

Quantum free energies of the conformers of glycine on an *ab initio* potential energy surface†

Thomas F. Miller, III and David C. Clary*

Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford, UK OX1 3QZ. E-mail: thomas.miller@chem.ox.ac.uk, david.clary@chem.ox.ac.uk

Received 13th November 2003, Accepted 2nd February 2004
First published as an Advance Article on the web 4th March 2004

The torsional path integral Monte Carlo (TPIMC) technique is used to study the five lowest-energy conformers of glycine. The theoretical method provides an anharmonic and quantum-mechanical description of conformational free energy and is used for the first time with an *ab initio* potential energy surface. The 3-dimensional torsional potential energy surface of glycine was obtained at the MP2/6-311++G** level of theory and is optimized with respect to the non-torsional degrees of freedom. Calculated conformer populations compare well with those reported in recent matrix-isolation infrared spectroscopy experiments. An additional conformer, not yet observed, is predicted to be heavily populated in the thermal equilibria probed by experiment, and a new explanation for its elusiveness is provided. Quantum effects, such as zero point energy, are found to substantially alter conformer populations, and an algorithm for estimating the role of non-torsional vibrations in the conformational thermodynamics of a molecule is introduced.

1 Introduction

Glycine, the simplest amino acid, is a benchmark system for the experimental and theoretical determination of molecular conformation. It is among the smallest compounds to exhibit intramolecular hydrogen bonds, which stabilize various structural configurations and give rise to a subtle potential energy surface.^{1–4} With flexible degrees of freedom and low-energy pathways between local minima, the potential energy surface of glycine contains the basic elements of the complex configurational landscapes which govern protein folding and other sophisticated biological processes.^{5,6}

In the gas phase, glycine exists in the non-zwitterionic isomer.^{7,8} Extensive theoretical analysis of the electronic structure of glycine has revealed multiple stable conformers arising from internal rotation about its three torsional degrees of freedom.^{9–18} High level *ab initio* calculations predict five glycine conformers to exist within 10 kJ mol⁻¹ of the global potential minimum.^{15,16} Fig. 1 presents the structures and energies of the lowest energy conformers of glycine first reported by Császár and reproduced in the current study. The torsional angles $\vec{\phi} = (\phi_1, \phi_2, \phi_3)$ in the figure indicate the coordinate of the twisting motions about the C–C bond (ϕ_1), the C–O single bond (ϕ_2), and the C–N bond (ϕ_3).

Conformer I is generally calculated to be the lowest energy conformer and to have a geometry of C_s symmetry.⁹ However, the level of theory employed alters the prediction of whether conformer II exhibits a symmetry plane. Early studies found the conformer to be symmetric, but the subsequent inclusion of electron correlation and larger basis sets revealed that the minimized geometry distorts in favor of strengthening hydrogen-bonding interactions.^{14–16} The heavy atoms of conformers III–V are all calculated to be non-planar. Every glycine structure of C₁ symmetry exhibits a mirror-image structure (with inverted torsional angles) of equivalent energy.¹³

The determination of glycine conformer populations is a challenging experimental problem. The structure of glycine conformers in the gas phase was first probed using microwave and millimeter wave spectroscopy.^{19,20} Because of its greater dipole moment, conformer II is more easily detected using rotational spectroscopy and was observed before the lower-energy conformer I.^{21,22} Other glycine conformers have not been found using millimeter wave or microwave spectroscopy, most likely because they interconvert to conformer I during the free jet expansion step of the experiments.^{23,24} Conformer interconversion is a shift in the population of molecules from one conformer to another during the cooling period between thermal equilibration and experimental observation.²⁵

A gas-phase electron diffraction study of the glycine conformers also identified at least two conformers of glycine.⁸ In

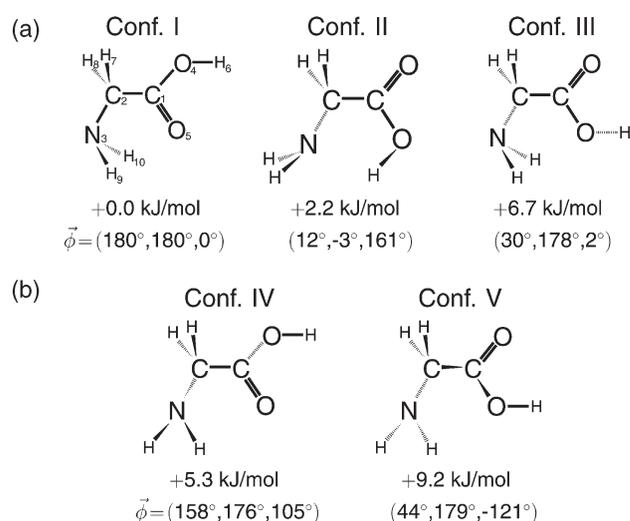


Fig. 1 The most stable conformers of glycine. Energies are reported at the MP2/6-311++G** level of theory.¹⁵ The torsional angles $\vec{\phi} = (\phi_1, \phi_2, \phi_3)$ correspond to intramolecular rotation about the C–C bond, C–O single bond, and C–N bond, respectively. Conformers in (a) have been observed experimentally.

† Electronic supplementary information (ESI) available: Calculated harmonic frequencies of the glycine conformers. See <http://www.rsc.org/suppdata/cp/b3/b314644h/>

principle, this technique enables direct observation of the thermal population distribution and avoids the problem of conformer interconversion. However, electron diffraction only determines the heavy atom positions in a molecule, so it has difficulty determining the relative population of conformers, such as II and III, which exhibit overlapping heavy-atom thermal distributions.

When coupled with the technique of matrix isolation, infrared (IR) spectroscopy has proven to be a valuable tool for studying the conformers of glycine.^{26–29} The analysis of glycine with conventional gas-phase IR spectroscopy is precluded by the low thermal stability of the molecule.²⁸ Matrix-IR spectroscopy is the only experimental technique for which all three conformers in Fig. 1a have been conclusively observed. Matrix deposition is a more gentle process than free jet expansion, so conformer interconversion is less problematic for this technique than for microwave spectroscopy.^{26,28} Also, IR spectroscopy directly probes the structure of the glycine molecules and discriminates between conformers more readily than electron diffraction. Nonetheless, the non-ideal behavior of the matrix gas can affect the measurement of conformer populations, and the absence of the low-energy conformer IV from the observations of matrix-IR studies requires explanation.

The accurate determination of glycine conformer populations also poses serious challenges for theory. It was previously mentioned that electron correlation and a flexible basis set are necessary for describing the intramolecular hydrogen-bonding structure of the glycine conformers. However, the theoretical difficulties of calculating the conformer populations of glycine extend beyond those of electronic structure theory. Glycine is a highly flexible molecule with shallow potential barriers separating minima of competitive energy.^{13,23,24} Moreover, the pathways of interconversion between conformers coincide with torsional degrees of freedom exhibiting low moments of inertia. Quantitative theoretical determination of the equilibrium population distribution of glycine therefore requires an anharmonic, quantum-mechanical description of the conformational thermodynamics. To interpret experiment, it is not sufficient to calculate only the potential energy of the conformers. The temperature-dependent free energy must also be computed.

