117.4 | 116.5 | 116.1 | 116.0 | 116.0 |       |
| $\angle(\text{C}_2\text{C}_1=\text{O}_5)$           | 125.4 | 125.4 | 125.6  | 122.0  | 122.6  | 122.6  | 124.8  | 125.5  | 126.4  | 125.0 | 124.8 | 125.0 | 125.1 | 126.0  | 126.3  | 125.0  | 125.7  | 123.9 | 123.8 | 121.7  | 122.3  | 123.2 | 123.0 | 123.0 | 123.5 | 123.5 |       |
| $\angle(\text{O}_4\text{C}_1=\text{O}_5)$           | 122.4 | 123.1 | 122.4  | 123.4  | 122.2  | 123.4  | 122.2  | 122.8  | 122.1  | 122.6 | 123.4 | 122.3 | 123.0 | 122.3  | 123.0  | 122.5  | 123.1  | 120.4 | 120.5 | 120.9  | 121.2  | 120.7 | 120.8 | 120.8 | 120.3 | 120.4 |       |
| $\angle(\text{C}_1\text{C}_2\text{C}_8)$            | 109.6 | 108.3 | 111.6  | 108.7  | 108.1  | 111.3  | 110.4  | 109.2  | 108.6  | 111.4 | 108.5 | 107.8 | 112.7 | 112.4  | 111.4  | 109.5  | 108.8  | 109.6 | 108.3 | 108.8  | 107.9  | 108.4 | 107.8 | 108.4 | 107.8 | 113.3 |       |
| $\angle(\text{C}_1\text{O}_4\text{H}_6)$            | 108.9 | 106.2 | 112.3  | 108.8  | 104.0  | 108.7  | 103.8  | 108.9  | 106.0  | 108.9 | 108.6 | 105.8 | 108.9 | 106.1  | 108.9  | 108.9  | 106.2  | 112.3 | 109.0 | 111.0  | 107.1  | 112.2 | 108.7 | 112.6 | 109.1 |       |       |
| $\angle(\text{C}_2\text{N}_3\text{H})$              | 110.5 | 109.2 | 113.3  | 112.1  | 111.6  | 111.7  | 111.8  | 110.8  | 111.7  | 111.2 | 109.4 | 111.5 | 110.4 | 111.1  | 110.0  | 111.2  | 109.7  | 107.8 | 108.8 | 116.4  | 115.4  | 111.2 | 108.9 | 111.6 | 110.2 |       |       |
| $\angle(\text{N}_3\text{C}_2\text{C}_1\text{O}_4)$  | 166.0 | 161.3 | 162.8  | 20.4   | 18.2   | 19.4   | 11.8   | 45.4   | 44.5   | 60.2  | 152.6 | 150.7 | 168.2 | 165.7  | -50.4  | -48.9  | -47.6  | 163.8 | 158.8 | 17.4   | 19.8   | 149.3 | 147.2 | 169.2 | 163.7 |       |       |
| $\tau(\text{N}_3\text{C}_2\text{C}_1\text{O}_4)$    | -14.5 | -20.5 | -161.4 | -164.4 | -163.0 | -169.7 | -135.9 | -137.3 | -119.3 | -30.1 | -32.5 | -14.7 | -16.9 | 131.8  | 131.1  | 133.4  | 135.2  | -17.4 | -23.9 | -163.7 | -162.5 | -33.9 | -37.1 | -13.8 | -19.5 |       |       |
| $\tau(\text{C}_8\text{C}_2\text{C}_1\text{O}_4)$    | -70.3 | -76.6 | -101.0 | -101.7 | 148.0  | 138.6  | -78.7  | -78.7  | -175.2 | -81.8 | -84.4 | 45.4  | 44.0  | -172.5 | -171.6 | 78.1   | 79.0   | -73.0 | -79.5 | -107.7 | -104.3 | -85.3 | -88.3 | 47.0  | 42.5  |       |       |
| $\tau(\text{C}_2\text{C}_1\text{O}_4\text{H}_6)$    | 178.7 | 176.9 | 180.0  | -4.3   | -5.1   | -4.6   | 178.3  | 177.4  | -179.1 | 177.0 | 175.6 | 177.1 | 176.9 | -179.8 | -179.9 | -179.6 | -178.7 | -2.8  | -6.2  | -3.3   | -5.1   | -7.8  | -10.1 | -4.2  | -6.5  |       |       |
| $\tau(\text{H}_6\text{N}_3\text{C}_2\text{C}_1)$    | 58.4  | 57.8  | 86.8   | 91.5   | 90.0   | 99.6   | 61.9   | 58.1   | 65.2   | 48.2  | 47.9  | 39.4  | 42.3  | 60.6   | 61.1   | 66.9   | 64.4   | 57.6  | 64.4  | 58.5   | 54.8   | 43.7  | 43.7  | 33.3  | 36.2  |       |       |
| $\tau(\text{H}_{10}\text{N}_3\text{C}_2\text{C}_1)$ | -60.2 | -58.4 | -151.8 | -148.7 | -149.3 | -140.9 | -59.4  | -59.7  | -56.5  | 169.8 | 166.3 | 161.7 | 162.0 | -178.7 | -179.5 | -172.2 | -177.8 | -60.5 | -57.7 | -78.0  | -77.8  | 165.4 | 162.1 | 156.1 | 156.1 |       |       |

<sup>a</sup> Distances ( $r$ ) in angstroms; angles ( $\angle$  and  $\tau$ ) in degrees. <sup>b</sup> Reference 11. For uncertainties of the parameters determined see the original publication. Parameters printed in italics were assumed and not refined.

$\alpha$ -alanine should be attributed to the existence of low-barrier relaxation pathways if sensitivity problems can be ruled out.

Partly on the basis of the measured dipole moment components for conformers **I** and **IIA** and the observed integrated areas of the strongest lines in the MMW spectra of the conformers, Godfrey et al.<sup>9</sup> report the following experimentally determined approximate lower limits for conformational energies relative to those of  $\alpha$ -alanine **I**: **IIA** =  $\sim 800$  cm<sup>-1</sup>, **IIIA** =  $\sim 700$  cm<sup>-1</sup>, and **VI** =  $\sim 1200$  cm<sup>-1</sup>. Thus, Godfrey et al. contributed the relatively easy detection of  $\alpha$ -alanine **IIA** to its large dipole moment (see, however, comments in the preceding paragraph). All the experimental estimates of Godfrey et al.<sup>9</sup> are significantly different from the high-quality correlated-level ab initio relative energy predictions of this report: the present calculations predict the energies of  $\alpha$ -alanine **I** and **II** to be about the same, while for conformers **III** the theoretical predictions are considerably lower and for **VI** considerably higher than their experimental counterparts. Although it is rather uncertain what error limits should be attached to the theoretical values, a somewhat conservative estimate would be  $\pm 200$  cm<sup>-1</sup>, suggesting that the differences between theory and experiment deserve further study.

