

Conformers of Gaseous Cysteine

Jeremiah J. Wilke,[†] Maria C. Lind,[†] Henry F. Schaefer III,[†] Attila G. Császár,^{*,‡} and Wesley D. Allen^{*,†}

Department of Chemistry and Center for Computational Chemistry, University of Georgia, Athens, Georgia 30602, and Laboratory of Molecular Spectroscopy, Institute of Chemistry, Eötvös University, H-1518 Budapest 112, P.O. Box 32, Hungary

Received January 3, 2009

Abstract: Structures, accurate relative energies, equilibrium and vibrationally averaged rotational constants, quartic and sextic centrifugal distortion constants, dipole moments, ¹⁴N nuclear quadrupole coupling constants, anharmonic vibrational frequencies, and double-harmonic infrared intensities have been determined from ab initio electronic structure computations for conformers of the neutral form of the natural amino acid L-cysteine (Cys). A systematic scan located 71 unique conformers of Cys using the MP2(FC)/cc-pVTZ method. The large number of structurally diverse low-energy conformers of Cys necessitates the highest possible levels of electronic structure theory to determine their relative energies with some certainty. For this reason, we determined the relative energies of the lowest-energy eleven conformers, accurate within a standard error (1σ) of about 0.3 kJ mol⁻¹, through first-principles composite focal-point analyses (FPA), which employed extrapolations using basis sets as large as aug-cc-pV(5+d)Z and correlation treatments as extensive as CCSD(T). Three and eleven conformers of L-cysteine fall within a relative energy of 6 and 10 kJ mol⁻¹, respectively. The vibrationally averaged rotational constants computed in this study agree well with Fourier-transform microwave spectroscopy results. The effects determining the relative energies of the low-energy conformers of cysteine are analyzed in detail on the basis of hydrogen bond additivity schemes and natural bond orbital analysis.

1. Introduction

Because amino acids are the building blocks of proteins and peptides, the structural investigation of them, extending from solids to the gas phase, has received considerable experimental and theoretical attention.¹ Cysteine (Cys) is the only amino acid with a reactive sulfur moiety. In this regard, cysteine contributes to diverse structures, including disulfide bonds, zinc fingers, and Fe–S coordination in iron–sulfur proteins.² Functionally, disulfide bonds formed from cysteine serve a central role in glutathione, a mediator of oxidative stress, and strong nucleophilicity also makes cysteine a key component of the active site in many other enzymes.^{3,4}

In the gas phase, amino acids are intrinsically flexible systems, occurring in their neutral form and exhibiting a large number of low-energy conformers. Even glycine, with only three rotatable single bonds, has eight conformers, five of which have relative energies less than 12 kJ mol⁻¹.^{1,5,6} The number of natural amino acids is limited, and the relatively small size of these molecules allows the application of highly sophisticated quantum chemical methods to study their equilibrium and dynamical structures and rotational–vibrational spectra. The number of local minima on the respective potential energy surfaces (PES) and the structural properties of the related conformers, including accurate relative energies, are available for a number of amino acids. A review summarizing results before 1999 is provided in ref 1. Given the accuracy of modern electronic structure techniques, characterization of the complex PESs of amino acids should precede and supplement related experimental

* Corresponding author e-mail: wdallen@uga.edu (W.D.A.); csaszar@chem.elte.hu (A.G.C.).

[†] University of Georgia.

[‡] Eötvös University.

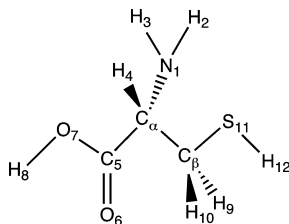


Figure 1. Labeling scheme for cysteine.

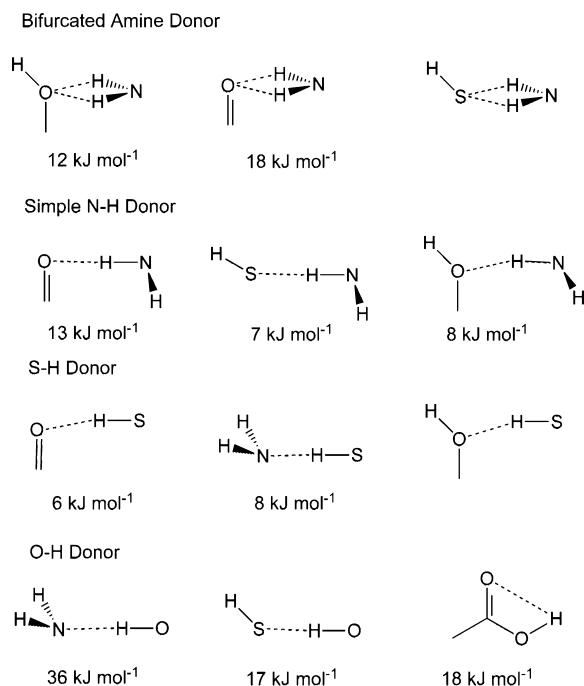


Figure 2. Common hydrogen bond motifs and approximate interaction strengths, where known.^{15,22} Approximate relative energies of conformers can be treated additively as the difference between the sum of near-atom interactions.

structural and spectroscopic studies. The most accurate equilibrium structures (in cases both Born–Oppenheimer and semiexperimental ones) and relative energies (obtained within the focal-point analysis (FPA) approach^{7–12}) are available for the amino acids glycine (Gly),^{5,13,14} alanine (Ala),^{13,15} threonine (Thr),¹⁶ and proline (Pro).^{17,18} The present high-level computational study expands the list of structurally well characterized amino acids by careful investigation of all of the low-energy conformers of L-cysteine.

Cysteine is a good representative of those amino acids with the added complexity of a side chain capable of substantial hydrogen-bonding interactions. Substitution of an -SH group for one of the H atoms of the methyl side chain of Ala introduces two new rotators of significance, those around the C_α–C_β and C_β–S bonds (Figure 1). A dramatic increase in the number of possible conformers results. The presence of three H-bond donors and four H-bond acceptors in Cys allows for the existence of twelve distinct types of hydrogen bonds (Figure 2), including (a) bifurcated H-bonds between -NH₂ and -COOH, similar to those found in the most stable conformers of Gly⁵ and Ala;¹⁵ (b) simple N–H donor bonds, like the adducts to side chain S–H and to carboxylic acid O–H and C=O; (c) side chain S–H interactions with nitrogen and the carbonyl or hydroxyl oxygen; and (d) O–H

donations to C=O and to side chain S–H and NH₂. While H-bonds are certainly the main secondary interactions determining the occurrence and relative energies of the conformers of Cys, other structural factors are also important. These include exchange, electrostatic, and hyperconjugative electronic effects, as well as steric and dispersive interactions. Detailed investigation of these structural factors is one of the main goals of this study. In principle, Cys could contain the same conformers as serine (Ser), its OH analogue. However, since the interactions in Cys are weaker than in Ser (the thiol group of the side chain has comparatively poorer H-bonding characteristics), the barriers separating the conformers are expected to be smaller and in some instances may even disappear. In such a case, Cys would exhibit fewer unique conformers than Ser.

Previous ab initio electronic structure computations performed on Cys include studies on its conformational behavior¹⁹ and on its various physical properties, including proton affinities and ionization potentials.²⁰ Schäfer et al.²¹ investigated 10 conformers of Cys at the RHF/4–21G level and established conformational trends. Gronert and O’Hair²² located 42 conformers at several levels of ab initio electronic structure theory, including RHF/6–31G* and MP2/6–31+G*.²³ Recently, Dobrowolski and co-workers²⁴ located 51 conformers using the B3LYP and MP2 methods in conjunction with the aug-cc-pVDZ (and in some cases aug-cc-pVTZ) basis set. The computed B3LYP/aug-cc-pVDZ frequencies were compared to IR matrix isolation spectra, suggesting the presence of between three and six L-cysteine conformers in the experiments.

The conformers of cysteine investigated previously range in relative energy by at most 50 kJ mol⁻¹. The six most stable conformers of Cys lie within 7 kJ mol⁻¹, while altogether 33 conformers have been identified within a 17 kJ mol⁻¹ range. Alonso et al.²⁵ recently identified five conformers within 10 kJ mol⁻¹ using laser ablation and Fourier-transform microwave spectroscopy (FTMW). It is clear that the highest possible levels of electronic structure theory must be employed to obtain definitive energetics for these structurally diverse but energetically similar conformers.

The present study yielded, as primary information, accurate equilibrium structures and relative energies, as well as copious spectroscopic molecular parameters related to the vibrational and rotational spectra of the most important conformers of Cys. In turn, the large number of computed molecular properties allowed the investigation of a number of interesting computational issues. These include systematic errors in the geometries, bracketing the errors in relative energies for different levels of theory, anharmonicity and zero-point vibrational corrections, and the electron correlation effects in properties such as quadrupole coupling constants.