In the current study, the recently developed torsional path integral Monte Carlo (TPIMC) technique is used to study the conformer free energies of glycine. The TPIMC technique provides a completely anharmonic description of the intramolecular torsional degrees of freedom and systematically includes quantum effects such as zero point energy.^{30,31} Although TPIMC has been successfully employed for the prediction of molecular conformation,³² this is the first study in which the technique is performed on an *ab initio* potential energy surface.

2 Methodology

2.1 Torsional path integral Monte Carlo

Torsional path integral Monte Carlo (TPIMC) is a quantum mechanical technique for calculating the equilibrium thermodynamics of molecules at finite temperature.^{30,31} It has been previously employed to investigate the quantum thermodynamics of large molecules^{30,31} and was recently shown to be a powerful tool for investigating the conformational preference of biomolecules.³² The TPIMC technique extends the applicability of the more general path integral Monte Carlo (PIMC) technique^{33–35} to larger molecules by explicitly treating only the intramolecular torsional degrees of freedom. Other implementations of PIMC have enabled the calculation of molecular thermodynamics for two- and three-atom systems using high-quality potentials,^{36,37} but the authors are unaware of any previous study in which the technique has been employed with an

ab initio potential to explore the conformer free energies of a biomolecule.

A convenient feature of path integral theory is that a single parameter n_b modulates the accuracy of the calculations between exact classical mechanics ($n_b = 1$) and exact quantum mechanics ($n_b = \infty$), neglecting Monte Carlo sampling error.³⁴ For intramolecular torsions, it has been shown that the bulk of quantum effects are recovered using only a value of $n_b \approx 5$.³¹ Upon varying the parameter n_b , TPIMC calculations explore the extent to which quantum effects are significant in the conformational thermodynamics of biomolecules.

The details of the TPIMC technique have been explained elsewhere,^{30,31} so we include only the major features of the calculations reported in this study. All three torsional degrees of freedom in the glycine molecule were treated explicitly. Simulations were performed at 100 K, 150 K, 200 K, 300 K, 400 K, 500 K, and 700 K. At each temperature, TPIMC calculations were performed with the parameter n_b set to 1, 5, and 15. We will subsequently refer to the TPIMC technique employed with $n_b = 1$ as “classical TPIMC” and to the TPIMC technique employed with $n_b = 15$ as “quantum TPIMC”. Torsional sampling was performed using the free rotor (FR-TPIMC) algorithm to ensure that the entire torsional configuration space of glycine was explored.³¹ Simulations were tightly converged to ensure minimal uncertainty and negligible estimator bias. Total enthalpies and free energies for glycine were calculated with estimated standard deviation of less than 0.0129 kJ mol⁻¹ and 0.0066 kJ mol⁻¹, respectively. In every case, estimator bias was calculated to be within the standard deviation of the corresponding value. The TPIMC program developed in our research group uses several subroutines from the TINKER software package.³⁸ The current TPIMC algorithm scales as n_b^3 , and the longest TPIMC calculations reported in this study were completed within eight days on a single Intel Pentium III processor (1.26 GHz).

2.2 *Ab initio* torsional potential energy surface

The TPIMC technique can be employed using any torsional potential energy surface. All previous studies have used atom-atom molecular mechanics force fields to describe intramolecular potential energy.^{30–32} The current study is the first application of the TPIMC technique on an *ab initio* potential energy surface.

All electronic structure calculations reported in this study were performed using second-order Møller–Plesset perturbation theory (MP2)³⁹ with the 6-311++G** basis set of Pople *et al.*,⁴⁰ which contains 145 contracted Gaussian functions for glycine. An important conclusion from previous theoretical work is that MP2 theory with a sufficiently large basis set accurately describes the relative energies of the conformers of glycine.^{15,16} Recent studies of glycine anharmonic vibrational frequencies further indicate that this level of theory provides a satisfactory description of the anharmonic regions of the potential energy surface.^{24,41,42} The Gaussian 98 package was employed for all electronic structure calculations.⁴³

The glycine potential energy surface calculated in this study has three dimensions corresponding to torsional angles $\vec{\phi} = (\phi_1, \phi_2, \phi_3)$. A $12 \times 12 \times 12$ grid was constructed by calculating the glycine energy at each $\phi_i = 30n^\circ$, $n = 1, 2, \dots, 12$. At each grid point, a constrained energy minimization was performed to optimize all non-torsional degrees of freedom. This minimization allowed the bond angles and stretches to relax with respect to the changing torsional angles.

An analytic potential energy surface was obtained by interpolating the grid points with an expansion of periodic functions of the form $g(\phi_1, \phi_2, \phi_3) = f_1(\phi_1) f_2(\phi_2) f_3(\phi_3)$, where $f_i(\phi_i) \in \{\sin(2\pi n\phi_i/360^\circ), \cos(2\pi n\phi_i/360^\circ)\}$ and $n \in \{0, 1, \dots, 6\}$. All 1728 orthogonal functions of this form were included in the expansion to yield an exact fit of the calculated grid points.

The potential energy surface fit described above accurately recovers the geometry and energy of almost every major glycine conformation. The only exception is conformer II, for which the surface is not sufficiently flexible to predict the slight out-of-plane relaxation of the geometry to $\vec{\phi} = (\pm 12^\circ, \mp 3^\circ, \pm 161^\circ)$. To include this feature, the analytical potential energy surface is supplemented with two Gaussian “dimple” functions of the form $g(\vec{\phi}) = \alpha \exp((\vec{\phi} - \vec{\phi}_0)^2 / \beta^2)$, where $\alpha = -1.0 \text{ kJ mol}^{-1}$, $\beta = 13^\circ$, and $\vec{\phi}_0 = (\pm 14.5^\circ, \mp 6^\circ, \pm 158^\circ)$. Although preliminary results indicate that this minor adjustment has negligible impact on the predicted conformer populations, the correction provides a better analytical fit of this important region of the potential energy surface.

In Table 1, we compare the stationary point energies of glycine on the torsional potential energy surface with reference energies from MP2/6-311++G** geometry optimizations. The conformer labels are consistent with those of Császár.¹⁵ For no stationary point does the torsional potential energy surface deviate from its reference value by more than 0.16 kJ mol⁻¹. The average RMS difference between the torsional potential and the reference energies for the stationary points is 0.0058 kJ mol⁻¹.

We also point out a technical detail regarding the definition of torsional angle ϕ_3 . In the calculation of the *ab initio* grid points, ϕ_3 was defined as the dihedral angle C₁-C₂-N₃-X, where X is a dummy atom between H₉ and H₁₀ (Fig. 1). The dihedral angle C₁-C₂-N₃-H₉ was fixed at $\phi_3 + 60^\circ$ so that changes in ϕ_3 correspond to changes in the position of real atoms. The torsional angle C₁-C₂-N₃-H₁₀ was left unconstrained. For no conformer in this study does the ϕ_3 angle differ from the bisector of the C₁-C₂-N₃-H₉ and C₁-C₂-N₃-H₁₀ dihedral angles by more than 2°.

2.3 Conformer coordinate ranges

The TPIMC technique uses quantum Boltzmann statistics to generate a multidimensional probability distribution function in terms of the torsional angles. Conformer populations are calculated by integrating this function over the range of coordinates ascribed to each particular conformer. The integral of the entire distribution function over all coordinates is unity. A similar energy distribution is also generated to enable the evaluation of conformer enthalpies.