It is also clear from the data of Table 1 that once a sufficiently flexible basis set (such as 6-311++G\*\*, applied throughout this study and in ref 18 as one of the smallest basis sets) is used for optimization of the geometry of the different conformers of glycine and  $\alpha$ -alanine, the choice of reference geometry selected for the subsequent single-point energy calculations is rather irrelevant. For example, Table 1 shows that, with the 6-311++G\*\* basis set used throughout, for conformers **IIIA**, **IVA**, **IVB**, **VI**, **VII**, and **VIIIA** of  $\alpha$ -alanine the differences between the relative energies at the RHF//RHF and RHF//MP2(full) levels (where // means "at the geometry of") are a mere +22, -9, -12, +7, +30, and -11 cm<sup>-1</sup>, respectively. Furthermore, if one compares the MP2//MP2 and MP2//RHF results, it turns out that the differences are comparably small, spreading from -8 (**IVB**) to +29 cm<sup>-1</sup> (**VII**). All these differences are small and certainly much smaller than the expected error limits of these calculations. Thus, they can be regarded as insignificant. The RHF//RHF and RHF//MP2, and the MP2//MP2 and MP2//RHF, energy differences between conformers **I** and **II** of  $\alpha$ -alanine are always around 130 cm<sup>-1</sup>. These differences, though still rather small, appear to be less tolerable. Thus, one can conclude that single point MP2 energies obtained at HF geometries for amino acids, at least for glycine and  $\alpha$ -alanine, are rather accurate, as conventional wisdom would imply. As detailed in ref 16, these results help to clarify some recent statements of Frey and co-workers<sup>27</sup> about the imprecision of RHF geometry optimizations of amino acids and their claim that single-point MP2 energies obtained at RHF-optimized geometries are "potentially highly inaccurate". The extensive relative energy results of Gronert and O'Hair<sup>34b</sup> obtained for a number of alanine, serine, and cysteine conformers also support the view that the warnings of Frey et al.<sup>27</sup> were somewhat unfounded. It is now obvious that although RHF theory completely fails in predicting the energy difference between N-H...O and O-H...N H-bonded conformers, for the other, less important, conformers the predicted relative energies seem to be reasonably accurate.

In summary, all the new and previous<sup>18</sup> theoretical results seem to indicate that there is a great similarity between the conformational behavior of glycine (an amino acid with R = H) and  $\alpha$ -alanine (R = CH<sub>3</sub>).

**Geometries.** As expected and revealed in Table 2, changes in bond lengths among the conformers of  $\alpha$ -alanine are rather

**TABLE 3: Theoretical (RHF, DFT(B3LYP), and MP2) and Experimental Rotational Constants, Quartic Centrifugal Distortion Constants, and Dipole Moments of  $\alpha$ -Alanine<sup>a</sup>**

| param          | <b>I</b> |        |        |                   | <b>IIA</b> |        |        |                   | <b>IIIB</b> |        |        | <b>IIIA</b> |        | <b>IIIB</b> |                  | <b>IVA</b> |        |
|----------------|----------|--------|--------|-------------------|------------|--------|--------|-------------------|-------------|--------|--------|-------------|--------|-------------|------------------|------------|--------|
|                | RHF      | DFT    | MP2    | expt <sup>b</sup> | RHF        | DFT    | MP2    | expt <sup>b</sup> | RHF         | DFT    | MP2    | RHF         | MP2    | RHF         | MP2 <sup>c</sup> | RHF        | MP2    |
| A/MHz          | 5160.0   | 5055.4 | 5084.8 | 5066.1            | 5079.4     | 4951.6 | 5003.8 | 4973.1            | 4983.9      | 4855.6 | 4872.1 | 5144.5      | 5078.4 | 5163.7      | 5125.0           | 5165.7     | 5063.6 |
| B/MHz          | 3136.1   | 3037.8 | 3053.9 | 3100.9            | 3182.5     | 3218.0 | 3202.7 | 3228.3            | 3456.6      | 3408.7 | 3447.6 | 2859.2      | 2839.4 | 3219.5      | 2889.0           | 3061.2     | 3028.5 |
| C/MHz          | 2283.9   | 2262.8 | 2305.6 | 2264.0            | 2370.5     | 2284.5 | 2351.0 | 2307.8            | 2172.0      | 2149.9 | 2175.0 | 2525.7      | 2503.8 | 2226.0      | 2400.0           | 2374.4     | 2381.3 |
| $D_J$ /kHz     | 1.552    |        |        | 1.567             | 1.253      |        |        | 1.411             | 0.469       |        |        | 2.771       |        | 2.428       |                  | 1.620      |        |
|                | (1.261)  |        |        |                   | (1.021)    |        |        |                   |             |        |        |             |        |             |                  |            |        |
| $D_{JK}$ /kHz  | -0.931   |        |        | -1.064            | -0.157     |        |        | -0.335            | 2.706       |        |        | -1.588      |        | -1.666      |                  | -0.986     |        |
|                | (-0.725) |        |        |                   | (-0.109)   |        |        |                   |             |        |        |             |        |             |                  |            |        |
| $D_K$ /kHz     | 1.038    |        |        | 0.87              | 0.561      |        |        | 0.84              | -1.861      |        |        | 0.529       |        | 1.056       |                  | 1.115      |        |
|                | (0.844)  |        |        |                   | (0.469)    |        |        |                   |             |        |        |             |        |             |                  |            |        |
| $d_1$ /kHz     | -0.584   |        |        | -0.571            | -0.335     |        |        | -0.408            | -0.250      |        |        | -0.483      |        | -1.022      |                  | -0.606     |        |
|                | (-0.471) |        |        |                   | (-0.270)   |        |        |                   |             |        |        |             |        |             |                  |            |        |
| $d_2$ /kHz     | -0.446   |        |        | -0.446            | -0.335     |        |        | -0.395            | -0.119      |        |        | -0.965      |        | -0.776      |                  | -0.448     |        |
|                | (-0.358) |        |        |                   | (-0.268)   |        |        |                   |             |        |        |             |        |             |                  |            |        |
| $\mu_a$ /D     | 0.6      |        |        | 0.622             | 5.4        |        |        | 4.924             | 5.0         |        |        | 0.2         |        | 0.7         |                  | 0.9        |        |
| $\mu_b$ /D     | 1.2      |        |        | 1.60              | 1.4        |        |        | 1.4               | 2.4         |        |        | 1.2         |        | 1.3         |                  | 2.2        |        |
| $\mu_c$ /D     | 0.5      |        |        | 0.339             | 0.5        |        |        | 0.279             | 0.4         |        |        | 1.3         |        | 0.2         |                  | 0.1        |        |
| $\mu_{tot}$ /D | 1.4      |        |        | 1.8               | 5.6        |        |        | 5.13              | 5.6         |        |        | 1.8         |        | 1.4         |                  | 2.4        |        |