2. Computational Details

Most of the atom-centered Gaussian basis sets selected for the electronic structure computations of this study contain both polarization and diffuse functions, both of which are needed for the determination of accurate structures and relative energies of H-bonded systems.²⁶ The subcompact 3-21G^{27,28} basis lacks these functions, and thus it was used

only for prescreening the conformers at the Hartree–Fock²⁹ level of theory. The correlation-consistent, polarized-valence (aug)-cc-p(C)V($n+d$)Z, $n = 2$ (D), 3 (T), 4 (Q) basis sets of Dunning and co-workers^{30–34} were employed extensively for geometry optimizations and single-point energy computations within the FPA approach.^{7–10} The augmented (aug) basis sets contain diffuse functions, while tight functions necessary for treating core correlation are contained in the core-polarized (C) basis sets. In addition, the “+d” notation indicates a set of tight d-functions for second-row atoms to rectify problems with the originally designed correlation-consistent sets and thus smooth basis set extrapolations for sulfur-containing molecules. For Cys, the aug-cc-pV(D+d)Z, aug-cc-pV(T+d)Z, aug-cc-pV(Q+d)Z, and aug-cc-pV(5+d)Z basis sets contain 233, 492, 891, and 1458 CGFs, respectively. Only pure spherical harmonics were employed in all basis sets used in this study.

Electronic wave functions were determined in this study by the single-configuration, self-consistent-field, restricted Hartree–Fock (RHF) method,^{29,35,36} by second-order Møller–Plesset perturbation theory (MP2),²³ and by coupled cluster (CC) methods,^{37,38} including all single and double excitations (CCSD),³⁹ as well as a perturbative correction for connected triple excitations [CCSD(T)].⁴⁰ In addition, energies and geometries were determined using the hybrid density functional B3LYP.^{41–43} Both the T_1 diagnostics of coupled cluster theory^{44,45} (~ 0.014) and qualitative bonding principles indicate that the conformers of Cys are well described by single-reference correlation methods. The seven lowest 1s-like orbitals along with the sulfur 2s and 2p orbitals were kept as frozen core (FC) in all post-Hartree–Fock treatments.

The electronic structure packages MAB-ACESII,⁴⁶ MPQC,^{47–49} MOLPRO,⁵⁰ and Gaussian03⁵¹ were used extensively in this study.

2.1. Geometry Optimizations. Initial structures for the geometry optimizations of the conformers of Cys were found by systematically varying the six most important dihedral angles (see Figure 1). The thiol carbon, $\tau(\text{S}_{11}-\text{C}_\beta-\text{C}_\alpha-\text{C}_5)$, and amine group, $\tau(\text{H}_3-\text{N}_1-\text{C}_\alpha-\text{C}_\beta)$, were rotated in 30° increments, while the carboxylic acid plane, $\tau(\text{O}_6-\text{O}_7-\text{C}_5-\text{C}_\alpha)$, thiol hydrogen, $\tau(\text{H}_{12}-\text{S}_{11}-\text{C}_\beta-\text{C}_\alpha)$, carboxyl hydrogen, $\tau(\text{H}_8-\text{O}_7-\text{C}_5-\text{O}_6)$, and $\text{C}_\alpha-\text{C}_\beta$ bond, $\tau(\text{S}_{11}-\text{C}_\beta-\text{C}_\alpha-\text{N}_1)$, were rotated in 120° increments, resulting in a preliminary set of 11 664 starting structures. The initial geometries were optimized at the HF/3-21G level until the Cartesian displacements between optimization steps were less than 10^{-4} bohr. Redundant conformers were identified by checking that energies and geometries were equivalent within a given threshold. Energies were considered to be the same if they were within $10^{-7} E_h$, while bond lengths and angles were required to be within 0.001 Å and 1.0° , respectively. In total, a set of 90 unique HF/3-21G conformers were found.

The HF/3-21G structures were reoptimized at the frozen-core MP2/cc-pVTZ level. When MP2(FC)/cc-pVTZ geometry optimizations were performed, some of the higher-energy HF/3-21G conformers disappeared, yielding a final set of 71 conformers for Cys, according to the same uniqueness criteria given above. The eleven MP2/cc-pVTZ conformers within 10.0 kJ mol^{-1} of the lowest-energy

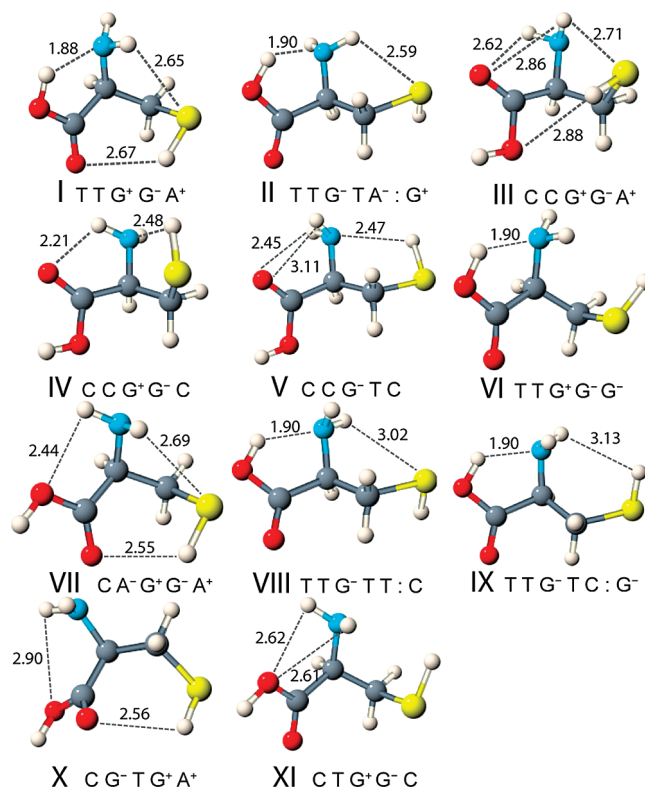


Figure 3. Pictorial representation of the eleven lowest-energy conformers of L-cysteine (Cys). See Figure 1 for numbering. The labeling scheme identifies conformers based on the series of dihedral angles $\tau(\text{H}_8-\text{O}_7-\text{C}_5-\text{O}_6)$, $\tau(\text{N}_1-\text{C}_\alpha-\text{C}_5-\text{O}_6)$, $\tau(\text{N}_1-\text{C}_\alpha-\text{C}_\beta-\text{S}_{11})$, $\tau(\text{C}_5-\text{C}_\alpha-\text{C}_\beta-\text{S}_{11})$, and $\tau(\text{N}_1-\text{C}_\alpha-\text{S}_{11}-\text{H}_6)$. A sixth identifier can be added for the “dihedral angle” of the nitrogen lone pair relative to the $\text{C}_\alpha-\text{C}_5$ bond in cases of ambiguity (e.g., II and VIII).

structure were chosen for a more detailed analysis. These structures were reoptimized at the frozen-core MP2/aug-cc-pV(T+d)Z level. Detailed information, including Cartesian coordinates and energies of all the conformers found in this study, is provided as Supporting Information. Following a scheme first employed for glycine,⁵ the conformers are numbered by Roman numerals (see Figure 3), reflecting the energy ordering determined at the MP2(FC)/aug-cc-pV(T+d)Z level. Similarly to Dobrowolski et al.,²⁴ we choose a series of dihedral angles to uniquely identify conformers, assigning each angle as C (cis, $-30^\circ < \tau < +30^\circ$), T (trans, $150^\circ < \tau < 210^\circ$), G⁺ (gauche, $+30^\circ < \tau < +90^\circ$), G[−] (gauche, $-90^\circ < \tau < -30^\circ$), A⁺ (antiperiplanar, $+90^\circ < \tau < +150^\circ$), or A[−] (antiperiplanar, $-150^\circ < \tau < -90^\circ$). This notation is equivalent to the Klyne–Prelog specification,⁵² but we use the terms cis, gauche, and trans instead of synperiplanar, synclinal, and antiperiplanar, respectively. In particular, we identify the conformers via the dihedral angles $\tau(\text{H}_8-\text{O}_7-\text{C}_5-\text{O}_6)$, $\tau(\text{N}_1-\text{C}_\alpha-\text{C}_5-\text{O}_6)$, $\tau(\text{S}_{11}-\text{C}_\beta-\text{C}_\alpha-\text{N}_1)$, $\tau(\text{S}_{11}-\text{C}_\beta-\text{C}_\alpha-\text{C}_5)$, and $\tau(\text{N}_1-\text{C}_\alpha-\text{S}_{11}-\text{H}_{12})$. $\tau(\text{H}_8-\text{O}_7-\text{C}_5-\text{O}_6)$ specifies the carboxyl group as cis or trans, $\tau(\text{N}_1-\text{C}_\alpha-\text{C}_5-\text{O}_6)$ identifies the type of carboxyl-amine hydrogen bond, $\tau(\text{S}_{11}-\text{C}_\beta-\text{C}_\alpha-\text{N}_1)$ and $\tau(\text{S}_{11}-\text{C}_\beta-\text{C}_\alpha-\text{C}_5)$ define the backbone of the molecule, and $\tau(\text{N}_1-\text{C}_\alpha-\text{S}_{11}-\text{H}_{12})$ gives the orientation of the thiol hydrogen relative to the amine. For some conformers, for example, Cys-II and Cys-VIII, the

orientation of the amine lone pair relative to the C₅–C_α bond must also be specified to uniquely identify the conformer.

The 71 conformers located here exceed those of Dobrowolski et al.,²⁴ and Gronert and O'Hair,²² who observed only 51 and 42 distinct Cys conformers, respectively. Their initial search tested 324 starting geometries at the AM1 level, resulting in a preliminary set of only 58 conformers. In contrast, our initial search was performed at the HF/3-21G level on an initial set of over 10 000 starting geometries, yielding a preliminary set of 90 conformers. Because it is highly likely that the 71 distinct MP2(FC)/cc-pVTZ conformers located here exist in reality, it seems that the 20 conformers were missed in previous work due to the AM1 preoptimization step.