To calculate the conformer populations (and other thermodynamic quantities), it is necessary to pre-define the coordinate

Table 1 Relative energies (kJ mol⁻¹) of the glycine stationary points^a

	Császár ^b	Current study ^c	Torsional PES ^d
Ip	0.00	0.00	0.00
Iip	2.45	2.46	2.48
IIn	2.20	2.21	2.24, 2.29
IIIp	6.66	6.66	6.67
IIIn	6.04	6.03	6.02, 6.08
IVp	19.57	—	19.54
IVn	5.29	5.29	5.15, 5.13
Vp	22.50	—	22.49
Vn	9.20	9.20	9.16, 9.16
VIp	23.87	—	23.89
VIIp	29.13	—	29.24
VIIIp	46.80	—	46.76
VIIIIn	29.64	—	29.51, 29.53

^a Conformer labels are consistent with those of Császár.¹⁵ ^b Single point MP2/6-311++G** energies calculated at MP2(full)/6-311++G** optimized geometries.¹⁵ ^c Energies calculated at MP2/6-311++G** optimized geometries. ^d For non-planar minima, the energies for both symmetry-related structures are reported. See text for details of the potential energy surface (PES) construction.

ranges of each conformer. In general, the torsional configuration space is partitioned into coordinate ranges approximately along the potential barriers that separate the minima of neighboring conformers. Because the probability distribution function is largest in the potential wells and smallest in the barrier regions, the exact choice of partition has little effect on calculated quantities. Informed by analysis of the *ab initio* torsional potential energy surface of glycine, we chose the conformer coordinate ranges reported in Table 2. The ϕ_1 and ϕ_2 ranges were divided into two partitions, and the ϕ_3 range was divided into three partitions. The ϕ_1 boundaries for conformers I and III were chosen carefully because the torsional distribution is non-zero in the barrier range separating these conformers at higher temperatures, as will be seen later. Other conformers, in addition to conformers I–V, were also analysed in this study. However, they were effectively unpopulated at the temperatures considered and are not discussed in detail.

2.4 Harmonic frequency calculations

For conformers I–V, unconstrained geometry optimizations and harmonic frequency calculations were also performed at the 6-311++G**/MP2 level of theory. Conformers II–V were allowed to converge to their non-planar minima so that all frequencies were positive. The conformer frequency vibrations are included in the electronic supplementary information (ESI).[†] Harmonic thermodynamic quantities were obtained from the standard statistical mechanical expressions for the harmonic oscillator.⁴⁴ Because experimental conformer populations for glycine are available primarily from infrared spectroscopy studies in which the sample molecules are fixed in an inert gas matrix, the translational and rotational partition functions of the glycine conformers are neglected throughout this study.²⁸

3 Results and discussion

3.1 TPIMC conformer populations

Fig. 2 displays the calculated conformer populations of glycine as a function of temperature. Calculations were performed using the quantum TPIMC technique and reflect the equilibrium thermal distribution of the glycine conformers. The populations of conformers II–V are expanded in Fig. 2b. The population of no higher-energy conformer is calculated to exceed 1% in this temperature range. In Fig. 2b, the estimated standard deviation of the reported populations are less than 0.1% for all conformers and all temperatures.

The TPIMC calculations predict that conformer I, which corresponds to the global minimum of the potential energy surface, is the most populated conformer at every considered temperature. Conformer II is the most thermally accessible of the other conformers and is found to contain 11% of the total glycine population at only 300 K (Fig. 2b). With increasing temperature, the population of conformer II reaches a maximum of 12.5% at about 400 K before decreasing as higher energy conformers become thermally accessible. The populations of conformers III and IV are calculated to increase

Table 2 Torsional angle ranges for the glycine conformers in the TPIMC calculations

	ϕ_1	ϕ_2	ϕ_3
I	[100°, 260°]	[90°, 270°]	[-60°, 60°]
II	[-90°, 90°]	[-90°, 90°]	[-180°, -60°] + [60°, 180°]
III	[-100°, 100°]	[90°, 270°]	[-60°, 60°]
IV	[90°, 270°]	[90°, 270°]	[-180°, -60°] + [60°, 180°]
V	[-90°, 90°]	[90°, 270°]	[-180°, -60°] + [60°, 180°]

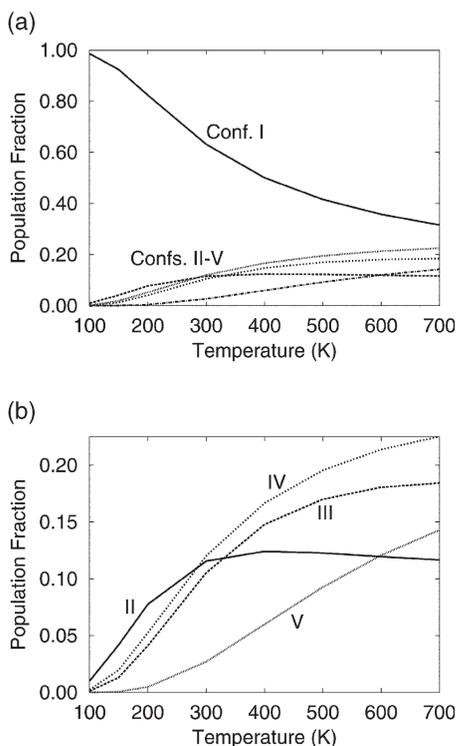


Fig. 2 (a) Glycine conformer populations calculated using the quantum TPIMC technique at various temperatures. (b) Same as (a) with lower range expanded.

rapidly between 200 K and 500 K and to surpass the population of conformer II at ≈ 400 K. Conformer IV, though it has not yet been experimentally observed, is calculated to have a slightly higher equilibrium population than conformer III at all temperatures. Conformer V is also found to be a substantial component of the thermal distribution above 300 K.

Growth in the population of conformers III–V with respect to conformer II as temperature rises (Fig. 2b) suggests that entropy plays a major role in the relative free energies of the conformers. Previous theoretical work has determined that conformer III exhibits high standard entropy, and this result is confirmed by the TPIMC results presented in Table 3.¹³ The table displays the calculated relative enthalpies and entropies of the glycine conformers at various temperatures. The conformer enthalpies are weakly dependent on changing temperature and reflect the relative energies of the optimized conformer structures in Fig. 1. The calculated entropies, however, are more dependent on temperature and vary substantially between the conformers. In particular, during the temperature

Table 3 Calculated relative standard enthalpies and entropies of the glycine conformers^a

T/K	$H^\circ/\text{kJ mol}^{-1}$					$S^\circ/\text{J mol}^{-1} \text{K}^{-1}$				
	I	II	III	IV	V	I	II	III	IV	V
100	0.0	3.8	6.5	5.4	9.6	0.0	-0.3	7.9	4.0	4.9
150	0.0	3.7	6.3	5.4	9.7	0.0	-1.2	6.9	4.3	5.4
200	0.0	3.5	6.2	5.4	9.8	0.0	-2.1	5.9	4.3	6.3
300	0.0	3.1	5.8	5.5	10.2	0.0	-3.7	4.6	4.5	7.7
400	0.0	2.9	5.5	5.7	10.3	0.0	-4.3	3.7	5.0	8.2
500	0.0	3.0	5.3	5.9	10.4	0.0	-4.1	3.2	5.5	8.3
600	0.0	3.3	5.1	6.1	10.3	0.0	-3.7	2.9	5.9	8.2
700	0.0	3.8	5.1	6.3	10.3	0.0	-2.8	2.7	6.2	8.1

^a Enthalpies and entropies calculated using the TPIMC technique with 15 beads.

range in which conformer III rapidly gains population (150 K to 400 K), its entropy is higher than the entropy of conformer II by about 8 kJ mol^{-1} and exceeds the entropy of conformer I by $4\text{--}8 \text{ kJ mol}^{-1}$. At higher temperatures, the fact that conformers IV and V become entropically favored coincides with their calculated rise in population (Fig. 2b).