| param          | <b>IVB</b> |        | <b>VA</b> |        | <b>VB</b> |        | <b>VI</b> |        | <b>VII</b> |        | <b>VIIIA</b> |        | <b>VIIIB</b> |        |
|----------------|------------|--------|-----------|--------|-----------|--------|-----------|--------|------------|--------|--------------|--------|--------------|--------|
|                | RHF        | MP2    | RHF       | MP2    | RHF       | MP2    | RHF       | MP2    | RHF        | MP2    | RHF          | MP2    | RHF          | MP2    |
| A/MHz          | 5145.1     | 5032.1 | 5188.8    | 5073.0 | 5123.5    | 5035.7 | 5102.3    | 5046.8 | 5048.6     | 4988.0 | 5112.3       | 5029.2 | 5078.6       | 4974.9 |
| B/MHz          | 3368.7     | 3377.7 | 3342.0    | 3341.3 | 2889.5    | 2892.5 | 3123.9    | 3020.9 | 3151.7     | 3101.7 | 3033.6       | 2982.2 | 3379.4       | 3391.3 |
| C/MHz          | 2177.8     | 2159.6 | 2188.8    | 2173.5 | 2541.9    | 2520.2 | 2301.9    | 2324.7 | 2299.3     | 2296.9 | 2406.1       | 2413.7 | 2175.6       | 2146.6 |
| $D_J$ /kHz     | 0.676      |        | 0.774     |        | 1.738     |        | 1.392     |        | 1.842      |        | 2.068        |        | 0.679        |        |
| $D_{JK}$ /kHz  | 1.423      |        | 0.926     |        | -0.029    |        | -0.643    |        | -0.708     |        | -1.618       |        | 2.504        |        |
| $D_K$ /kHz     | -0.598     |        | -0.166    |        | 0.012     |        | 0.803     |        | 0.420      |        | 1.263        |        | -1.790       |        |
| $d_1$ /kHz     | -0.317     |        | -0.338    |        | -0.327    |        | -0.492    |        | -0.608     |        | -0.767       |        | -0.324       |        |
| $d_2$ /kHz     | -0.142     |        | -0.163    |        | -0.520    |        | -0.386    |        | -0.591     |        | -0.618       |        | -0.198       |        |
| $\mu_a$ /D     | 0.6        |        | 0.7       |        | 2.2       |        | 1.7       |        | 3.6        |        | 3.1          |        | 3.3          |        |
| $\mu_b$ /D     | 1.5        |        | 1.9       |        | 0.2       |        | 2.7       |        | 2.6        |        | 3.5          |        | 2.9          |        |
| $\mu_c$ /D     | 1.7        |        | 1.7       |        | 1.9       |        | 1.7       |        | 0.7        |        | 1.3          |        | 2.3          |        |
| $\mu_{tot}$ /D | 2.4        |        | 2.7       |        | 2.9       |        | 3.6       |        | 4.5        |        | 4.9          |        | 4.9          |        |

<sup>a</sup> All theoretical values were obtained with the 6-311++G\*\* basis set. The quartic centrifugal distortion (QCD) constants correspond to Watson's  $S$ -reduced Hamiltonian. The theoretical QCD constants reported in parentheses were obtained from unscaled quadratic force fields. All other QCD constants were obtained from crude SQM<sup>71,72</sup> force fields with scale factors of 0.9 for all stretches and 0.8 for all bends. <sup>b</sup> Reference 9. For uncertainties of the rotational and QCD constants and dipole moment components see the original publication. <sup>c</sup> See footnote *b* to Table 1.

small. Still, the spread of bond distances for all characteristic single bonds (C–N, C–O, C–C) is larger than 0.02 Å. The most significant differences can be observed between the C–N and C–O distances of  $\alpha$ -alanine **I** and **II**. There are some notable changes among the bond angles; probably most important is the dependence of the CCN and CCO angles on the orientation of the NH<sub>2</sub> and OH groups, respectively. The CCN angle changes between 107.7° (**VIIIA**) and 117.5° (**VII**), while the CCO angle changes between 111.4° (**I**) and 116.5° (**VII**). While part of this spread in the bond angles should be attributed to sizable through-space repulsions, the *trans*-angle rule proposed by Räsänen et al.,<sup>68</sup> stating that “if in a conformer of a primary alcohol or amine a CC or CH bond is *trans* to an XH bond (X = O,N), the corresponding XCC or XCH angle will be considerably smaller than that for other configurations”, is clearly valid for the conformers of  $\alpha$ -alanine. For example, the CCO angle of **I** is 111.4°, while that of **IIA** is 114.0°, in agreement with the “*trans*” vs “*cis*” arrangement of the CC bond as compared to the OH bond. The tilt and asymmetry of the methyl group are similar to those described by Boggs et al.<sup>69</sup> in asymmetric environments.

Comparison of Table 2 of this study and Table III of ref 18 suggests that there are only a few characteristic differences between the structures of the related conformers of  $\alpha$ -alanine and glycine. 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 experimental methods for molecules of this size and of low symmetry. Some angles change more noticeably; especially pronounced is a decrease of 1.9° in the CCN angle of  $\alpha$ -alanine **I** as compared

to glycine **Ip**.<sup>18</sup> Note that the CCN angles of the conformers of  $\alpha$ -alanine seem to be smaller by about 2–3° than the CCN angles of the respective glycine conformers.

In contrast to all other conformers, **IIIB** might not correspond to a minimum on the true potential energy surfaces (PES) of  $\alpha$ -alanine. The PES is extremely flat around the **IIIB** minimum, and position of the minimum greatly depends on the level of theory at which it is located. Cao et al.<sup>25</sup> did not recognize this and speculated that “at the MP2 computational level [this] conformer is in good agreement with [a] set of [experimental] rotational constants and dipole moments”. However, as Table 3 shows, the fortuitous agreement between experiment and the 6-311G\*\* MP2 rotational constants disappears if diffuse functions are added to the basis. Thus, it appears that there is no ambiguity in assigning the measured rotational constants to conformer **IIA**.

The only experimental attempt so far to determine geometries for  $\alpha$ -alanine conformers is the GED study of Iijima and Beagley.<sup>11,70</sup> During determination of the structure of the lowest energy conformer of  $\alpha$ -alanine Iijima and Beagley concluded, on the basis of model refinements assuming internal rotation only around the C–C bond, that “the vapour of  $\alpha$ -alanine consists of one conformer with a high potential barrier around the C–C bond”, and consequently, they carried out the structural refinements assuming the presence of only one conformer in the gas, that of conformer **I**. The energy results of Table 1 suggest a substantially different picture: the vapor of  $\alpha$ -alanine should contain, especially at 500 K, the temperature of the GED experiment, several conformers in noticeable concentration. As the geometry results obtained for  $\alpha$ -alanine indicate (see Table

**TABLE 4: 6-311++G\*\* RHF Vibrational Frequencies, Infrared Intensities, and Zero-Point Vibrational Energies (ZPVEs) of  $\alpha$ -Alanine Conformers<sup>a</sup>**