2.2. Focal-Point Analysis (FPA). To obtain relative energies as accurately as possible, the focal point analysis (FPA) approach^{7–12} was utilized. The eleven lowest-energy structures of Cys (Figure 3), obtained at the MP2(FC)/aug-cc-pV(T+d)Z level, were included in the FPA investigation. Extrapolation of the energies to the complete basis set (CBS) limit at the RHF and MP2 levels was performed, as part of the FPA approach. For HF, the total energy was extrapolated using the formula $E_n = E_{\text{CBS}} + A \exp(-Bn)^{53–55}$ with $n \in \{3, 4, 5\}$, where A and B are adjustable parameters, E_n is the RHF total energy for a correlation-consistent basis set aug-cc-pV(n +d)Z, and E_{CBS} is the Hartree–Fock limit. For MP2, the correlation energies (ϵ_n) were extrapolated according to $\epsilon_n = \epsilon_{\text{CBS}} + Bn^{-3}$.⁵⁶ Coupled-cluster energy corrections were treated additively because the electron correlation contributions to the conformational energies at higher levels of theory do not change significantly with the size of the basis set.

For the auxiliary corrections normally included within the FPA approach, the core correlation term was obtained at the MP2/aug-cc-pCVTZ level, while the relativistic⁵⁷ and diagonal Born–Oppenheimer corrections (DBOC)⁵⁸ were deemed negligible. In similar FPA studies performed on the conformers of proline¹⁸ and threonine,¹⁶ the relativistic corrections to the relative conformational energies turned out to be minuscule.

2.3. Spectroscopic Parameters. For the eleven lowest-energy conformers of Cys, quadratic force constants were computed at the B3LYP/aug-cc-pVTZ level using geometries optimized at the same level. In addition, quartic force fields in the normal coordinate space were determined at the B3LYP/6-31G* level. No scaling of the force fields or of the resulting vibrational frequencies was attempted. The use of fully optimized reference geometries during the force field determinations helps to avoid the nonzero force dilemma.⁵⁹ Anharmonic, fundamental vibrational frequencies of the Cys conformers were computed by applying the VPT2 formalism^{60–63} to the quartic force fields. Whenever a Fermi resonance appeared, the corresponding contribution to the fundamental frequency was evaluated by eliminating the associated terms in the expression for the anharmonic constants and then explicitly diagonalizing the 2×2 Hamiltonian matrix for the resonating states.

The optimized structures obtained at the MP2(FC)/aug-cc-pVTZ level determine the equilibrium rotational constants, while the quadratic and cubic force fields from B3LYP/6-

31G* yield the quartic and sextic centrifugal distortion constants in the A -reduced representation, respectively. In addition, quadrupole coupling constants are reported at the MP2 level using a “locally dense” basis set composed of the standard cc-pVTZ functions on carbon, sulfur, oxygen, and hydrogen and the cc-pCV5Z functions on nitrogen. The quadrupole coupling constant is determined by the electric field gradient at nitrogen so that the combination of a locally dense basis and MP2 generally produces good results (see Supporting Information).

2.4. Zero-Point Vibrational Corrections. Zero-point vibrational energies (ZPVE) were obtained as harmonic and anharmonic values at the B3LYP/aug-cc-pVTZ and B3LYP/6-31G* levels, respectively. Some ambiguity exists in the computation of zero-point energies based on anharmonic force fields. For molecules with many normal modes, resonances occur in the energy denominators of the related second-order vibrational perturbation theory (VPT2)^{60–63} expressions. Here, we employ the approach of Allen et al.,⁶⁴ in which the oft-neglected G_0 term is included to obtain an expression for the zero-point vibrational energy completely devoid of resonance denominators. The working equations are

$$\text{ZPVE} = \frac{1}{2} \sum_i \omega_i - \frac{1}{32} \sum_{ijk} \frac{\phi_{ijk} \phi_{ijk}}{\omega_k} - \frac{1}{48} \sum_{ijk} \frac{\phi_{ijk}^2}{\omega_i + \omega_j + \omega_k} + \frac{1}{32} \sum_{ijj} \phi_{ijj} + Z_{\text{Kinetic}} \quad (1)$$

and

$$Z_{\text{Kinetic}} = -\frac{1}{4} \sum_{\alpha} B_e^{\alpha} \left\{ 1 + \sum_{i>j} (\zeta_{ij}^{\alpha})^2 \frac{B_e^{\alpha}(\omega_i + \omega_j) - (\omega_i - \omega_j)^2}{\omega_i \omega_j} \right\} \quad (2)$$

where the ϕ_{ijk} and ϕ_{ijkl} are cubic and quartic force constants in reduced normal coordinates, B_e^{α} denotes the rotational constant for axis α , and the ζ_{ij}^{α} are Coriolis coupling constants. The total zero-point vibrational energies obtained directly from Gaussian 03 were quite similar to those from eq 1, but contained discrepancies as large as 30 cm^{−1} for some conformers. While both our method and Gaussian 03 include the G_0 term, Gaussian 03 includes anharmonic effects by summing over the average of the harmonic and fundamental frequencies⁶⁵

$$\text{ZPVE} = \frac{1}{4} \sum_i (\omega_i + \nu_i) + G_0 - \frac{1}{4} \sum_i x_{ii} \quad (3)$$

where x_{ii} is the diagonal anharmonic constant for mode i . The anharmonic corrections to the ZPVE in Gaussian 03 therefore avoid resonance denominators by including the explicit 2×2 matrix diagonalization for resonating states (section 2.3), in contrast to our approach which avoids resonance denominators by computing ZPVE directly in terms of cubic and quartic force constants.

Anomalously large VPT2 anharmonic shifts are observed for rotation of the S–H bond in the conformers **Cys-IV**, **Cys-VI**, and **Cys-XI**. In general, for conformers **Cys-IV**, **Cys-VI**, and **Cys-XI**, the thiol hydrogen is near the amino group,

Table 1. Summary of Focal Point Analysis for the Relative Energies of the Eleven Most Stable Conformers of Cysteine^a

	Basis	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
$\Delta E(\text{RHF})^a$	DZ	0.00	3.82	-3.91	-0.09	-5.29	7.62	-2.85	3.23	4.57	-2.75	-0.17
	TZ	0.00	3.67	-4.00	-0.39	-5.40	7.22	-2.34	3.15	4.27	-2.52	-0.13
	QZ	0.00	3.59	-4.02	-0.50	-5.54	7.13	-2.33	2.97	4.06	-2.54	-0.17
	5Z	0.00	3.57	-4.04	-0.54	-5.59	7.11	-2.36	2.94	4.02	-2.57	-0.23
	CBS	0.00	3.56	-4.06	-0.55	-5.61	7.09	-2.39	2.93	4.01	-2.60	-0.25
$\delta[\text{MP2}]^a$	DZ	+0.00	+2.20	+10.67	+8.37	+12.77	+1.80	+11.14	+7.24	+6.90	+12.64	+11.25
	TZ	+0.00	+2.83	+11.43	+8.98	+14.38	+2.07	+11.64	+7.87	+7.50	+14.46	+12.32
	QZ	+0.00	+2.94	+11.71	+9.00	+14.40	+2.28	+11.72	+7.82	+7.49	+14.60	+12.62
	5Z	+0.00	+2.98	+11.74	+8.98	+14.41	+2.32	+11.69	+7.84	+7.52	+14.65	+12.63
	CBS	+0.00	+3.02	+11.78	+8.95	+14.43	+2.37	+11.66	+7.86	+7.55	+14.71	+12.64
$\delta[\text{CCSD}]$	DZ	+0.00	-0.66	-2.90	-1.84	-3.20	-1.03	-2.92	-2.60	-2.56	-3.32	-3.04
	TZ	+0.00	-0.70	-3.05	-2.03	-3.64	-1.00	-3.14	-2.83	-2.76	-3.60	-3.30
$\delta[\text{CCSD(T)}]$	DZ	+0.00	+0.42	+1.88	+1.64	+2.04	+0.30	+1.80	+1.08	+1.01	+1.84	+1.85
	TZ	+0.00	+0.49	+2.06	+1.78	+2.33	+0.32	+1.96	+1.15	+1.07	+2.21	+2.10
$\Delta E(\text{CCSD(T)/CBS})$		0.00	6.37	6.73	8.15	7.51	8.78	8.09	9.11	9.87	10.72	11.19
core correction		+0.00	+0.13	+0.17	+0.04	+0.33	+0.04	+0.17	+0.21	+0.21	+0.33	+0.21
harmonic ZPVE		+0.00	-0.01	-2.19	-2.38	-2.24	-0.71	-1.73	-0.93	-1.00	-1.85	-1.76
anharmonic correction		+0.00	+0.08	+0.08	+0.14	+0.21	-0.55	-0.01	+0.21	+0.05	+0.32	-0.10
$\Delta E(\text{FPA})$		0.00	6.58	4.79	5.95	5.81	7.56	6.52	8.60	9.13	9.52	9.54

^a All values given in kJ mol^{-1} . ΔE denotes a relative energy between conformers. δ denotes an increment or correction to ΔE with respect to the preceding level of theory in the hierarchy $\text{RHF} \rightarrow \text{MP2} \rightarrow \text{CCSD} \rightarrow \text{CCSD(T)}$.

and the thiol rotation is strongly coupled to the NH_2 wag. Because of this coupling, VPT2 breaks down for these modes, resulting in unphysically large anharmonic shifts. In eq 1, we exclude those force constants involving at least one mode for which VPT2 gives an anomalous fundamental frequency.