The calculated torsional distributions in Fig. 3 provide insight into the relative entropies of the experimentally observed conformers I, II, and III. In the figure, the 3-dimensional torsional probability distribution obtained with quantum TPIMC is reduced along the C–C torsional angle, ϕ_1 . The conformer populations reported in Fig. 2 may be recovered by integrating the distribution curves with respect to ϕ_1 .

Because of its symmetric minimum energy structure, it is expected that the torsional distribution of conformer I has a single peak at $\phi_1 = 180^\circ$. However, the minimum energy geometry of conformer II deviates from C_s symmetry with $\phi_1 = \pm 12^\circ$. It is therefore notable that the torsional distribution of conformer II exhibits only one maxima in the heavy atom positions at $\phi_1 = 0$ for every temperature (Figs. 3c and 3d). This is consistent with the findings of microwave spectroscopy and indicates that the shallow barrier separating the local minima of conformer II is easily overcome.^{22,23,45}

As was previously noted in regard to Fig. 2b, the population of conformer II rises with temperature until approximately 400 K, after which it diminishes. This is again illustrated in Figs. 3c and 3d. Between 300 K and 400 K, as the torsional distribution spreads, Fig. 3d shows that the distribution peak height lowers despite a slight increase in integrated conformer population.

Conformer III is seen to have a flatter torsional distribution than those of conformers I and II (Fig. 3b). The flexibility of conformer III and the large volume of configuration space occupied by its distribution are consistent with the high entropy reported for the conformer in Table 3. It is seen that conformer III exhibits two peaks in the calculated torsional distribution which correspond to the symmetric, non-planar minima of the conformer at $\phi_1 = \pm 30^\circ$. However, the peaks increasingly overlap at higher temperatures as the distribution converges toward a broad single peak spanning $\phi_1 = [-60^\circ, 60^\circ]$.

This behavior explains the relative entropy of conformers I and III as a function of temperature (Table 3). Conformer I exhibits a single distribution peak at every temperature and therefore has a statistical degeneracy number of one. At low temperatures, when the distribution peaks of conformer III are fairly distinct, it will exhibit two-fold statistical degeneracy. The greater degeneracy of conformer III contributes to its larger entropy. The degeneracy number of conformer III will reduce to one as the distribution peaks merge at higher temperatures, so the corresponding decrease in the relative conformer entropies is to be expected.

Unlike conformer III, the equivalent minima for both conformers IV and V are well separated on the potential energy surface, so they exhibit two-fold statistical degeneracy at all temperatures. As a result, the relative entropies of conformers IV and V rise with respect to conformer III as temperature increases.

For higher temperatures, it is seen in Fig. 3b that the distribution of conformer III is non-zero at $\phi_1 = \pm 100^\circ$, the position of the potential barrier preventing interconversion of conformer III to conformer I. Therefore, depending on how the glycine molecules are cooled for experimental analysis, the ideal equilibrium populations of conformers I and III will not necessarily be observed.

3.2 Comparison with experiment

Fig. 4 displays the calculated and experimentally observed populations of glycine conformers II and III as a function of temperature. The relative populations of conformers I, II,

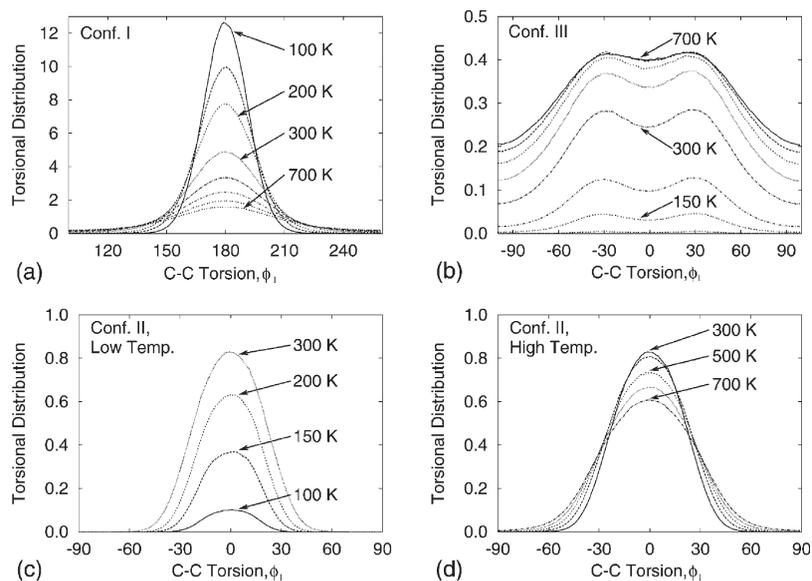


Fig. 3 Probability distribution functions for C–C torsional angle ϕ_1 calculated at various temperatures using the quantum TPIMC technique. (a) Conformer I. (b) Conformer III. (c) Conformer II at low temperature. (d) Conformer II at high temperature.

and III have been reported in two recent matrix-IR studies performed using neon and argon matrices.^{28,29} Since conformers IV and V have not been conclusively observed experimentally, it is assumed that molecules in these conformations at equilibrium interconvert to conformer I under experimental conditions. The population of conformer I is therefore complementary to the sum of the populations in the figure.

The neon matrix-IR populations were obtained by measuring the absorption intensities of the conformer C=O stretching vibrations for two different temperatures.²⁹ In this study, the conformer populations ratios are reported explicitly at 358 K, and the shift in the conformer equilibrium at 438 K is estimated. The argon matrix-IR conformer populations were obtained by comparing the calculated and observed intensities of the C=O stretching vibrations at 443 K.²⁸ No error estimate was given for the measurements in argon.

The conformer populations calculated using the TPIMC technique are in close agreement with the experimental results. For conformer III, the interpolated theoretical results coincide with the experimental population of $\approx 13\%$ determined at 358 K in the Ne matrix and remain within the uncertainty of that study at 438 K. The calculated population for conformer III at 443 K is also in precise agreement with the measurement in the Ar matrix.

For conformer II, the calculated results are within the error bars of the Ne matrix-IR experiments at 358 K and 438 K. At ≈ 440 K, it is seen that the measured population of conformer

II differs by approximately 5% depending on whether the neon or argon matrix is employed. However, it is likely that the experiments employing the lighter neon gas provides a more appropriate basis for comparison with the gas-phase calculations.