| $\nu_i$       | <b>I</b> |     | <b>IIA</b> |     | <b>IIIB</b> |     | <b>IIIA</b> |     | <b>IIIB</b> |     | <b>IVA</b> |     | <b>IVB</b> |     | <b>VA</b> |     | <b>VB</b> |     | <b>VI</b> |     | <b>VII</b> |     | <b>VIIIA</b> |     | <b>VIIIB</b> |     |
|---------------|----------|-----|------------|-----|-------------|-----|-------------|-----|-------------|-----|------------|-----|------------|-----|-----------|-----|-----------|-----|-----------|-----|------------|-----|--------------|-----|--------------|-----|
|               | freq     | int | freq       | int | freq        | int | freq        | int | freq        | int | freq       | int | freq       | int | freq      | int | freq      | int | freq      | int | freq       | int | freq         | int | freq         | int |
| 1             | 4115     | 124 | 4034       | 208 | 4033        | 221 | 4117        | 130 | 4117        | 127 | 4114       | 128 | 4118       | 11  | 4119      | 130 | 4119      | 132 | 4175      | 87  | 4170       | 107 | 4180         | 94  | 4193         | 92  |
| 2             | 3810     | 8   | 3825       | 13  | 3822        | 14  | 3821        | 9   | 3822        | 44  | 3827       | 12  | 3839       | 14  | 3829      | 10  | 3822      | 8   | 3809      | 10  | 3871       | 15  | 3827         | 15  | 3842         | 20  |
| 3             | 3736     | 4   | 3745       | 7   | 3735        | 6   | 3738        | 4   | 3734        | 3   | 3741       | 5   | 3752       | 6   | 3747      | 3   | 3738      | 3   | 3735      | 6   | 3764       | 7   | 3741         | 8   | 3754         | 9   |
| 4             | 3266     | 21  | 3268       | 15  | 3275        | 13  | 3259        | 27  | 3261        | 25  | 3241       | 46  | 3266       | 26  | 3269      | 24  | 3253      | 38  | 3262      | 21  | 3275       | 15  | 3243         | 31  | 3255         | 23  |
| 5             | 3242     | 39  | 3231       | 42  | 3227        | 44  | 3249        | 37  | 3248        | 39  | 3234       | 35  | 3239       | 35  | 3242      | 28  | 3231      | 30  | 3232      | 43  | 3243       | 36  | 3224         | 39  | 3225         | 38  |
| 6             | 3207     | 13  | 3201       | 14  | 3186        | 14  | 3230        | 6   | 3195        | 15  | 3214       | 6   | 3177       | 25  | 3176      | 28  | 3228      | 12  | 3172      | 20  | 3223       | 7   | 3164         | 29  | 3161         | 37  |
| 7             | 3170     | 21  | 3168       | 23  | 3167        | 22  | 3172        | 19  | 3181        | 18  | 3164       | 22  | 3099       | 47  | 3104      | 51  | 3163      | 21  | 3165      | 29  | 3163       | 25  | 3158         | 35  | 3088         | 49  |
| 8             | 1993     | 431 | 2019       | 440 | 2016        | 422 | 1994        | 440 | 1997        | 403 | 2000       | 412 | 1996       | 436 | 2005      | 415 | 2005      | 459 | 2026      | 377 | 2018       | 386 | 2031         | 358 | 2026         | 394 |
| 9             | 1811     | 28  | 1800       | 47  | 1799        | 42  | 1805        | 47  | 1800        | 42  | 1784       | 62  | 1783       | 63  | 1794      | 45  | 1791      | 41  | 1813      | 32  | 1787       | 47  | 1783         | 61  | 1782         | 58  |
| 10            | 1617     | 10  | 1618       | 12  | 1617        | 15  | 1616        | 3   | 1610        | 4   | 1625       | 20  | 1621       | 17  | 1617      | 9   | 1619      | 14  | 1618      | 13  | 1615       | 4   | 1621         | 21  | 1623         | 9   |
| 11            | 1613     | 6   | 1615       | 6   | 1608        | 3   | 1610        | 8   | 1607        | 5   | 1611       | 8   | 1612       | 5   | 1608      | 7   | 1609      | 4   | 1606      | 6   | 1614       | 9   | 1610         | 6   | 1612         | 24  |
| 12            | 1552     | 13  | 1536       | 25  | 1541        | 154 | 1540        | 10  | 1554        | 2   | 1572       | 29  | 1553       | 19  | 1571      | 15  | 1546      | 34  | 1554      | 9   | 1535       | 12  | 1565         | 30  | 1546         | 10  |
| 13            | 1525     | 4   | 1527       | 396 | 1536        | 151 | 1520        | 13  | 1522        | 51  | 1528       | 4   | 1540       | 7   | 1537      | 39  | 1529      | 5   | 1531      | 2   | 1502       | 24  | 1532         | 5   | 1541         | 41  |
| 14            | 1492     | 31  | 1487       | 41  | 1509        | 158 | 1475        | 68  | 1511        | 4   | 1479       | 23  | 1476       | 7   | 1479      | 20  | 1480      | 64  | 1481      | 150 | 1471       | 372 | 1471         | 109 | 1479         | 17  |
| 15            | 1411     | 6   | 1469       | 63  | 1440        | 67  | 1455        | 16  | 1419        | 13  | 1427       | 16  | 1439       | 37  | 1429      | 25  | 1460      | 21  | 1408      | 210 | 1444       | 61  | 1424         | 322 | 1424         | 415 |
| 16            | 1385     | 2   | 1375       | 2   | 1334        | 13  | 1374        | 15  | 1364        | 11  | 1354       | 63  | 1339       | 48  | 1357      | 5   | 1362      | 75  | 1381      | 95  | 1351       | 21  | 1339         | 31  | 1329         | 4   |
| 17            | 1290     | 198 | 1315       | 3   | 1318        | 7   | 1324        | 178 | 1269        | 261 | 1269       | 286 | 1311       | 148 | 1294      | 298 | 1307      | 111 | 1267      | 6   | 1289       | 3   | 1267         | 37  | 1282         | 10  |
| 18            | 1249     | 151 | 1233       | 17  | 1249        | 16  | 1237        | 18  | 1242        | 73  | 1257       | 13  | 1248       | 162 | 1248      | 60  | 1204      | 76  | 1235      | 44  | 1245       | 26  | 1239         | 18  | 1241         | 40  |
| 19            | 1173     | 33  | 1149       | 42  | 1177        | 62  | 1140        | 91  | 1166        | 30  | 1187       | 15  | 1199       | 5   | 1186      | 14  | 1179      | 33  | 1169      | 42  | 1140       | 34  | 1186         | 14  | 1199         | 5   |
| 20            | 1100     | 1   | 1093       | 1   | 1097        | 1   | 1099        | 3   | 1099        | 8   | 1115       | 11  | 1105       | 1   | 1102      | 1   | 1105      | 26  | 1101      | 2   | 1075       | 1   | 1114         | 8   | 1104         | 1   |
| 21            | 1004     | 34  | 1008       | 47  | 1012        | 36  | 997         | 22  | 1004        | 9   | 990        | 2   | 1010       | 1   | 1009      | 4   | 992       | 10  | 1004      | 82  | 997        | 10  | 988          | 5   | 1009         | 4   |
| 22            | 976      | 125 | 947        | 93  | 949         | 112 | 943         | 150 | 947         | 139 | 917        | 173 | 913        | 107 | 942       | 65  | 918       | 153 | 981       | 83  | 874        | 17  | 922          | 179 | 911          | 117 |
| 23            | 866      | 38  | 875        | 5   | 881         | 5   | 857         | 30  | 882         | 53  | 864        | 42  | 872        | 71  | 872       | 97  | 861       | 53  | 873       | 10  | 814        | 10  | 872          | 18  | 876          | 53  |
| 24            | 819      | 41  | 823        | 5   | 811         | 9   | 850         | 47  | 817         | 37  | 835        | 38  | 801        | 63  | 811       | 72  | 851       | 28  | 812       | 4   | 700        | 81  | 825          | 11  | 801          | 16  |
| 25            | 691      | 16  | 782        | 141 | 790         | 142 | 665         | 44  | 657         | 75  | 693        | 36  | 658        | 74  | 646       | 63  | 659       | 56  | 701       | 10  | 596        | 30  | 708          | 15  | 637          | 14  |
| 26            | 618      | 127 | 685        | 12  | 600         | 12  | 623         | 134 | 609         | 84  | 628        | 106 | 601        | 66  | 600       | 58  | 611       | 105 | 538       | 4   | 570        | 41  | 545          | 1   | 532          | 5   |
| 27            | 535      | 20  | 580        | 7   | 572         | 2   | 575         | 13  | 545         | 34  | 540        | 26  | 533        | 9   | 534       | 26  | 572       | 13  | 474       | 116 | 530        | 165 | 480          | 123 | 476          | 83  |
| 28            | 409      | 14  | 415        | 2   | 440         | 7   | 397         | 13  | 404         | 3   | 397        | 3   | 445        | 14  | 428       | 7   | 388       | 5   | 411       | 17  | 414        | 39  | 397          | 6   | 450          | 36  |
| 29            | 330      | 8   | 352        | 10  | 382         | 10  | 312         | 25  | 339         | 13  | 338        | 3   | 359        | 1   | 346       | 4   | 326       | 19  | 331       | 18  | 368        | 57  | 339          | 9   | 365          | 12  |
| 30            | 265      | 3   | 300        | 29  | 301         | 18  | 290         | 22  | 310         | 38  | 267        | 10  | 292        | 2   | 306       | 46  | 268       | 28  | 265       | 14  | 331        | 15  | 266          | 24  | 313          | 9   |
| 31            | 246      | 47  | 263        | 3   | 293         | 13  | 256         | 1   | 264         | 2   | 256        | 0   | 251        | 20  | 270       | 3   | 258       | 2   | 243       | 55  | 278        | 21  | 258          | 3   | 249          | 20  |
| 32            | 242      | 1   | 249        | 3   | 250         | 1   | 249         | 1   | 252         | 1   | 236        | 42  | 241        | 26  | 243       | 5   | 248       | 6   | 234       | 2   | 248        | 2   | 247          | 40  | 241          | 29  |
| 33            | 55       | 3   | 71         | 2   | 62          | 1   | 45          | 1   | 34          | 1   | 63         | 2   | 61         | 1   | 59        | 1   | 62        | 1   | 62        | 9   | 50         | 5   | 56           | 5   | 52           | 6   |
| $\Delta$ ZPVE | 0        |     | +122       |     | +108        |     | +10         |     | -4          |     | -25        |     | -32        |     | -18       |     | -24       |     | -72       |     | -124       |     | -97          |     | -98          |     |