2.5. Natural Bond Orbital Analysis. The natural bond orbital (NBO) method transforms the molecular orbital picture into a localized description based on the intuitive Lewis structures of molecules.^{66–70} The NBO scheme decomposes the 1-particle density matrix into formally occupied 1-center core orbitals and lone pairs (n_x) and 2-center bonding orbitals (σ_{X-Y} , π_{X-Y}), which correspond naturally to bonds drawn in a Lewis diagram. The localized, Lewis density is transformed to the exact density through delocalizations into formally unoccupied 2-center antibonding orbitals (σ_{X-Y}^* , π_{X-Y}^*) and unoccupied 1-center Rydberg orbitals. These delocalizations can be interpreted physically as energy-stabilizing donor–acceptor interactions between localized orbitals, including conjugation (e.g., $\pi \rightarrow \pi^*$) or hyperconjugation ($\sigma \rightarrow \sigma^*$). Applying perturbation theory gives the leading, second-order energy correction due to these donor–acceptor interactions as

$$E_2 = -\frac{F_{ij}^2}{\epsilon_i - \epsilon_{j^*}} \quad (4)$$

where F_{ij} is the Fock matrix element between occupied orbital i and unoccupied orbital j^* , and ϵ_i and ϵ_{j^*} are the corresponding diagonal Fock matrix elements. Here i and j^* do not represent canonical doubly occupied and virtual molecular orbitals, but rather “almost doubly occupied” and “almost unoccupied” orbitals, respectively. In this regard, the Brillouin condition does not apply so that the matrix elements F_{ij} , while small, are not rigorously zero. Written this way, the E_2 values indicate the most important relaxation from an idealized, local density to the exact density. For cysteine, we can therefore assess the importance of hyperconjugation in conformational transitions through the NBO

formalism. Because E_2 is based on perturbation theory, it may overestimate the absolute magnitude of strong interactions in which the energy denominator is small or the Fock matrix element is large. However, the overall trends should be consistent between conformers.

3. Results and Discussion

3.1. Relative Energies of Conformers. As observed repeatedly for amino acids, the introductory Hartree–Fock level of electronic structure theory, independent of the basis set used, is unable to yield the correct relative energies of the conformers of Cys (Table 1). RHF/CBS theory places six conformers (**Cys-III**, **-IV**, **-V**, **-VII**, **-X**, and **-XI**) lower in energy than the global minimum, **Cys-I**. While in Gly and Ala the inclusion of electron correlation tends to decrease the energy differences between the conformers,^{5,15} in the case of Cys it increases the energy differences in almost all cases, which might be attributed to the relatively weak interactions present in Cys. Because sulfur is much more polarizable than oxygen or carbon, the dipole–induced-dipole and dispersion forces should be generally more important in Cys than in Ala, Ser, or Pro, while S–H hydrogen bonds should be weaker. Accordingly, the MP2 correlation energy destabilizes all the conformers considered relative to **I**, as signified by the positive $\delta[\text{MP2}]$ values in Table 1, which can be as large as 15 kJ mol^{-1} . This observation also serves as a warning that the theoretical results obtained with small basis sets and simple electronic structure methods might change considerably once more rigorous techniques are employed.

Compared to $\delta[\text{MP2}]$, the $\delta[\text{CCSD}]$ and $\delta[\text{CCSD(T)}]$ energy increments are relatively small but can affect relative energies as much as 3.6 and 2.3 kJ mol^{-1} for CCSD and CCSD(T), respectively. Such amounts are clearly substantial when so many conformers are within a window of a few kJ mol^{-1} . Interestingly, but again in line with earlier work on the conformers of Thr¹⁶ and Pro,¹⁸ the relative energies are barely affected by core correlation. Even the largest change is smaller than 0.35 kJ mol^{-1} . In contrast, ZPVE corrections

Table 2. Comparison of Conformational Energies of Cysteine (kJ mol^{-1}) for Different Levels of Theory without Zero-Point Vibrational Correction

	CCSD(T)/CBS ^{a,b}	MP2/aug-cc-pV(T+d)Z ^b	CCSD/aug-cc-pV(T+d)Z ^b	RHF/3-21G ^c	B3LYP/aug-cc-pVTZ ^d
Cys-I	0.00	0.00	0.00	1.26	0.00
Cys-II	6.37	6.49	5.82	11.84	3.93
Cys-III	6.73	7.45	4.35	4.35	7.15
Cys-IV	8.15	8.58	6.57	2.43	8.79
Cys-V	7.51	9.00	5.36	0.00	5.44
Cys-VI	8.78	9.29	8.28	12.68	7.53
Cys-VII	8.09	9.29	6.15	5.23	6.36
Cys-VIII	9.11	11.00	8.16	9.08	6.36
Cys-IX	9.87	11.76	9.00	11.30	7.15
Cys-X	10.72	11.97	8.37	13.05	8.91
Cys-XI	11.19	12.18	8.87	10.04	11.09

^a CCSD(T)/CBS denotes the extrapolated value from the focal point analysis. See Table 1. ^b Computed at the MP2/aug-cc-pV(T+d)Z reference geometries. ^c Computed at the RHF/3-21G reference geometries. ^d Computed at the B3LYP/aug-cc-pVTZ reference geometries.

can affect relative energies on the order of 2 kJ mol^{-1} . Anharmonic corrections to the relative energies were less than 0.21 kJ mol^{-1} for all conformers except **Cys-VI** and **Cys-X**, for which these shifts were -0.55 and $+0.32 \text{ kJ mol}^{-1}$, respectively.

The MP2(FC)/aug-cc-pV(T+d)Z single-point energies happen to be quite accurate (Table 2) because the CCSD and CCSD(T) increments are usually of opposite sign and partially cancel (Table 1). In this regard, many of the MP2 relative energies in Table 2 are closer to the FPA results than their CCSD counterparts. The effect of higher-order correlation from the CCSD(T) perturbative triples is not negligible, however, and the definitive FPA scheme alters the MP2 energy ordering of the conformers. In general, B3LYP performs reasonably well for most conformers but can be in error by as much as 2.7 kJ mol^{-1} , as seen in **Cys-VIII**. While density functional theory can be useful for zero-point corrections and geometry optimizations, obtaining the correct energy ordering of so many conformers in such a small energy range (10 kJ mol^{-1}) clearly requires better accuracy than B3LYP provides. A rigorous energy ordering therefore necessitates correlation treatments as extensive as CCSD(T) and also considerations inherent in the FPA scheme.

The FPA scheme allows errors to be bracketed based on the observed convergence to the basis set and correlation limits. The RHF relative energies are converged to better than 0.05 kJ mol^{-1} with the aug-cc-pV(5+d)Z basis, and thus there should be virtually no basis set error in our final RHF/CBS results. Similarly, the MP2/aug-cc-pV(5+d)Z correlation increments match the extrapolated values within 0.1 kJ mol^{-1} for all conformers, and the associated basis set errors should again be negligible. For the coupled-cluster increments, the aug-cc-pV(T+d)Z result matches the aug-cc-pV(D+d)Z result within 0.4 kJ mol^{-1} . In previous work, the CCSD and CCSD(T) increments are essentially converged with a TZ basis,⁷¹ so that the basis set error in the coupled-cluster values should not be greater than 0.2 kJ mol^{-1} .

Assessing the error caused by the higher-order correlation and zero-point vibrational corrections is more difficult. Corrections due to quadruple and higher excitations are generally an order of magnitude less than CCSD(T) corrections.^{71–78} Since most CCSD(T) corrections here are

on the order of $1\text{--}2 \text{ kJ mol}^{-1}$, the neglect of higher excitations should introduce an error to the cysteine relative energies of at most 0.2 kJ mol^{-1} . Zero-point vibrational corrections are generally insensitive to the level of theory. For example (see Supporting Information, Table S1), even MP2 harmonic zero-point corrections with the Huzinaga–Dunning DZP++ basis⁷⁹ (double- ζ plus polarization and diffuse functions) match the B3LYP/aug-cc-pVTZ values within 0.3 kJ mol^{-1} . Since we have accounted for anharmonicity in the present work, the zero-point error should therefore not be greater than $0.2\text{--}0.3 \text{ kJ mol}^{-1}$. In summary, we estimate a standard error (1σ) of 0.3 kJ mol^{-1} for our predicted conformational energies, corresponding to a 95% confidence interval (2σ) of $\pm 0.6 \text{ kJ mol}^{-1}$. We emphasize that these uncertainties hold only for the relative energies due to cancelation of errors, and the uncertainty in the absolute energies will therefore be much larger.

Two recent studies published relative energies for the lowest-energy conformers of cysteine as summarized in the Supporting Information. All eight conformers considered by Dobrowolski et al.²⁴ were also studied in the current work. Three conformers from Alonso et al.²⁵ were not within 10 kJ mol^{-1} of the global minimum after optimization at the MP2(FC)/cc-pVTZ level, and were therefore not included in our rigorous focal point analyses. All three studies agree in the structure of the two most stable conformers of cysteine, **Cys-I** and **Cys-III** in our notation. The ordering of the other low-lying conformers is similar but not the same in the three studies, and the relative energies vary by as much as 1.3 kJ mol^{-1} for **Cys-III** and 1.9 kJ mol^{-1} for **Cys-IX**.