Despite the good agreement between theory and experiment in Fig. 4, quantitative comparison of conformer populations determined from gas-phase theoretical calculations and matrix-IR spectroscopy experiments is vulnerable to several sources of error. In addition to the possible flaws of the potential energy surface and assumptions of the theoretical technique, the identity of the matrix is known to affect experimentally determined conformer populations.²⁹ Neon and argon atoms, although chemically inert, contain electron density that will alter the local environment of the sample molecules in the matrix. Therefore, the matrix can induce real changes in conformer populations as well as perceived changes arising from distortion of the IR spectra. The development of techniques for doping helium droplets with glycine molecules improves the possibility of obtaining effectively gas-phase conformer populations of glycine in the near future.⁴⁶

For the matrix-IR results discussed above, the substrate temperature was kept very low (5 K for Ne, 8 K for Ar) with the intention of minimizing changes in the glycine conformer populations during matrix deposition. However, the disappearance of conformer peaks during sample cooling is difficult to avoid, particularly when low-energy pathways of interconversion are available.²⁵ It was mentioned previously that conformer III of glycine is not found in microwave spectroscopy experiments because the potential barrier for III \rightarrow I conformer interconversion is small enough to be overcome during the collisional cooling of free jet expansion.²³ Although the presence of conformer III in the matrix-IR experiments suggests that matrix deposition cooling is more gentle than free jet expansion, the possibility of interconversion cannot be discounted.

It was seen in Fig. 2b that the TPIMC technique predicts $\approx 15\%$ of the total glycine population to be in conformer IV in the equilibrium distributions at 350–450 K. The absence of this conformer in the matrix-IR spectra, particularly in light of the otherwise good agreement between theory and experiment, suggests that a large portion of the equilibrium population of conformer IV is collapsing to conformer I during matrix cooling. Previous theoretical studies have dismissed the likelihood of isolating conformer IV on the basis that the

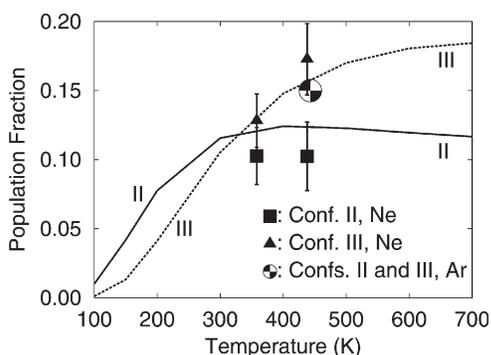


Fig. 4 Glycine conformer populations calculated using the quantum TPIMC technique (lines) and measured experimentally using matrix-IR spectroscopy in neon²⁹ and argon.²⁸

calculated barrier for IV \rightarrow I interconversion was substantially lower than that for III \rightarrow I interconversion. However, these electronic structure calculations either relied on small basis sets¹³ or did not optimize the molecular geometry as a function of the interconversion coordinate.²⁴ The potential energy surface employed in the current study has neither of these deficiencies.

Fig. 5 displays the calculated potential curves along the coordinate of interconversion for conformers III and IV to conformer I. The III \rightarrow I interconversion takes place along the C–C torsion, ϕ_1 , and IV \rightarrow I occurs along the C–N torsion, ϕ_3 . In obtaining the curves, the torsional coordinate of interconversion was fixed and the remaining torsions were allowed to relax on the potential energy surface. The figure reveals barrier heights of 3.15 kJ mol⁻¹ and 2.97 kJ mol⁻¹ for interconversions III \rightarrow I and IV \rightarrow I, respectively. These barrier heights are markedly closer in energy than has been previously reported, so an alternative explanation for the absence of conformer IV in the experiments is required.

We suggest two possible explanations for the observation that conformer IV is more likely to interconvert to conformer I than conformer III. In the first explanation, we note that the minimum energy cross section of the potential energy surface along the reaction coordinate (Fig. 5) does not necessarily provide an adequate dynamical description of conformer interconversion. Among other things, this minimized surface neglects the entropic effects and internal energy arising from the other degrees of freedom. It is well known from transition state theory that the free energy surface, or the potential mean force, incorporates these thermodynamic effects in an averaged fashion and may provide a more appropriate set of barrier heights.⁴⁷ Once we have obtained the glycine torsional distribution functions from the TPIMC calculations, it is a simple task to convert them into free energy surfaces for coordinates ϕ_1 and ϕ_3 .⁴⁸ Doing so for a temperature of 400 K, we find that the free energy barriers for III \rightarrow I and IV \rightarrow I are 3.72 kJ mol⁻¹ and 2.98 kJ mol⁻¹, respectively. That is, the difference in barrier heights on the free energy surface is over four times the difference in energy on the minimized potential energy

surface. Using a simple Arrhenius analysis, we find that the ratio of interconversion rates $k_{IV \rightarrow I}/k_{III \rightarrow I}$ is about 20% higher using the free energy surface rather than the potential energy surface. In agreement with experiment, this finding suggests that conformer III will be less likely to interconvert than conformer IV.

The second factor favoring the observation of conformer III in preference to conformer IV is statistical. Note from Fig. 5 that conformers III and IV each have two corresponding local minima. The minima of conformer III are close neighbors on the potential energy surface separated by a low barrier, but the minima of conformer IV are separated by either a larger barrier or the deep potential well of conformer I. Therefore, in the event that a glycine molecule in one of the potential wells corresponding to conformer IV undergoes a collision strong enough to induce interconversion, the only likely outcome is that it falls into the neighboring potential well of conformer I. On the other hand, if a glycine molecule in one of the potential wells of conformer III undergoes interconversion, there are two possible results: (i) the deep potential well of conformer I or (ii) the other shallow well of conformer III. Because the population of conformer III can shuttle between its two local minima, it is expected to survive more collisions than the population of conformer IV which can only collapse to conformer I. Although the rotationally inelastic scattering of glycine and the quantum dynamics of the ϕ_1 torsional angle have been recently explored theoretically,^{49,50} verification of these hypothesis requires a comparative dynamical study of the III \rightarrow I and IV \rightarrow I conversion pathways and is currently in progress.

It is noted that other previous studies have used experimental techniques to estimate the energetics and equilibrium populations of the glycine conformers. The populations at ≈ 500 K (219 °C) have been reported using gas-phase electron diffraction by Iijima *et al.*⁸ This study determined the combined population of conformers II and III to be 24%, which bears only qualitative agreement with the matrix-IR data and the TPIMC value of 31% in Fig. 4. However, to interpret the electron diffraction pattern, the authors relied on structural parameters obtained from early theoretical calculations and highly simplified expressions for the glycine potential energy surface. Other theoretical studies have had similar difficulties in trying to reproduce the conformer structures and energies predicted in this study.^{15,16,18}

In 1980, Suanram and Lovas estimated the relative minimum energies of conformers I and II to be 5.9 ± 1.8 kJ mol⁻¹ using millimeter wave spectroscopy.²¹ This study neglected the glycine intramolecular partition function and relied on a crude estimate of the conformer dipole moments.²¹ The discrepancy of this energy difference with subsequent theoretical evidence^{15,16} most likely arises from the fact that the experimental estimation did not account for the collapse of the substantial populations of conformers III and IV to conformer I. Therefore, the measured equilibrium population of conformer I would have been too large, leading to an over-estimation of the energy difference between the conformers. More recent microwave spectroscopy experiments sacrificed frequency response in favor of sensitivity, so the relative population and energetics of conformers I and II observed in these experiments was not determined.⁵¹