<sup>a</sup> Frequencies ( $\nu_i$ ; freq) in  $\text{cm}^{-1}$ ; intensities (int) in  $\text{km/mol}$ ; relative zero-point vibrational energies ( $\Delta$ ZPVE), measured with respect to the ZPVE of **I**,  $25407 \text{ cm}^{-1}$ , in  $\text{cm}^{-1}$ . All theoretical values were obtained at the 6-311++G\*\* RHF level at the respective fully optimized reference geometries. The scale factor of 0.89 might be a reasonable estimate for frequency (and consequently ZPVE) corrections.

**TABLE 5: Ab Initio Data ( $\text{cm}^{-1}$ ) for the Energy Difference between *cis*- and *trans*-Formic Acid<sup>a</sup>**

|                           | RHF  | $\Delta$ MP2 | $\Delta$ MP3 | $\Delta$ MP4 | $\Delta$ MP $\infty$ | MP $\infty$ | $\Delta$ CCSD | $\Delta$ CCSD(T) | CCSD(T) |
|---------------------------|------|--------------|--------------|--------------|----------------------|-------------|---------------|------------------|---------|
| 6-31G** (55)              | 2110 | -159         | -59          | +22          | +4                   | 1918        | -28           | -6               | 1917    |
| 6-311++G** (80)           | 1885 | -262         | -43          | -7           | -3                   | 1570        | -15           | -21              | 1587    |
| cc-pVTZ (118)             | 1741 | -174         | -43          | +15          | 0                    | 1539        | -11           | -12              | 1544    |
| aug-cc-pVTZ (184)         | 1658 | -164         | -23          | -2           | -2                   | 1467        | +1            | -12              | 1483    |
| cc-pVQZ (225)             | 1668 |              |              |              |                      |             |               |                  |         |
| CV(TZ) (262) <sup>b</sup> | 1725 | -161         |              |              |                      |             |               |                  |         |

<sup>a</sup> See the section Computational Details for basis set and theoretical level notations. For each basis set the total number of contracted Gaussian functions is listed in parentheses. Contributions to the energy difference relative to the next lower level of theory are denoted by the symbol  $\Delta$ . The  $\Delta$ CCSD values are relative to those obtained at the MP2 level. All energy data reported were obtained at 6-311++G\*\* MP2(full) optimized geometries, which are as follows for the *cis*(*trans*) conformers:  $r(\text{C}=\text{O}) = 1.2046(1.1978) \text{ \AA}$ ;  $r(\text{C}-\text{O}) = 1.3468(1.3533) \text{ \AA}$ ;  $r(\text{O}-\text{H}) = 0.9685(0.9635) \text{ \AA}$ ;  $r(\text{C}-\text{H}) = 1.0957(1.1033) \text{ \AA}$ ;  $\angle(\text{H}-\text{C}=\text{O}) = 125.26(123.98)^\circ$ ;  $\angle(\text{O}-\text{C}=\text{O}) = 125.24(122.68)^\circ$ ;  $\angle(\text{C}-\text{O}-\text{H}) = 106.34(108.38)^\circ$ . <sup>b</sup> See refs 66c and 67 for a description of the basis set.

2), the unfounded assumption about the presence of only one conformer in the gas during the course of the GED experiments resulted in considerable differences between Iijima and Beagley's structural results and the present high-quality ab initio results including their conclusion that "the methyl group strongly hinders the rotation of the acid group". As one should not place too much emphasis on absolute values of calculated geometry parameters, it is probably more telling that experimental changes in the GED geometry of  $\alpha$ -alanine **I**<sup>11</sup> as compared to glycine **I**<sup>10</sup> are hardly ever reproduced by the 6-311++G\*\* MP2(full) geometries of this study and of ref 18. It is therefore suggested that a new joint structural refinement of GED and now available microwave data, predicated toward correlated-level *ab initio* structural parameters, quadratic force fields, and relative ener-

gies, be performed for  $\alpha$ -alanine, just as it was suggested for glycine.<sup>16,18</sup>

The accuracy of the theoretical geometry parameters determined for conformers **I** and **IIA** of  $\alpha$ -alanine in this study is supported by the good agreement between the corresponding calculated and experimental<sup>9</sup> rotational constants (see Table 3), albeit the agreement between theory and experiment is not as impressive as observed for glycine<sup>18</sup> (correction of the experimental rotational constants for effects of vibrations might improve this agreement as it did for glycine **I**). The improved quality of the 6-31++G\*\* MP2(full) rotational constants over the 6-31G\*\* RHF constants determined by Godfrey et al.<sup>9</sup> can be seen from the fact that deviation of the 6-311++G\*\* MP2(full) values from their experimental counterparts is

considerably more systematic. The 6-311++G\*\* DFT(B3LYP) rotational constants are in the best overall agreement with the measured values.

Table 3 also contains theoretical and experimental quartic centrifugal distortion (QCD) constants and dipole moments. For  $\alpha$ -alanine **I** and **IIA**, where comparison of theory and experiment is possible, a nice agreement can be observed. While the directly calculated QCD constants of **I** are only in reasonable agreement with the measured values,<sup>9</sup> if the 6-311++G\*\* RHF quadratic force field is transformed into internal coordinate space and a crude, simple but well-established scaling of the force constants, similar to the one advocated in the scaled quantum mechanical (SQM) force field approach,<sup>71,72</sup> is applied (0.9 for the stretching and 0.8 for the bending coordinates), the agreement between theory and experiment becomes excellent, in line with previous theoretical studies of QCD constants.<sup>73–79</sup> It can also be seen from this and previous studies<sup>73,75–77</sup> at the HF level of theory that a simple multiplication of the directly calculated QCD constants by 1.2 brings the calculated values in much better agreement with experiment and with QCD constants calculated from SQM force fields. Note that QCD constants obtained only from scaled force fields are reported in Table 3 for the conformers for which no experimental data are available. Somewhat surprisingly, the agreement between theory and experiment is less satisfactory, even after scaling, for **IIA**. The reason for this is not obvious. The theoretical estimates given for the yet unobserved conformers should be most valuable during search for these conformers by MW and MMW spectroscopies.