3.2. Geometric Structures. The sophisticated laser ablation FTMW experiments of Alonso et al.²⁵ yielded rotational constants of several conformers of cysteine, but only for the parent isotopologues. Therefore, the type of refinement on collections of isotopologues which has yielded semiexperimental equilibrium structures for Gly¹⁴ and Pro¹⁷ cannot be executed at present for any of the conformers of Cys. Consequently, one must rely on otherwise rather accurate^{15,17} computed structures when analyzing structural trends among the conformers of Cys.

Two major factors seem to determine the general type of conformation that Cys can assume. First, the thiol, amine, and hydroxyl groups adopt different orientations about the $\text{C}_\alpha\text{--C}_\beta$ bond as either gauche or trans. In the discussion to

follow, unless stated otherwise, gauche and trans identify the orientation about the C_α - C_β bond. Second, the carboxyl group may assume a cis or trans conformation. Depending on these orientations, different hydrogen bonding patterns can form of the types $O-H\cdots N$, $N-H\cdots O=C$, and $N-H\cdots OH$, as clearly seen in Figure 3. For the trans carboxyl, a strong $O-H\cdots N$ interaction can form as found in conformers **Cys-I**, **Cys-II**, **Cys-VI**, **Cys-VIII**, and **Cys-IX**. For the cis carboxyl, bifurcated $N-H\cdots O$ bonds form to either the carbonyl oxygen in **Cys-III**, **Cys-IV**, and **Cys-V** or to the hydroxyl oxygen in **Cys-VII** and **Cys-XI**. The gauche conformers are more sterically crowded than the trans conformers. However, the gauche conformation brings the thiol group closer to the carboxyl group, allowing $S-H\cdots O$ interactions. A ring of hydrogen bonds can therefore form, as in **Cys-I**, **Cys-III**, and **Cys-VII**. For trans conformers, only the amine interacts with the thiol, for example in **Cys-II**.

As observed previously for other amino acids,^{5,15,18} bond lengths and bond angles change little among the low-lying conformers. Most bonds have a standard deviation of less than 0.004 Å, while most bond angles have a standard deviation of less than 2.0°. There are, however, a few notable exceptions. The $C=O$ distance has a standard deviation of 0.007 Å with the largest deviation from the mean being 0.012 Å, occurring in **Cys-III**. The $C=O$ bond in **Cys-III** forms a bifurcated hydrogen bond with the amine group, lengthening the bond and redshifting the carbonyl stretching frequency. The C_α - C_5 bond has a standard deviation of 0.008 Å, with the largest deviation from the mean (0.012 Å) occurring in **Cys-IV**. In general, three strong hyperconjugative interactions are possible for the C_α - C_5 bond with the lone pairs from nitrogen, the hydroxyl oxygen, and the carbonyl oxygen. The hyperconjugation is strongest when the lone pair is trans to the C_α - C_5 bond (see below), and it will lengthen this bond by increasing the antibonding occupation. In **Cys-IV**, both the amine and hydroxyl groups are unfavorably placed for hyperconjugation, and the C_α - C_5 bond distance is only 1.514 Å. In contrast, both the hydroxyl and amine are favorably placed for hyperconjugation in **Cys-I**, lengthening the C_α - C_5 bond to 1.534 Å. Similarly, in **Cys-III**, the amine lone pair is trans to the C_α - C_5 bond, but the hydroxyl lone pair is unfavorably placed cis. The C_α - C_5 bond therefore has an intermediate value of 1.521 Å.

For bond angles, the largest standard deviations are for the $N-C-C$ angles. The standard deviation for $N_1-C_\alpha-C_\beta$ is 2.8° with the largest deviation from the mean being 4.0° for **Cys-VI**. Similarly, the standard deviation for $N_1-C_\alpha-C_5$ is 2.8° with the largest deviation from the mean value being 5.4° for **Cys-XI**. The large spread of $N-C-C$ angles is consistent with the trans angle rule.⁸⁰ In general, in primary alcohols and amines, if a $C-C$ bond is trans to the $X-H$ bond, the $X-C-C$ angle will be smaller, because of both reduced bond repulsion relative to the gauche conformer and weaker hyperconjugation from the nitrogen lone pair. The change in angle depending on bond orientation is clearly evident in cysteine. For conformers **Cys-I**, **Cys-VI**, **Cys-VIII**, and **Cys-IX**, the $N-H$ bonds are gauche to the C_α - C_β bond, and the angles are in the range 115°–117° (see

Supporting Information Table S3). In contrast, for conformers **Cys-II**, **-III**, **-IV**, **-V**, **-VII**, **-X**, and **-XI**, the $N-H$ bond is trans to the C_α - C_β bond, and the $N-C-C$ angles range from 109° to 111°. The same general trends are observed for the $N_1-C_\alpha-C_5$ angles. The angle variations are consistent with strong hyperconjugation from the nitrogen lone pair to the $C-C$ antibonding orbital ($n_N \rightarrow \sigma_{C-C}^*$). Following arguments rationalizing tilting of the methyl group,^{81,82} the $C-C$ axis generally tilts away from the $C-N$ axis to maximize overlap between the nitrogen lone pair and the backside lobe of the $C-C$ antibonding orbital, strengthening the $n_N \rightarrow \sigma_{C-C}^*$ hyperconjugative stabilization. In particular, for both C_α - C_5 and C_α - C_β , large hyperconjugative interactions ($E_2 > 20.0$ kJ mol⁻¹, Supporting Information Table S3) are observed with large $C-C-N$ angles, while weaker interactions ($E_2 < 16.0$ kJ mol⁻¹) are observed with smaller $C-C-N$ angles.

In the same way for the carboxyl group, if the carboxyl group assumes a cis conformation ($O-H$ bond trans to the $C-H$ bond), the $O-C-C$ angle is much smaller. This is observed in conformers **Cys-III**, **Cys-IV**, **Cys-V**, and **Cys-VII**, all of which have $O-C-C$ angles of approximately 111.5°. In contrast, for conformers **Cys-I**, **Cys-II**, **Cys-VI**, and **Cys-VIII** with trans carboxyl ($O-H$ bond cis to the $C-C$ bond), the angles are larger, between 113.5° and 114°. In general, the trans effect seems weaker for the $O-H$ bond than for the $N-H$ bond. The weaker dependence may be attributed to the stronger basicity of the amine and therefore stronger $n \rightarrow \sigma^*$ hyperconjugation. The $O-H\cdots N$ hydrogen bonding also seems to offset the trans effect, closing the $O-C-C$ angle to maximize the $O-H\cdots N$ interaction.

3.3. Structural Effects on Relative Energies. As emphasized previously for alanine¹⁵ and serine,²² approximate values for the strength of certain types of hydrogen bonds can be computed and used to rationalize energy differences among the conformers. Specifically, the energy of each conformer can be approximated as a sum of stabilizations from near-atom interactions. The interaction strengths are then fit through a linear regression to match as closely as possible the conformational energies. Approximate hydrogen bond strengths are given in Figure 2 for the common bonding motifs, as reported in ref 15. In general, the hydrogen bond donors can be ranked in the order $O-H > N-H > S-H$, and the hydrogen bond acceptors can be ranked in the order $N > O > S$. Additionally, an additive approximation can be applied to the conformation of the carboxyl group. Based on the formic acid prototype,¹⁵ the cis carboxyl is intrinsically more stable than the trans carboxyl by approximately 18.5 kJ mol⁻¹ irrespective of hydrogen bonds to other functional groups.

As shown in Figure 2, the $O-H\cdots N$ arrangement is the strongest hydrogen bond, matching the strongest hydrogen bond donor, OH, with the best acceptor, N. We therefore find that the structure of the global minimum, **Cys-I**, is stabilized by a strong $O-H\cdots N$ hydrogen bond between the amino and carboxyl groups. The Cys global minimum is in contrast to that of serine,²² for which the lowest energy conformer is analogous to **Cys-V**, exhibiting a strong $O-H\cdots N$ hydrogen bond with the side chain. The $S-H$ bond in Cys is comparatively a much weaker hydrogen bond

Table 3. Anharmonic Vibrational Fundamentals in cm^{-1} and Double-Harmonic Infrared Relative Intensities (%) of the Eleven Lowest-Energy Conformers of L-Cysteine for Regions Characteristic of Particular Hydrogen Bond Patterns^a

	O–H, N–H stretch			S–H stretch	C=O stretch	O–H bend		C–O stretch	
	ν_1	ν_2	ν_3	ν_7	ν_8	ν_{11}	ν_{12}	ν_{17}	ν_{18}
Cys-I	3397 (4)	3278 (18)	3216 (47)	2542 (0)	1793 (75)	1356 (100)	1353 (2)	1129 (2)	1073 (3)
Cys-II	3383 (6)	3300 (15)	3226 (43)	2543 (0)	1807 (87)	1337 (100)	1337 (8)	1144 (3)	1068 (3)
Cys-III	3539 (22)	3387 (4)	3327 (1)	2551 (0)	1776 (100)	1350 (0)	1314 (3)	1121 (2)	1086 (78)
Cys-IV	3554 (30)	3417 (8)	3374 (5)	2557 (1)	1775 (100)	1385 (4)	1314 (9)	1114 (69)	1084 (22)
Cys-V	3541 (21)	3397 (3)	3324 (1)	2552 (1)	1773 (100)	1380 (2)	1278 (15)	1124 (15)	1084 (58)
Cys-VI	3397 (4)	3378 (54)	3238 (12)	2539 (0)	1798 (84)	1349 (100)	1353 (12)	1130 (4)	1065 (2)
Cys-VII	3545 (28)	3407 (5)	3392 (1)	2548 (0)	1762 (100)	1350 (0)	1330 (8)	1114 (69)	1102 (22)
Cys-VIII	3396 (3)	3305 (10)	3244 (62)	2548 (0)	1794 (100)	1336 (59)	1366 (82)	1120 (3)	1084 (2)
Cys-IX	3410 (2)	3395 (18)	3254 (32)	2525 (0)	1798 (71)	1332 (1)	1362 (100)	1112 (5)	1088 (1)
Cys-X	3538 (25)	3398 (2)	3301 (0)	2560 (0)	1766 (100)	1344 (10)	1354 (0)	1111 (90)	1056 (21)
Cys-XI	3538 (27)	3401 (3)	3315 (0)	2545 (0)	1767 (100)	1359 (2)	1309 (10)	1122 (70)	1102 (18)