3.3 Quantum effects

Up to this point, only the TPIMC calculations which provide a quantum mechanical description of torsional motion have been discussed. However, it is of interest to consider how the glycine conformer populations and torsional distributions are altered by quantum effects. Fig. 6 reports the population of conformer II at various temperatures as determined by TPIMC calculations performed with the parameter n_b set to

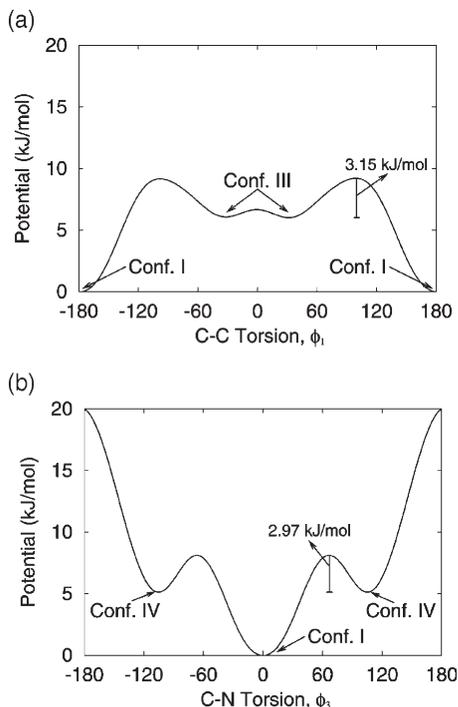


Fig. 5 Cross sections of the glycine potential energy surface minimized with respect to (a) the C–C torsional angle ϕ_1 and (b) the C–N torsional angle ϕ_3 .

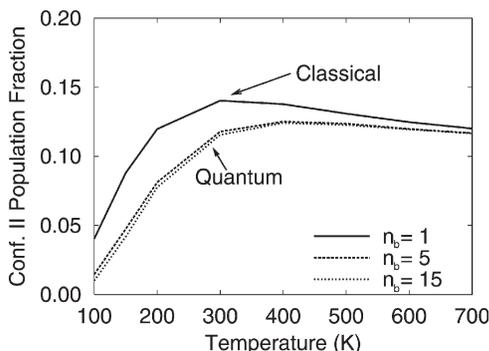


Fig. 6 Population of conformer II calculated at various temperature using $n_b = 1$ (classical TPIMC), $n_b = 5$, and $n_b = 15$ (quantum TPIMC).

1, 5, and 15. TPIMC calculations with $n_b = 1$ are equivalent to a classical mechanical description, and the calculations $n_b = 15$ are expected to recover almost all quantum effects for glycine in this temperature range. In the figure, the very small differences between the $n_b = 5$ and $n_b = 15$ results support this conclusion.

Fig. 6 shows that quantum effects are greatest at lower temperatures, whereas the quantum and classical descriptions converge at high temperatures. At 200 K, the classical result over-estimates the population of conformer II by a factor of ≈ 1.5 . Even at the higher temperatures probed by the matrix-IR experiments (350–450 K), differences arising from quantum effects remain of similar magnitude to differences attributed to the identity of the matrix gas (Fig. 4). Necessarily, quantum effects changed the populations of the other glycine conformers, but to a lesser extent than seen here for conformer II.

Quantum effects on the torsional distribution of the glycine conformers are illustrated in Fig. 7. The 3-dimensional torsional distribution of glycine is reduced over $\phi_1 = [-90, 90]$ and plotted as a function of ϕ_2 and ϕ_3 . For this range of ϕ_1 , the torsional distributions of conformers II, III, and V can be simultaneously displayed. The left column displays the classical TPIMC distribution, and the right column displays the quantum TPIMC distribution.

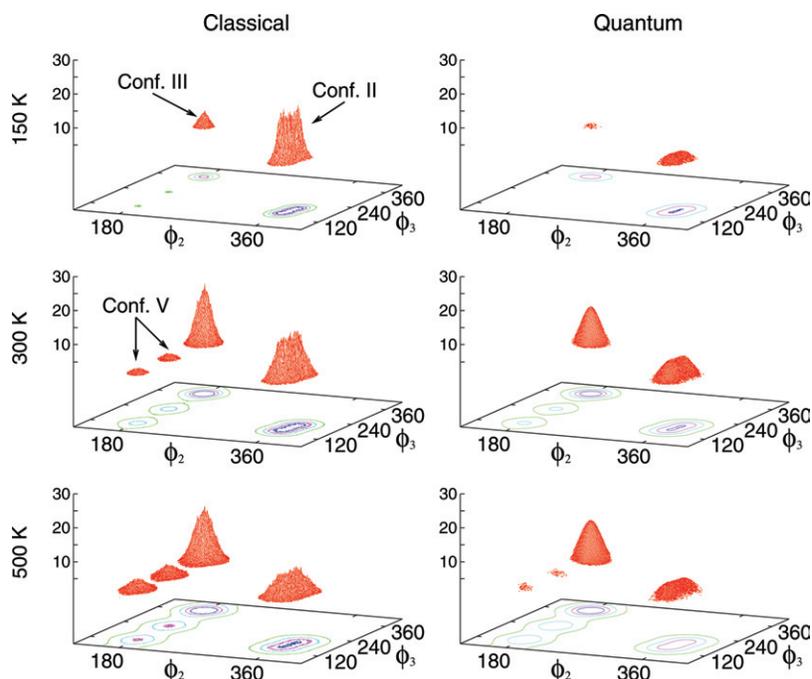


Fig. 7 Torsional population distributions for glycine conformers II, III, and V calculated using quantum and classical TPIMC at various temperatures and plotted as a function of the C–N torsional angle ϕ_2 and C–O torsional angle ϕ_3 .

As was noted before in Fig. 6, the role of quantum mechanics is particularly striking for conformer II. At each temperature, the quantum distribution of conformer II about $(\phi_2, \phi_3) = (360^\circ, 180^\circ)$ is diminished and broadened in comparison to the corresponding classical distribution. These two-dimensional plots offer further insight into the reason for the sensitivity of conformer II to quantum effects. In the classical plots, the distribution of conformer II exhibits two slight peaks corresponding to symmetric, non-planar minima of the conformer at $\phi_3 \approx 160^\circ$ and 200° . In the quantum distributions, this feature is largely washed out as the light $-\text{NH}_2$ torsion tunnels between the neighboring minima, resulting in a single, broad distribution peak for conformer II at $\phi_3 = 180^\circ$. This tunneling effect gives rise to the notably quantum behavior of conformer II.

3.4 Non-torsional modes

The TPIMC technique enables the efficient quantum mechanical analysis of large molecules by reducing the total number of intramolecular degrees of freedom to the number of single-bond torsions.^{30,31} By not explicitly including the thermodynamic contributions of non-torsional motions, such as bond stretches and angle bends, the method assumes that these contributions are identical for every conformer and cancel in the evaluation of relative quantities.

The motivation for this assumption is best understood within the context of the harmonic oscillator approximation. For the quantum harmonic oscillator, the contribution of a particular vibrational mode to the free energy of a conformer is composed of two terms. The first is the zero point energy, which is independent of temperature and proportional to the frequency of the vibrational mode. The second term vanishes exponentially for large frequencies and becomes infinite as the frequency approaches zero. It follows that high-frequency modes will make free energy contributions approximately equal to the zero point energy, and low-frequency modes will make potentially huge contributions to the free energy arising from the second term.