**Vibrational Spectra.** The plethora of information contained in the 6-311++G\*\* RHF theoretical vibrational frequencies and infrared intensities presented in Table 4 for the  $\alpha$ -alanine conformers would allow interpretation of a carefully executed experimental investigation of the gas-phase vibrational spectrum of  $\alpha$ -alanine. As, to our best knowledge, no such study has been performed, this task is left for the future. Furthermore, due to the lack of experimental data, no elaborate scaling, like the one resulting in a scaled quantum mechanical (SQM) force field,<sup>71,72,80</sup> is attempted here. (Note that the uniform scale factor of 0.89 is a reasonable estimate for direct frequency corrections at the 6-311++G\*\* RHF level and can be applied when predictions of this study are to be compared with experiments.) The frequency results obtained confirm that all optimized structures **I–VIII B** correspond to minima on the potential energy surface of  $\alpha$ -alanine.

The results summarized in Table 4 reveal that there are some normal modes which do not change significantly from one conformer to another (including  $\nu_8(\text{C}=\text{O}$  str), which changes between 1993 (**I**) and 2031 (**VIIIA**)  $\text{cm}^{-1}$ , and the usually rather intense band  $\nu_{22}$ ), while some normal modes having high IR intensity shift considerably. Most notably, conformer **IIA** has characteristic, high-intensity bands:  $\nu_{13}$  at 1527  $\text{cm}^{-1}$  and  $\nu_{25}$  at 782  $\text{cm}^{-1}$ . High-intensity normal modes whose frequency does not change from one conformer to another should serve as indicators of the presence of  $\alpha$ -alanine, while characteristic shifts should help identification of  $\alpha$ -alanine conformers in its gas.

The frequency data of Table 4 also contain important information about the internal bondings of the conformers and about the relative strength of the different H-bonds. For example, the O–H stretch ( $\nu_1$ ) frequency shifts considerably from its free value of about 4115  $\text{cm}^{-1}$  (**I**) to 4034  $\text{cm}^{-1}$  (**IIA**) when involved in H-bonding. Some of this shift, however, is due to the change from a *cis* to a *trans* COOH arrangement, which itself<sup>15</sup> produces an about 60  $\text{cm}^{-1}$  shift. (Note that the

upward shift in the predominantly C=O stretch mode,  $\nu_8$ , from  $\alpha$ -alanine **I** to **II** can also be observed in formic acid. This change is also in line with the almost 0.007 Å shorter C=O bond length in *trans*-formic acid as compared to *cis*-formic acid.) A less pronounced shift in the N–H stretch frequencies ( $\nu_2$  and  $\nu_3$ ) can be observed upon H-bonding, also indicating that the N–H $\cdots$ O H-bond is considerably weaker than the O–H $\cdots$ N H-bond.

## Conclusions

The large amount of high-level *ab initio* data presented in this and previous<sup>16,18,25,34b</sup> studies not only shapes our understanding of the conformational behavior of simple free amino acids but also provides vital data for different types of experiments applicable to probe these molecules. In particular, the following conclusions can be drawn about the shape and spectra of free  $\alpha$ -alanine and about their determination by experimental techniques and by methods of electronic structure theory.

1. It seems that introduction of the methyl group for the amino acid  $\alpha$ -alanine with R = CH<sub>3</sub> to replace one of the hydrogens of glycine (R = H) has a rather small effect on either the geometry or the conformational preferences of  $\alpha$ -alanine. As a consequence, 13 stable conformers of  $\alpha$ -alanine have been located on  $\alpha$ -alanine's potential energy surface. Note that in two previous theoretical studies only six and 10 out of the 13 conformers have been found by Godfrey et al.<sup>9</sup> and by Gronert and O'Hair,<sup>34b</sup> respectively, while Cao et al.<sup>25</sup> located all 13 of them. Since characterization of conformationally flexible molecules, such as  $\alpha$ -alanine, requires the identification of all major conformers, these problems emphasize the need to conduct extremely careful, systematic surveys of all possible conformers at relatively high levels of theory when the conformational behavior of these species is to be determined.

2. There is a general tendency at the restricted Hartree–Fock (RHF) level of theory to overestimate the relative energies of all conformers of  $\alpha$ -alanine. This overestimation is, however, not systematic. More importantly, as calculations employing almost complete basis sets indicate, the RHF level of theory completely fails in predicting the energy difference between N–H $\cdots$ O and O–H $\cdots$ N H-bonds. (Unfortunately, these are just the type of interactions responsible for stabilization of the lowest energy conformers of glycine,  $\alpha$ -alanine, and most similar amino acids.) For the other, less important, conformers not only the predicted geometries but even the predicted relative energies seem to be reasonably accurate. The large set of data presented in this and previous papers<sup>16,18,34b</sup> do not seem to support earlier conclusions of Frey et al.<sup>27</sup> about the imprecision of RHF geometry optimizations of amino acids and their claim that single-point MP2 energies obtained at RHF-optimized geometries are “potentially highly inaccurate”. Still, the use of correlated levels of theory in the calculation of most of the relative energies of amino acid conformers is strongly advocated, as interest will probably be focused on the low-energy, H-bond-stabilized conformers.

3. The DFT(B3LYP) calculations on glycine<sup>39</sup> and  $\alpha$ -alanine suggest that at this relatively inexpensive level of electronic structure theory characterization of the conformational behavior of amino acids can be accomplished. Use of the B3LYP functional with sufficiently large basis sets may open the way to reliable studies on related medium-sized biomolecules, e.g., oligopeptides and their analogues.

4. Partly on the basis of the measured dipole moment components for conformers **I** and **IIA** and the observed integrated areas of the strongest lines of different conformers,

Godfrey et al.<sup>9</sup> reported experimentally determined approximate lower limits for conformational energies relative to those of  $\alpha$ -alanine **I**. All of their estimates are significantly different from the high-quality theoretical relative energy predictions of this report. The discrepancies deserve further study.

5. Energy estimates, based on the approximate, qualitative picture obtained by completely neglecting the effects of steric strain and lone electron pair repulsion and assigning independent energy contributions to the different possible H-bonds and the carboxylic functional group, clearly assess the importance of the related energy effects as follows: (a) the O—H $\cdots$ N H-bond has the largest stabilization effect; (b) the stabilization contribution of the N—H $\cdots$ O H-bond is only about half as large as that of the O—H $\cdots$ N H-bond, even though it's made up by two bifurcated H-bonds; (c) the N—H $\cdots$ OH bifurcated H-bond seems to be even weaker; (d) the energy difference between the *cis* and *trans* COOH arrangements is 1550(100) cm<sup>-1</sup>, favoring the *cis* form. Appropriate sums of the corresponding simple estimates agree within about 250 cm<sup>-1</sup> with the relative energy values obtained from high-level theoretical computations.

6. A GED analysis by Iijima and Beagley<sup>11</sup> resulted in the heavy-atom structure of the lowest energy conformer of  $\alpha$ -alanine. Their model for structure refinement, however, should be revised based on the newly available theoretical and experimental data, and the analysis should be repeated since significant changes in some structural parameters are expected.

7. Agreement between the 6-311++G\*\* DFT(B3LYP) and MP2(full) rotational constants and experiment is excellent for the two low-energy forms of  $\alpha$ -alanine, **I** and **IIA**. The rotational constants determined for the other, yet unobserved, conformers should have a comparable accuracy and thus aid the search for these conformers by rotational spectroscopy. Failure to detect **IIb**, a conformer having a large dipole moment and a relative energy similar to that of **IIA**, deserves special attention.