^a Harmonic frequencies and intensities were computed at the B3LYP/aug-cc-pVTZ level. Anharmonic corrections were computed using B3LYP/6-31G*. Intensities are reported as a percentage of the most intense peak for a given conformer.

donor than the O–H side chain in serine so that the trans to cis isomerization of the carboxyl is not enough to offset the weaker hydrogen bond. In addition, **Cys-I**, because of its gauche conformation, can form three hydrogen bonds whose charge polarization will cooperatively reinforce each other. We emphasize that our focal point conformational energies should be accurate within a standard error of 0.3 kJ mol^{-1} (1σ) or a 95% confidence interval of $\pm 0.6 \text{ kJ mol}^{-1}$ (2σ). In contrast to previous studies,^{24,25} we can therefore definitively say that the energy differences are real physical effects rather than errors in the underlying computational methods.

In addition to hydrogen bonding, gauche and trans conformations are also affected by steric repulsion and hyperconjugation. The gauche conformers of cysteine have all three bulky substituents in close vicinity, increasing steric repulsion. For butane, the trans–gauche difference is 2.6 kJ mol^{-1} .⁸³ Consequently, steric effects are certainly not negligible in Cys since its eleven lowest-energy conformers lie in an energy range of 10 kJ mol^{-1} . Hyperconjugation is stronger in the gauche configuration since the strongly electronegative groups (amine, carboxyl, thiol) are all anti-periplanar to the electropositive hydrogens. In this regard, the better electron donor orbitals ($\sigma_{\text{C–H}}$) are matched to the better electron acceptors ($\sigma_{\text{C–N}}^*$, $\sigma_{\text{C–S}}^*$). The leading hyperconjugative interactions in cysteine conformers are listed in Table S4, Supporting Information. For example, the $\sigma_{\text{C–H}} \rightarrow \sigma_{\text{C–S}}^*$ interaction is 20.7 kJ mol^{-1} in the gauche **Cys-I**, while the equivalent $\sigma_{\text{C–H}} \rightarrow \sigma_{\text{C–H}}^*$ interaction is only 11.3 kJ mol^{-1} in **Cys-V**, although the effect is offset somewhat by stronger $\sigma_{\text{C–C}} \rightarrow \sigma_{\text{C–S}}^*$ hyperconjugation in **Cys-V** relative to the $\sigma_{\text{C–C}} \rightarrow \sigma_{\text{C–H}}^*$ hyperconjugation in **Cys-I**. This gauche effect has been observed previously for difluorosubstituted hydrocarbons and hydroxyproline.^{84,85} Hyperconjugation and steric effects will therefore tend to offset each other. For some conformers, the relative energies will therefore be dictated mainly by the hydrogen bonding interactions, owing to fortuitous cancelation of competing electronic effects.

The importance of hyperconjugation can be seen in the transformation from **Cys-IV** to **Cys-V**, wherein the amine switches from a simple N–H \cdots O bond to a bifurcated N–H \cdots O bond, while the orientation about the $\text{C}_\alpha\text{–C}_\beta$ bond simultaneously goes from gauche to trans. Assuming simple hydrogen bond additivity, **Cys-V** should lie 5 kJ mol^{-1} below

Cys-IV due to the larger bifurcated hydrogen bond energy (Figure 2). In fact, the energy difference is less than 1 kJ mol^{-1} . The S–H \cdots N bond distances are basically equivalent in **Cys-IV** and **Cys-V** (2.48 \AA versus 2.47 \AA), so the S–H \cdots N interaction should not contribute significantly to the energy difference. The discrepancy seems to lie in hyperconjugative stabilization of **Cys-IV**. While steric repulsion is greater in **Cys-IV**, the much stronger $\sigma_{\text{C–H}} \rightarrow \sigma_{\text{C–S}}^*$ and $\sigma_{\text{C–H}} \rightarrow \sigma_{\text{C–C}}^*$ donations in **Cys-IV** relative to the $\sigma_{\text{C–C}} \rightarrow \sigma_{\text{C–S}}^*$ and $\sigma_{\text{C–S}} \rightarrow \sigma_{\text{C–C}}^*$ delocalizations in **Cys-V** (Table S4, Supporting Information) preferentially stabilize **Cys-IV**.

Perhaps the most interesting conformer is **Cys-X**, which forms no typical hydrogen bonds in the sense of near-linear X–H \cdots Y arrangements. The carboxyl plane is perpendicular to the C–N bond. In this way, the amine seems to form both weak N–H \cdots O–H and N–H \cdots O=C interactions. Despite positioning the C–N bond trans to the C–S bond, **Cys-X** exhibits strong $\sigma_{\text{C–S}} \rightarrow \sigma_{\text{C–N}}^*$ hyperconjugation, contributing to its unusual stability.

3.4. Vibrational Fundamentals. Anharmonic vibrational frequencies in characteristic hydrogen bonding regions of the infrared spectra are given in Table 3 for the eleven lowest-energy conformers of Cys. Complete sets of vibrational fundamentals for these conformers are provided in Table S5 of Supporting Information. The combination of B3LYP/aug-cc-pVTZ harmonic frequencies and B3LYP/6-31G* anharmonic corrections employed here should be accurate on average to within 30 cm^{-1} , although some deviations may be substantially larger.⁸⁶

Various fundamentals can be found where either the intensities or the frequencies distinguish between structural features. For example, **Cys-I** is the only conformer that has all of the following: two medium/strong bands between $3200\text{--}3300 \text{ cm}^{-1}$, no strong bands above 3300 cm^{-1} , and only two medium/weak bands below 1300 cm^{-1} . In matrix isolation experiments on Cys, Dobrowolski et al.⁸⁷ noted that broadening and clustering of peaks often prevented assignment of absorptions to individual conformers. In general, the matrix isolation vibrational spectra were only able to distinguish between conformers as being with or without certain intramolecular hydrogen bonds. As a natural extension, Dobrowolski et al. suggested that vibrational circular dichroism (VCD) may be more informative. Because the

relatively accurate computation of VCD spectra is straightforward, such experiments hold promise for the detection of conformers present in the gas before matrix deposition.

Despite possible difficulties, four different regions should provide important fingerprints for cysteine conformers to distinguish between O–H···N or N–H···O=C bonding patterns. No conformers with strong N–H···O–H hydrogen bonds appeared in the current study. The first region is the O–H stretch region between 3200–3600 cm^{−1}. For conformers such as **Cys-I**, **Cys-II**, **Cys-VI**, and **Cys-VIII** with O–H···N hydrogen bonds, the O–H stretches are red-shifted so that the highest frequency peaks are the N–H stretches near 3400 cm^{−1}. For free hydroxyl groups, the O–H stretch appears near 3540 cm^{−1}, as seen in conformers **Cys-III**, **Cys-IV**, **Cys-V**, **Cys-VII**, **Cys-X**, and **Cys-XI**. This general band structure was observed in the matrix IR study,⁸⁷ and our anharmonic fundamentals agree with the experimental absorptions within 10–20 cm^{−1}, consistent with the uncertainty estimate given above. As noted by Dobrowolski et al.,⁸⁷ the C–O single-bond stretch also provides an important diagnostic through differing intensities. In conformers **Cys-I**, **Cys-II**, **Cys-VI**, and **Cys-VII** with O–H···N bonds, the C–O stretching region does not exhibit any high intensity peaks, presumably because the oscillator strength is smeared out among several modes. In contrast, for hydroxyl groups that do not form hydrogen bonds, the C–O stretch in **Cys-III**, **Cys-IV**, **Cys-V**, **Cys-VII**, **Cys-X**, and **Cys-XI** shows strong features between 1080 and 1125 cm^{−1}, again in excellent agreement with the experimentally observed frequencies.

As usual, the most telling band is the carbonyl stretch, ν_8 , which ranges from 1762–1807 cm^{−1}. For conformers **Cys-III**, **Cys-IV**, **Cys-V**, **Cys-VII**, **Cys-X**, and **Cys-XI**, which contain a N–H···O=C hydrogen bond, the stretch is red-shifted, appearing between 1762–1776 cm^{−1}. In contrast, the free carbonyls in **Cys-I**, **Cys-VI**, **Cys-VIII**, and **Cys-IX** all appear in a narrow range around 1795 cm^{−1}. The highest frequency occurs for **Cys-II** at 1807 cm^{−1}. However, no peak appears above 1800 cm^{−1} in the experimental spectrum.⁸⁷ **Cys-II** is essentially identical to **Cys-VIII** except for a 60° rotation of the amine group to form a N–H···S hydrogen bond, but the carbonyl stretch for **Cys-VIII** appears at 1794 cm^{−1}. It is therefore very surprising both that **Cys-II** and **Cys-VIII** have such different stretching frequencies, and also that **Cys-II**, one of the lowest energy conformers, is absent from the matrix.