We can now consider the relative contribution of a particular vibrational mode to the free energy of two different

conformers. Although high-frequency vibrations, such as bond stretches and most angle bends, will certainly vary depending on conformation, the relative free energy contribution is only proportional to the size of this frequency shift. A frequency shift of the same magnitude in the low-frequency range can lead to much larger relative contributions. It is therefore appealing to assume that relative conformer free energies are dominated by the contributions from low-frequency modes.

Comparison of TPIMC calculations with experiment has supported the assumption that the torsional degrees of freedom dominate conformational thermodynamics. Quantitative agreement between calculated and observed conformer populations is reported above for glycine and in a previous study for the 2-amino-1-phenylethanol molecule.³² However, it is important to recognize that, in addition to torsional modes, flexible molecules exhibit other low frequency motions that can be conformer dependent and substantially affect the relative free energy. It is therefore worthwhile to estimate the influence of non-torsional motions on the glycine conformer populations calculations reported in this study.

A straightforward algorithm for estimating the effect of non-torsional vibrations on the calculated conformer populations involves (i) calculating the harmonic frequencies of the conformers of interest, (ii) identifying the vibrational motions that most resemble torsions, (iii) using the remainder of the vibrational frequencies to calculate harmonic free energy corrections, and (iv) using these free energy corrections to adjust the conformer populations. The quantitative reliability of these corrections depends on the extent to which the torsional modes couple with other vibrations and the accuracy of the harmonic oscillator approximation for the non-torsional motions. Nonetheless, comparison of the calculated populations before and after correction will indicate the extent to which non-torsional vibrations are relevant in the current study.

The left side of Table 4 compares the quantum TPIMC conformer populations before and after non-torsional correction at the experimentally relevant temperature of 400 K. Harmonic frequencies for the five conformers used in the corrections were calculated at the MP2/6-311++G** level of theory and are listed in the ESI.† The close agreement between the second and third columns of Table 4 indicates that the free energy contributions from the non-torsional modes largely cancel in the various conformers of glycine at 400 K. The non-torsional harmonic correction generally changes the calculated TPIMC populations 1–2% at 400 K. It is concluded that the above comparison of quantum TPIMC populations with experiment is not significantly affected by contributions arising from non-torsional motions.

The harmonic vibrational frequencies of glycine, including the torsional modes, can also be used to directly calculate conformer populations. This is a standard electronic structure

theory approach for analyzing intramolecular thermodynamics. Harmonic vibrational analysis approximates the potential energy surface about a local minimum with a quadratic function. To avoid large errors in the evaluation of conformer populations, the technique requires *a priori* specification of each conformer that will be populated and the statistical degeneracy number g_i of each conformer. Conformer I, which has a planar energy minimum, is singly degenerate ($g_I = 1$). The symmetric, nonplanar minima of conformers IV and V are well separated on the potential energy surface, so we can reliably assign $g_{IV} = 2$ and $g_V = 2$. The appropriate degeneracy numbers of conformers II and III, however, are less obvious. It was shown in Fig. 3b that the torsional distribution of conformer III has two maxima at low temperature which converge toward a single maximum at high temperature. Similarly, despite the fact that the quantum torsional distribution of conformer II exhibits only a single peak (Fig. 7), the conformer has a symmetric pair of minimum energy structures at $\phi = (\pm 12^\circ, \mp 3^\circ, \pm 161^\circ)$. It is therefore difficult to make an *a priori* selection of appropriate degeneracy numbers for conformers II and III in the harmonic vibrational thermodynamic analysis of glycine.

The right side of Table 4 presents the glycine conformer populations calculated using three different sets of degeneracy numbers. In each set, it is assumed that $g_I = 1$ and $g_{IV} = g_V = 2$. In the fourth column of the table (HO_1), we set $g_{II} = g_{III} = 1$. For HO_2 , we set $g_{II} = 1$ and $g_{III} = 2$, and for HO_3 , we set $g_{II} = g_{III} = 2$.

In qualitative agreement with the TPIMC results, the HO_X columns in Table 4 predict the majority of molecules to be found in conformer I and smaller populations for conformers II–V. The HO_2 degeneracy number choice of $g_{II} = 1$ and $g_{III} = 2$ recovers the relative ordering of conformer populations, a finding that is consistent with the numbers of maxima in the TPIMC ϕ_1 torsional distributions (Fig. 3). However, the harmonic approximation can severely hinder the calculation of conformer populations,³² and it is seen in the table that the results vary depending on the choice of degeneracy numbers. Without prior knowledge of the TPIMC results, it would have been difficult in this case to reliably compare the harmonic approximation calculations with experiment.

4 Conclusions

By combining an anharmonic, quantum-mechanical description of the glycine torsions with an *ab initio* description of the electronic structure of the molecule, the current study achieves a level of accuracy not previously available in the theoretical study of molecular conformation. The population of the three experimentally observed conformers are calculated in good agreement with recent matrix-IR experiments, and the presence of a fourth conformer is predicted at equilibrium. Contrary to previous conclusions, it is suggested that the absence of conformer IV from experimental observation is not primarily due to the height of the $IV \rightarrow I$ potential barrier but rather to other aspects of the dynamics of interconversion. Quantum effects are found to substantially alter the population of conformer II, especially at low temperature, and the role of the non-torsional vibrations of glycine is not found to significantly alter the comparison of the TPIMC calculations with experiment.

To achieve optimal accuracy, the study employs a high-quality potential energy surface, accounts for the contributions of non-torsional modes, and utilizes conservative parameters for numerical convergence criteria and TPIMC quantum convergence. These measures have led to calculated results that are in strong agreement with available experimental data, but they introduce considerable computational cost. The longest TPIMC calculations presented took approximately eight days

Table 4 Glycine conformer populations at 400 K calculated using Q-TPIMC, Q-TPIMC with non-torsional correction, and the harmonic approximation^a

	Q-TPIMC ^b	Q-TPIMC/ HO^c	HO_1^d	HO_2^d	HO_3^d
I	50.1	48.4	63.0	58.7	53.7
II	12.4	14.2	10.0	9.3	17.0
III	14.8	15.6	7.4	13.7	12.6
IV	16.7	15.6	15.0	14.0	12.8
V	6.0	6.2	4.6	4.3	3.9

^a Conformer populations reported in percentage of total population.

^b Quantum TPIMC conformer populations, as shown in Fig. 2.

^c Quantum TPIMC conformer populations corrected with the harmonic approximation of non-torsional vibrational modes described in the text.

^d Harmonic vibrational analysis with $g_{II} = g_{III} = 1$, with $g_{II} = 1$ and $g_{III} = 2$, and with $g_{II} = g_{III} = 2$, respectively.

to complete on a single processor, and the construction of a global *ab initio* potential energy surface would soon become unfeasible with more degrees of freedom. Nonetheless, previous studies show that the TPIMC technique, when used with a molecular mechanics potential, can be affordably applied to systems of at least 75 atoms³¹ and that the approach can quantitatively reproduce experimental findings.³² Furthermore, the quantum anharmonic description provided by the TPIMC technique will be critical in future applications that consider the explicit interaction of a biomolecule with its solvent. When employed with a reliable potential energy surface, the TPIMC technique is a promising new method for the prediction of biomolecular conformation that is expected to gain future application on more complicated systems of biological interest.