8. In line with previous theoretical studies<sup>73-79</sup> the calculated quartic centrifugal distortion (QCD) constants are in excellent agreement with experiment for **I** once the quadratic force field is scaled. Accuracy of the calculated QCD constants for **IIA** is considerably lower.

9. The calculated vibrational spectra of the conformers of  $\alpha$ -alanine reveal that different conformers have some high-intensity bands in different regions of the spectrum. These considerable shifts in the normal modes would probably allow identification of a number of conformers in the gas-phase or in the low-temperature inert gas matrix vibrational spectrum of  $\alpha$ -alanine.

**Acknowledgment.** Most of the research described was carried out during a stay of the author, supported by an Eötvös Fellowship of the Hungarian Ministry of Education, with Professor Ian M. Mills (Department of Chemistry, University of Reading), whose hospitality is gratefully acknowledged. The research described was also supported by the Scientific Research Foundation of Hungary (OTKA T017604). Most calculations reported have been performed on an IBM SP/2 computer system whose purchase was made possible by a grant from "Catching up with European Higher Education" (FEFA III/1052).

## References and Notes

- Levy, H. A.; Corey, R. B. *J. Am. Chem. Soc.* **1941**, *63*, 2095.
- Donohue, J. *J. Am. Chem. Soc.* **1950**, *72*, 949.
- Takagi, S.; Chihara, H.; Seki, S. *Bull. Chem. Soc. Jpn.* **1959**, *32*, 84.
- Junk, G.; Sevec, H. *J. Am. Chem. Soc.* **1963**, *85*, 839.
- Brown, R. D.; Godfrey, P. D.; Storey, J. W. V.; Bassez, M. P. *J. Chem. Soc., Chem. Commun.* **1978**, 547.
- (a) Suenram, R. D.; Lovas, F. J. *J. Mol. Spectrosc.* **1978**, *72*, 372. (b) Suenram, R. D.; Lovas, F. J. *J. Am. Chem. Soc.* **1980**, *102*, 7180.
- Brown, R. D.; Godfrey, P. D.; Storey, J. W. V.; Bassez, M. P.; Robinson, B. J.; Batchelor, R. A.; McCulloch, M. G.; Rydbeck, O. E. H.; Hjalmarsen, A. G. *Mon. Not. R. Astron. Soc.* **1979**, *186*, 5.
- Godfrey, P. D.; Brown, R. D. *J. Am. Chem. Soc.* **1995**, *117*, 2019.
- Godfrey, P. D.; Firth, S.; Hatherley, L. D.; Brown, R. D.; Pierlot, A. P. *J. Am. Chem. Soc.* **1993**, *115*, 9687.
- Iijima, K.; Tanaka, K.; Onuma, S. *J. Mol. Struct.* **1991**, *246*, 257.
- Iijima, K.; Beagley, B. *J. Mol. Struct.* **1991**, *248*, 133.
- (a) Reva, I. D.; Stepanian, S. G.; Plokhotnichenko, A. M.; Radchenko, E. D.; Sheina, G. G.; Blagoi, Y. P. *J. Mol. Struct.* **1994**, *318*, 1. (b) Reva, I. D.; Plokhotnichenko, A. M.; Stepanian, S. G.; Ivanov, A. Y.; Radchenko, E. D.; Sheina, G. G.; Blagoi, Y. P. *Chem. Phys. Lett.* **1995**, *232*, 141.
- (a) Debies, T. P.; Rabalais, J. W. *J. Electron Spectrosc. Relat. Phenom.* **1974**, *3*, 315. (b) Klasinc, L. *J. Electron Spectrosc. Relat. Phenom.* **1976**, *8*, 161.
- (a) Grenie, Y.; Lassegues, J. C.; Lagrange, C. G. *J. Chem. Phys.* **1970**, *53*, 2980. (b) Grenie, Y.; Lagrange, C. G. *J. Mol. Spectrosc.* **1972**, *41*, 240.
- Hu, C.-H.; Shen, M.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1993**, *115*, 2923.
- Császár, A. G. *J. Mol. Struct.* **1995**, *346*, 141.
- Wright, L. R.; Borkman, R. F. *J. Am. Chem. Soc.* **1980**, *102*, 6207.
- Császár, A. G. *J. Am. Chem. Soc.* **1992**, *114*, 9569.
- Palla, P.; Petrongolo, C.; Tomasi, J. *J. Phys. Chem.* **1980**, *84*, 435.
- van Alsenoy, C.; Scarsdale, J. N.; Sellers, H. L.; Schäfer, L. *Chem. Phys. Lett.* **1981**, *80*, 124.
- van Alsenoy, C.; Kulp, S.; Siam, K.; Klimkowski, V. J.; Ewbank, J. D.; Schäfer, L. *J. Mol. Struct. (THEOCHEM)* **1988**, *50*, 169.
- Schäfer, L.; Siam, K.; Klimkowski, V. J.; Ewbank, J. D.; van Alsenoy, C. *J. Mol. Struct. (THEOCHEM)* **1990**, *63*, 361.
- Schäfer, L.; Kulp-Newton, S. Q.; Siam, K.; Klimkowski, V. J.; van Alsenoy, C. *J. Mol. Struct.* **1990**, *209*, 373.
- (a) Sellers, H. L.; Schäfer, L. *J. Am. Chem. Soc.* **1978**, *100*, 7728. (b) Sellers, H. L.; Schäfer, L. *Chem. Phys. Lett.* **1979**, *63*, 609. (c) Siam, K.; Klimkowski, V. J.; Ewbank, J. D.; Van Alsenoy, C.; Schäfer, L. *J. Mol. Struct. (THEOCHEM)* **1984**, *110*, 171.
- Cao, M.; Newton, S. Q.; Pranata, J.; Schäfer, L. *J. Mol. Struct. (THEOCHEM)* **1995**, *332*, 251.
- Schäfer, L.; Sellers, H. L.; Lovas, F. J.; Suenram, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 6566.
- Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schäfer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369.
- Clementi, E.; Cavallone, F.; Scordamaglia, R. *J. Am. Chem. Soc.* **1977**, *99*, 5531.
- Dixon, D. A.; Lipscomb, W. N. *J. Biol. Chem.* **1976**, *251*, 5992.
- (a) Ramek, M. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1990**, *17*, 45. (b) Ramek, M.; Cheng, V. K. W.; Frey, R. F.; Newton, S. Q.; Schäfer, L. *J. Mol. Struct. (THEOCHEM)* **1991**, *235*, 1.
- (a) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 7917. (b) Jensen, J. H.; Baldrige, K. K.; Gordon, M. S. *J. Phys. Chem.* **1992**, *96*, 8340.
- Vishveshwara, S.; Pople, J. A. *J. Am. Chem. Soc.* **1977**, *99*, 2422.
- Destro, R.; Marsh, R. E.; Bianchi, R. *J. Phys. Chem.* **1988**, *92*, 966.
- (a) O'Hair, R. A. J.; Bowie, J. H.; Gronert, S. *Int. J. Mass Spectrom. Ion Processes* **1992**, *117*, 23. (b) Gronert, S.; O'Hair, R. A. J. *J. Am. Chem. Soc.* **1995**, *117*, 2071.
- Sapse, A. M.; Mallah-Levy, L.; Daniels, S. B.; Erickson, B. W. *J. Am. Chem. Soc.* **1987**, *109*, 3526.
- Yu, D.; Armstrong, D. A.; Rauk, A. *Can. J. Chem.* **1992**, *70*, 1762.
- Ding, Y.; Krogh-Jespersen, K. *Chem. Phys. Lett.* **1992**, *199*, 261.
- Vijay, A.; Sathyanarayana, D. N. *J. Phys. Chem.* **1992**, *96*, 10735.
- Barone, V.; Adamo, C.; Lelj, F. *J. Chem. Phys.* **1995**, *102*, 364.
- Frisch, M. J.; Pople, J. A.; Del Bene, J. E. *J. Phys. Chem.* **1985**, *89*, 3664.
- Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.
- (a) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650. (b) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294. (c) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.
- Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- Partridge, H. *Near-Hartree-Fock Quality Gaussian Type Orbital Basis Sets for the First- and Third-Row Atoms*. NASA Technical Memorandum 101044; 1989.
- (a) Huzinaga, S. *J. Chem. Phys.* **1965**, *42*, 1293. (b) Allen, W. D.; Schaefer, H. F., III. *J. Chem. Phys.* **1988**, *89*, 329.
- Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69; *Ibid.* **1960**, *32*, 179.
- Pulay, P. *Mol. Phys.* **1969**, *17*, 197.