The S–H stretch varies only over a narrow range of 2539–2560 cm^{−1}. This is consistent with the geometries since S–H seems to form only very weak hydrogen bonds. In general, the absorptions are also predicted to be quite weak, so that the S–H peak is not likely to be useful in distinguishing conformers.

3.5. Rotational Spectra. Alonso et al.²⁵ recently reported the identification of six low-energy conformers of Cys through Fourier transform microwave spectroscopy, providing a good opportunity here to compare the computed and experimental results. Computed rotational constants, centrifugal distortion constants, and dipole moments are reported in Table 4 along with experimental values, where available.

Table 4. Equilibrium (A_e , B_e , C_e) and Ground-State (A_0 , B_0 , C_0) Rotational Constants, Quartic Centrifugal Distortion Constants in the A-Reduced Representation, and Dipole Moments (μ_a , μ_b , μ_c) of the 11 Lowest-Energy Conformers of L-Cysteine^a

	Cys-I	Cys-II	Cys-III	Cys-IV	Cys-V	Cys-VI	Cys-VII	Cys-VIII	Cys-IX	Cys-X	Cys-XI
A_e	3059.2	4376.2	2886.6	2855.8	4235.8	3094.6	3182.5	4534.7	4505.9	2977.2	2984.0
B_e	1641.1	1189.3	1654.3	1715.8	1195.0	1605.7	1607.4	1181.2	1185.4	1558.3	1610.3
C_e	1357.6	1028.6	1390.6	1430.2	1018.5	1335.4	1302.7	972.8	970.7	1234.9	1370.0
A_0	3039.4(3071.1)	4341.7(4359.2)	2852.1(2889.4)	2812.4	4204.9(4235.6)	3055.1	3174.6(3216.2)	4502.1	4462.2	2977.1(3004.2)	2953.9
B_0	1623.0(1606.5)	1180.1(1178.3)	1642.6(1623.0)	1698.8	1187.1(1187.3)	1601.9	1587.9(1572.7)	1172.0	1176.4	1538.7(1527.4)	1605.2
C_0	1344.5(1331.8)	1018.5(1015.3)	1383.6(1367.8)	1421.0	1007.5(1003.1)	1331.3	1287.8(1276.8)	965.2	963.2	1218.3(1210.7)	1360.5
D_{JK}	−434	898	−449	−891	1036	−1054	−2621	706	763	−1145	−1416
D_J	471	89	628	628	99	614	562	59	59	438	749
D_K	1757	835	2049	2014	1379	2809	9180	1028	1187	5316	5164
μ_a	−1.55	2.84	1.00	1.08	1.69	2.52	2.06	−2.56	−2.46	2.30	1.22
μ_b	4.08	−2.62	−1.34	0.73	−0.34	−4.94	0.47	2.83	3.80	−0.36	−1.84
μ_c	−1.27	1.16	1.45	2.51	−0.66	1.74	0.04	0.29	0.66	−0.24	0.18

^a Rotational constants given in MHz and belong to the structures optimized at the MP2(FC)/aug-cc-pV(T+d)Z level. Quartic centrifugal distortion constants are given in Hz. Dipole moments are given in Debye and computed at the B3LYP/aug-cc-pVTZ level. Where available, experimental values²⁵ are given in parentheses. Ground-state rotational constants are determined from MP2(FC)/aug-cc-pVTZ equilibrium structures with anharmonic B3LYP/6-31G* corrections.

Some of the conformers (*e.g.*, **Cys-I**, **Cys-II**, and **Cys-VI**) have substantial dipole moments along the principal axes, thus helping (a) the observation of the related rotational transitions, and (b) the assignment of conformers based on information about which rotational constants correspond to intense transitions. Generally, we can divide the cysteine conformers into two groups based on the orientation about the C_α – C_β bond. In gauche conformers, the thiol is gauche to both the amine and carboxyl groups. The overall geometry is therefore more compact about the *B* axis, which is reflected in the larger B_e rotational constants for **Cys-I** and **Cys-III** in comparison to **Cys-II** and **Cys-V** (Table 4). In contrast, in a trans conformation, the thiol is positioned antiperiplanar to either the amine or carboxyl group. The overall geometry for the trans conformers is therefore extended along the *A* axis, leading to much larger A_e rotational constants for **Cys-II** and **Cys-V** in comparison to **Cys-I** and **Cys-III**.

At the MP2(FC)/aug-cc-pV(T+d)Z level, the rotational constants corresponding to the optimized equilibrium structures should be accurate enough to be useful to deduce the presence of conformers when interpreting experimental microwave spectra. As seen in Table 4, this is indeed the case. The mean absolute deviations from experiment are 15, 25, and 21 MHz for A_e , B_e , and C_e , respectively. Vibrational corrections for the rotational constants can be evaluated within the realm of second-order vibrational perturbation theory (VPT2)^{60–63} by taking one-half the sum of the lowest-order vibration–rotation interaction constants α_i . The mean absolute deviations for B_0 and C_0 become only 11 and 9 MHz.

For the *A* axis, the theoretical rotational constants consistently underestimate the experimental ones by as much as 40 MHz. For most conformers, the cysteine molecule lies along the *A* axis, with the C–S bond running roughly perpendicular along the *C* axis. The moment of inertia about the *A* axis is greatly affected by the position of the sulfur atom, and A_0 is therefore very sensitive to the C–S bond length. If the C–S bond length is systematically overestimated, then the computed A_0 constants will be too small. For example, shortening the C–S bond length by 0.005 Å in **Cys-I** increases A_e by 30 MHz, bringing the corresponding A_0 into nearly exact agreement with experiment. Such bond length discrepancies may be attributed to a number of factors, including neglect of core correlation, basis set incompleteness, or higher excitations that would be included in CCSD or CCSD(T) geometry optimizations. The errors in B_0 and C_0 also appear to be systematic, with both being consistently overestimated. In all cases, the zero-point vibrational corrections lower the rotational constant, consistent with the vibrationally averaged bond lengths being longer than their equilibrium values. Zero-point vibrational corrections to the rotational constants therefore improve agreement for B_0 and C_0 , but actually diminish the agreement for A_0 . The source of the systematic error for B_0 and C_0 is more difficult to assess than for A_0 , especially without the empirical refinement that was performed for conformers of glycine and proline.^{14,17}

Table 5. Quadrupole Coupling Constants for Conformers of Cysteine Computed at the MP2/cc-pVTZ-LD Level (See Text)^a

Conformer	χ_{aa}	χ_{bb}	χ_{cc}
Cys-I	−3.14 (−3.12)	2.44 (2.44)	0.70 (0.68)
Cys-II	−0.18 (−0.41)	2.19 (2.23)	−2.01 (−1.83)
Cys-III	−0.01 (−0.15)	0.34(0.44)	−0.32 (−0.30)
Cys-IV	−3.09	2.74	0.35
Cys-V	−4.39	2.74	1.65
Cys-VI	−3.26	2.36	0.9
Cys-VII	0.06 (0.00)	−0.48 (−0.45)	0.42 (0.45)
Cys-VIII	−3.02	1.51	1.51
Cys-IX	−3.22	1.61	1.61
Cys-X	0.51	−1.99	1.49
Cys-XI	−1.27	0.88	0.39

^a Experimental values²⁵ where known are given in parentheses. All values given in MHz.

In the experiments of Alonso et al.,²⁵ some ambiguity still remained in differentiating conformers with similar rotational constants. For example, the B_0 and C_0 rotational constants of **Cys-I** and **Cys-III** match within 40 MHz while A_0 matches within 180 MHz. Furthermore, the computed rotational constants for **Cys-I** lie in between the observed values. For example, C_0 for **Cys-I** is computed to be 1344.5 MHz, in between the observed values of 1331.8 and 1367.8 MHz. Quadrupole coupling constants of the nitrogen nucleus are therefore necessary to uniquely identify such conformers. The quadrupole coupling constants (χ_{aa}) are given as⁸⁸

$$\chi_{aa} = eq_{aa}Q \quad (5)$$

where q_{aa} is the electric field gradient along the α -axis at the nitrogen nucleus, e is the fundamental charge, and Q is the nuclear quadrupole moment. For the nitrogen quadrupole moment, we use the literature value of 20.44 mb.⁸⁹ As seen for ammonia (Table S6, Supporting Information), the accurate computation of electric field gradients at the nuclei presents a difficult theoretical problem. Similar difficulties hold for spin-dependent properties that depend on contact terms since the amplitude and shape of the wave function near the nuclei must be very accurately described.^{90,91} In particular, Gaussian basis functions have the incorrect shape at the nuclei, so that extremely flexible and carefully designed basis sets are required for accurate results.

In general, double- ζ basis sets and Hartree–Fock methods are not flexible enough to yield good results for quadrupole coupling constants. Since the coupling constant depends only on the nitrogen nucleus, it is possible to use a locally dense basis on the nitrogen atom.⁹² The combination of a cc-pCV5Z basis on nitrogen and cc-pVTZ basis on all other atoms (denoted cc-pVTZ-LD) very closely matches both the full cc-pCV5Z result and the experimental coupling constant (4.09 MHz)⁹³ for ammonia (Table S2, Supporting Information). Furthermore, probably through fortuitous cancelation of errors, MP2 matches the CCSD(T) results well, better than even CCSD. The combination of MP2 and the locally dense basis therefore seems to offer the best combination of accuracy and efficiency for computing the quadrupole coupling constants of cysteine.