Supporting information available

Tables listing the calculated harmonic frequencies of the glycine conformers.

Acknowledgements

T.F.M. and D.C.C acknowledge fellowships from the National Science Foundation and the Leverhulme Trust, respectively. The work was supported by the Oxford Supercomputing Centre and the Engineering and Physical Sciences Research Council. The authors appreciate the helpful comments of Dr. M. Mella, Dr. P. D. Godfrey, Prof. S. G. Stepanian, and Prof. J. P. Simons and his research group.

References

- R. J. Lavrich, C. R. Torok and M. J. Tubergen, *J. Phys. Chem. A*, 2002, **106**, 8013.
- E. G. Robertson and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2001, **3**, 1.
- L. F. Pacios, O. Gálvez and P. C. Gómez, *J. Phys. Chem. A*, 2001, **105**, 5232.
- W. Wang, X. Pu, W. Zheng, N.-B. Wong and A. Tian, *Chem. Phys. Lett.*, 2003, **370**, 147.
- P. Wolynes, J. N. Onuchic and D. Thirumalai, *Science*, 1995, **267**, 1619.
- C. M. Dobson, A. Sali and M. Karplus, *Angew. Chem., Int. Ed. Engl.*, 1998, **37**, 868.
- G. Junk and H. Svec, *J. Am. Chem. Soc.*, 1963, **85**, 839.
- K. Iijima, K. Tanaka and S. Onuma, *J. Mol. Struct.*, 1991, **246**, 257.
- S. Vishveshwara and J. A. Pople, *J. Am. Chem. Soc.*, 1977, **99**, 2422.
- H. L. Sellers and L. Schäfer, *J. Am. Chem. Soc.*, 1978, **100**, 7728.
- L. Schäfer, H. L. Sellers, F. J. Lovas and R. D. Suenram, *J. Am. Chem. Soc.*, 1980, **102**, 6566.
- M. Ramek, V. K. W. Cheng, R. F. Frey, S. Q. Newton and L. Schäfer, *J. Mol. Struct. (THEOCHEM)*, 1991, **235**, 1.
- J. H. Jensen and M. S. Gordon, *J. Am. Chem. Soc.*, 1991, **113**, 7917.
- R. F. Frey, J. Coffin, S. Q. Newton, M. Ramek, V. K. W. Cheng, F. A. Momany and L. Schäfer, *J. Am. Chem. Soc.*, 1992, **114**, 5369.
- A. G. Császár, *J. Am. Chem. Soc.*, 1992, **114**, 9568.
- C. H. Hu, M. Shen and H. F. Schaefer, III, *J. Am. Chem. Soc.*, 1993, **115**, 2923.
- F. Barone, C. Adamo and F. Leij, *J. Chem. Phys.*, 1995, **102**, 364.
- A. G. Császár, *J. Mol. Struct.*, 1995, **346**, 141.
- R. D. Brown, P. D. Godfrey, J. W. V. Storey and M. P. Bassez, *J. Chem. Soc. Chem. Commun.*, 1978, 547.
- R. D. Suenram and F. J. Lovas, *J. Mol. Spec.*, 1978, **72**, 372.
- R. D. Suenram and F. J. Lovas, *J. Am. Chem. Soc.*, 1980, **102**, 7810.
- P. D. Godfrey and R. D. Brown, *J. Am. Chem. Soc.*, 1995, **117**, 2019.
- P. D. Godfrey, R. D. Brown and F. M. Rodgers, *J. Mol. Struct.*, 1996, **376**, 65.
- O. Bludský, J. Chocholoušová, J. Vacek, F. Huisken and P. Hobza, *J. Chem. Phys.*, 2000, **113**, 4629.
- R. S. Ruoff, T. D. Klots, T. Emilsson and H. S. Gutowsky, *J. Chem. Phys.*, 1990, **93**, 3142.
- I. D. Reva, A. M. Plokhotnichenko, S. G. Stepanian, A. Y. Ivanov, E. D. Radchenko, G. G. Sheina and Y. P. Blagoi, *Chem. Phys. Lett.*, 1995, **232**, 141.
- A. Y. Ivanov, A. M. Plokhotnichenko, V. Izvekov, G. G. Sheina and Y. P. Blagoi, *J. Mol. Struct.*, 1997, **408/409**, 459.
- S. G. Stepanian, I. D. Reva, E. D. Radchenko, M. T. S. Rosado, M. L. T. S. Duarte, R. Fausto and L. Adamowicz, *J. Phys. Chem. A*, 1998, **102**, 1041.
- A. Y. Ivanov, G. G. Sheina and Y. P. Blagoi, *Spectrochim. Acta*, 1999, **55**, 219.
- T. F. Miller, III and D. C. Clary, *J. Chem. Phys.*, 2002, **116**, 8262.
- T. F. Miller, III and D. C. Clary, *J. Chem. Phys.*, 2003, **119**, 68.
- T. F. Miller, III and D. C. Clary, *J. Phys. Chem. B*, 2004 in press.
- R. P. Feynman, *Statistical Mechanics*, Benjamin, New York, 1972.
- D. Chandler and P. G. Wolynes, *J. Chem. Phys.*, 1981, **74**, 4078.
- B. J. Berne and D. Thirumalai, *Quantum Simulations of Condensed Matter Phenomena*, World Scientific, Singapore, 1990.
- S. L. Mielke and D. G. Truhlar, *J. Chem. Phys.*, 2001, **115**, 652.
- K. R. Glaesemann and L. E. Fried, *J. Chem. Phys.*, 2003, **118**, 1596.
- J. W. Ponder, *TINKER Software Tools for Molecular Design, Version 3.8*, Washington University in St Louis, St :Louis, MO, 2000.
- C. Möller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618.
- R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, **72**, 5639.
- G. M. Chaban, J. O. Jung and R. B. Gerber, *J. Phys. Chem. A*, 2000, **104**, 10 035.
- V. Alexandrov, S. Stepanian and L. Adamowicz, *Chem. Phys. Lett.*, 1998, **291**, 110.
- M. J. Frisch, *et al.*, *Gaussian 98, Revision A.11.1*, Gaussian, Inc., Pittsburgh, PA, 2001.
- D. A. McQuarrie, *Statistical Mechanics*, Harper and Row, New York, 1976.
- S. J. McGlone, P. S. Elmes, R. D. Brown and P. D. Godfrey, *J. Mol. Struct.*, 1999, **485/486**, 225.
- F. Huisken, O. Werhahn, A. Y. Ivanov and S. A. Krasnokutski, *J. Chem. Phys.*, 1999, **111**, 2978.
- J. Gao and D. G. Truhlar, *Annu. Rev. Phys. Chem.*, 2002, **53**, 467.
- D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, Oxford, 1987.
- D. W. Zhang, M. L. Wang and J. Z. H. Zhang, *J. Chem. Phys.*, 2003, **118**, 2716.
- D. W. Zhang, M. L. Wang and J. Z. H. Zhang, *J. Phys. Chem. A*, 2004, in press.
- P. D. Godfrey, personal communication.