- (49) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley-Interscience: New York, 1986.
- (50) (a) Möller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618. (b) Krishnan, R.; Frisch, M. J.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 4244.
- (51) (a) Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256. (b) Čížek, J.; Paldus, J. *Phys. Scr.* **1980**, *21*, 251.
- (52) Bartlett, R. J.; Dykstra, C. E.; Paldus, J. In *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*; Dykstra, C. E., Ed.; Reidel: Dordrecht, 1984, p 127.
- (53) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
- (54) (a) Scuseria, G. E.; Scheiner, A. C.; Lee, T. J.; Rice, J. E.; Schaefer, H. F., III. *J. Chem. Phys.* **1986**, *86*, 2881. (b) Scuseria, G. E.; Lee, T. J. *J. Chem. Phys.* **1990**, *93*, 5851. (c) Lee, T. J.; Taylor, P. R. *Int. J. Quantum Chem. Symp.* **1989**, *23*, 199. (d) Lee, T. J.; Scuseria, G. E. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer: Dordrecht, 1995; p 47.
- (55) (a) Dirac, P. A. M. *Proc. Cambridge Philos. Soc.* **1930**, *26*, 376. (b) Becke, A. D. *Phys. Rev. B* **1988**, *38*, 3098. (c) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (d) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. (e) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
- (56) Pople, J. A.; Frisch, M. J.; Luke, B. T.; Binkley, J. S. *Int. J. Quantum Chem., Symp.* **1983**, *17*, 307.
- (57) Handy, N. C.; Knowles, P. J.; Somasundram, K. *Theor. Chim. Acta* **1985**, *68*, 87.
- (58) Allen, W. D.; Császár, A. G. *J. Chem. Phys.* **1993**, *98*, 2983.
- (59) (a) Osamura, Y.; Yamaguchi, Y.; Saxe, P.; Fox, D. J.; Vincent, M. A.; Schaefer, H. F., III. *J. Mol. Struct.* **1983**, *103*, 183. (b) Handy, N. C.; Schaefer, H. F., III. *J. Chem. Phys.* **1984**, *81*, 5031.
- (60) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem., Symp.* **1979**, *13*, 225.
- (61) Stanton, J. F.; Gauss, J.; Watts, J. D.; Lauderdale, W. J.; Bartlett, R. J. *Int. J. Quantum Chem., Symp.* **1992**, *26*, 879.
- (62) Frisch, M. J.; et al. *Gaussian 94*, Revision B.2; Gaussian Inc.: Pittsburgh, PA, 1995.
- (63) *PSI Version 2.0*; Psitech, Inc.: Watkinsville, GA, 1991.
- (64) (a) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165. (b) *Turbomole User Guide*, version 2.3.0; Biosym. Technologies: San Diego, 1993.
- (65) Hedberg, L.; Mills, I. M. *J. Mol. Spectrosc.* **1993**, *160*, 117.
- (66) (a) Allen, W. D.; East, A. L. L.; Császár, A. G. *Structures and Conformations of Non-Rigid Molecules*; Laane, J., Dakkouri, M., van der Veken, B., Oberhammer, H., Eds.; Kluwer: Dordrecht, 1993; p 343. (b) Császár, A. G. *J. Phys. Chem.* **1994**, *98*, 8823. (c) Császár, A. G.; Allen, W. D. *J. Chem. Phys.* **1996**, *104*, 1. (d) Martin, J. M. L. *Chem. Phys. Lett.* **1995**, *242*, 343.
- (67) To evaluate core correlation effects the cc-pVTZ basis was completely uncontracted and polarized into the core region by adding two sets of tight d and f polarized manifolds to it, resulting in an extended basis denoted, similarly to previous studies,<sup>66c</sup> as CV(TZ). By freezing the C and O 1s core and all virtual orbitals greater than 150 au in energy, the *cis-trans* energy difference of formic acid was found to be 1564.2 cm<sup>-1</sup> at the MP2 level. An analogous set of MP2 calculations with all orbitals correlated subsequently gave an energy difference of 1564.9 cm<sup>-1</sup>, indicating a negligible, -0.7 cm<sup>-1</sup>, contribution of core-core and core-valence correlation to the *cis-trans* energy difference.
- (68) Räsänen, M.; Aspiala, A.; Homanen, L.; Murto, J. *J. Mol. Struct.* **1982**, *96*, 81.
- (69) Flood, E.; Pulay, P.; Boggs, J. E. *J. Am. Chem. Soc.* **1977**, *99*, 5570.
- (70) It is noted here that, using GED data, Ewbank and co-workers (Ewbank, J. D.; Klimkowski, V. J.; Siam, K.; Schäfer, L. *J. Mol. Struct.* **1987**, *160*, 275) determined the structure of the lowest energy conformer of the methyl ester of  $\alpha$ -alanine. As expected, also on the basis of results presented in this study, the molecular intensities measured proved to be consistent with a form corresponding to  $\alpha$ -alanine I.
- (71) Pulay, P.; Fogarasi, G.; Pongor, G.; Boggs, J. E.; Vargha, A. *J. Am. Chem. Soc.* **1983**, *105*, 7037.
- (72) Allen, W. D.; Császár, A. G.; Horner, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 6834.
- (73) Császár, A. G.; Fogarasi, G. *J. Chem. Phys.* **1988**, *89*, 7646.
- (74) Clabo, D. A., Jr.; Allen, W. D.; Remington, R. B.; Yamaguchi, Y.; Schaefer, H. F., III. *J. Chem. Phys.* **1988**, *123*, 187.
- (75) Császár, A. G.; Fogarasi, G.; Boggs, J. E. *J. Phys. Chem.* **1989**, *93*, 7644.
- (76) Włodarczyk, G.; Burie, J.; Demaison, J.; Vormann, K.; Császár, A. G. *J. Mol. Spectrosc.* **1989**, *134*, 297.
- (77) Császár, A. G. *Chem. Phys. Lett.* **1989**, *162*, 361.
- (78) Allen, W. D.; Yamaguchi, Y.; Császár, A. G.; Clabo, D. A., Jr.; Remington, R.; B.; Schaefer, H. F., III. *J. Chem. Phys.* **1990**, *145*, 427.
- (79) Bogey, M.; Bolvin, H.; Cordonnier, M.; Demuyneck, C.; Destombes, J. L.; Császár, A. G. *J. Chem. Phys.* **1994**, *100*, 8614.
- (80) An early application of SQM theory for molecules of biological interest was the prediction of the fundamental vibrations of uracil (Harsányi, L.; Császár, P.; Császár, A.; Boggs, J. E. *Int. J. Quantum Chem.* **1986**, *29*, 799).