The computed quadrupole coupling constants for the conformers of cysteine are presented in Table 5. The MP2/

cc-pVTZ-LD approximation generally performs quite well, yielding most coupling constants within 0.1 MHz of experiment with the largest deviation being 0.23 MHz. In particular, the ambiguity in assignment based on rotational constants is now removed. For example, **Cys-I** has a strong χ_{aa} quadrupole coupling while **Cys-III** exhibits almost no χ_{aa} coupling, in agreement with the experimental results of Alonso et al.²⁵

4. Summary

In the present work, we performed a comprehensive study of the important structural features and spectroscopic signatures of the amino acid L-cysteine. Through the focal point approach, we have established definitive relative energies of the eleven lowest conformers to within a standard error of 0.3 kJ mol⁻¹ (1 σ) or 95% confidence interval of \pm 0.6 kJ mol⁻¹ (2 σ).

Because of the added flexibility of the thiol side chain, cysteine exhibits 71 unique conformers (fully specified in Supporting Information) and eleven conformers within 10 kJ mol⁻¹ of the lowest minimum. As observed previously,¹ Hartree–Fock energies are inaccurate. Inclusion of electron correlation with B3LYP greatly improves results, but still fails by more than 2.5 kJ mol⁻¹ for some conformers, which becomes significant when so many conformers lie within a narrow 10 kJ mol⁻¹ range. In general, B3LYP performs well for geometries and zero-point vibrational corrections, but inclusion of correlation through at least MP2 seems necessary for accurate energies. Definitive energies to 0.5 kJ mol⁻¹ accuracy still require corrections through CCSD(T).

While hydrogen bonding and electrostatics are the most important factors determining structures and energetics, the bond length, bond angle, and energy changes between conformers depend strongly on subtle electronic effects, including hyperconjugation, steric repulsion, hydrogen-bond cooperativity, and dispersion forces. In contrast to previous work, we therefore emphasize features such as the trans angle rule⁸⁰ and the gauche effect.^{84,85} An additive picture of hydrogen bonds may therefore be overly simplistic for cysteine.

Harmonic frequencies were computed at the B3LYP/aug-cc-pVTZ level with anharmonic corrections at the B3LYP/6-31G* level (Tables 3 and S4, Supporting Information). The vibrational perturbation theory generally performs well (within 20 cm⁻¹ of experiment), although it breaks down for a few large-amplitude motions with very low frequencies. The computed fundamentals should aid future IR spectroscopy studies. Since we are aiming for accuracy near 0.5 kJ mol⁻¹ in the conformational energies, rigorous anharmonic zero-point vibrational corrections are necessary instead of simply scaling harmonic frequencies.

The extensive ab initio results reported here should serve as an important reference both for calibrating more approximate theoretical methods or future experiments, including circular dichroism or infrared and microwave spectroscopy of isotopologues of cysteine. As more empirical data becomes available (e.g., rotational constants of isotopo-

logues), the structures and energies can be further refined by empirical fitting, as done previously for glycine¹⁴ and proline.¹⁷

Acknowledgment. This work was supported by NSF grant CHE-0749868, an NSF-MTA-OTKA grant, and the Hungarian Scientific Research Fund (OTKA, K72885, IN77954). Most computations were run at the Pittsburgh Supercomputing Center under TeraGrid grant TG-CHE070039N.

Supporting Information Available: Comparison of cysteine conformational energies with previously computed values, zero-point vibrational energy comparison for B3LYP/aug-cc-pVTZ and MP2/DZP++ levels of theory, summary of N–C–C bond angle and hyperconjugation changes due to amine rotation, NBO summary of hyperconjugation for varying conformations about the C α –C β bond, complete table of anharmonic vibrational fundamentals and double-harmonic infrared relative intensities, electron correlation and basis set dependence of computed quadrupole coupling constants of ammonia, optimized geometries and energies at MP2(FC)/aug-cc-pV(T+d)Z level of theory for eleven lowest energy conformers of L-cysteine, and optimized geometries and energies at MP2(FC)/cc-pVTZ level of theory for 71 unique conformers of L-cysteine. This information is available free of charge via the Internet at <http://pubs.acs.org/>.

References

- (1) Császár, A. G.; Perczel, A. *Prog. Biophys. Mol. Biol.* **1999**, *71*, 243–309.
- (2) Cox, M. M.; Nelson, D. L., *Principles of Biochemistry*, 4th ed; W.H. Freeman and Company: New York, 2005.
- (3) Lu, X. F.; Galkin, A.; Herzberg, O.; Dunaway-Mariano, D. *J. Am. Chem. Soc.* **2004**, *126*, 5374–5375.
- (4) Li, H. M.; Thomas, G. J. *J. Am. Chem. Soc.* **1991**, *113*, 456–462.
- (5) Császár, A. G. *J. Am. Chem. Soc.* **1992**, *114*, 9568–9575.
- (6) Hu, C. H.; Shen, M. Z.; Schaefer, H. F. *J. Am. Chem. Soc.* **1993**, *115*, 2923–2929.
- (7) Allen, W. D.; East, A. L. L.; Császár, A. G., *Structures and Conformations of Non-Rigid Molecules*; Kluwer: Dordrecht, The Netherlands, 1993; p 343.
- (8) Császár, A. G.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **1998**, *108*, 9751–9764.
- (9) East, A. L. L.; Allen, W. D. *J. Chem. Phys.* **1993**, *99*, 4638–4650.
- (10) Császár, A. G.; Tarczay, G.; Leininger, M. L.; Polyansky, O. L.; Tennyson, J.; Allen, W. D. *Spectroscopy from Space*; Demaison, J., Sarka, K., Eds.; Kluwer: Dordrecht, The Netherlands, 2001; p 317–339.
- (11) Gonzales, J. M.; Pak, C.; Cox, R. S.; Allen, W. D.; Schaefer, H. F.; Császár, A. G.; Tarczay, G. *Chem.—Eur. J.* **2003**, *9*, 2173–2192.
- (12) Schuurman, M. S.; Muir, S. R.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **2004**, *120*, 11586–11599.
- (13) Császár, A. G. *J. Mol. Struct.* **1995**, *346*, 141–152.

- (14) Kasalová, V.; Allen, W. D.; Schaefer, H. F.; Czinki, E.; Császár, A. G. *J. Comput. Chem.* **2007**, *28*, 1373–1383.
- (15) Császár, A. G. *J. Phys. Chem.* **1996**, *100*, 3541–3551.
- (16) Szidarovszky, T.; Czakó, G.; Császár, A. G. *Mol. Phys.* **2009**, DOI: 10.1080/00268970802616350.
- (17) Allen, W. D.; Czinki, E.; Császár, A. G. *Chem.—Eur. J.* **2004**, *10*, 4512–4517.
- (18) Czinki, E.; Császár, A. G. *Chem.—Eur. J.* **2003**, *9*, 1008–1019.
- (19) Laurence, P. R.; Thomson, C. *Theor. Chim. Acta* **1981**, *58*, 121–124.
- (20) Wright, L. R.; Borkman, R. F. *J. Am. Chem. Soc.* **1980**, *102*, 6207–6210.
- (21) Schäfer, L.; Kulpnewton, S. Q.; Siam, K.; Klimkowski, V. J.; Vanalsenoy, C. *J. Mol. Struct. (THEOCHEM)* **1990**, *68*, 373–385.
- (22) Gronert, S.; O’Hair, R. A. J. *J. Am. Chem. Soc.* **1995**, *117*, 2071–2081.
- (23) Krishnan, R.; Frisch, M. J.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 4244–4245.
- (24) Dobrowolski, J. C.; Rode, J. E.; Sadlej, J. *J. Mol. Struct. THEOCHEM* **2007**, *810*, 129–134.
- (25) Sanz, M. E.; Blanco, S.; Lopez, J. C.; Alonso, J. L. *Angew. Chem., Int. Ed.* **2008**, *47*, 6216–6220.
- (26) Frisch, M. J.; Pople, J. A.; Delbene, J. E. *J. Phys. Chem.* **1985**, *89*, 3664–3669.
- (27) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939–947.
- (28) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797–2803.
- (29) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.
- (30) Peterson, K. A.; Dunning, T. H., Jr. *J. Chem. Phys.* **2002**, *117*, 10548–10560.
- (31) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (32) Dunning, T. H. Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (33) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1995**, *103*, 4572–4585.
- (34) Wilson, A. K.; Dunning, T. H., Jr. *J. Chem. Phys.* **2003**, *119*, 11712–11714.
- (35) Pulay, P. *Mol. Phys.* **1969**, *17*, 197–204.
- (36) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley-Interscience: New York, 1986.
- (37) Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256–4266.
- (38) Crawford, T. D.; Schaefer, H. F. *Rev. Comp. Chem.* **2000**, *14*, 33–136.
- (39) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
- (40) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (41) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (42) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (43) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (44) Jayatilaka, D.; Lee, T. J. *J. Chem. Phys.* **1993**, *98*, 9734–9747.
- (45) Taylor, P. R.; Lee, T. J. *Int. J. Quantum Chem. Symp.* **1989**, *23*, 199.
- (46) Stanton, J. F.; Gauss, J.; Watts, J. D.; Szalay, P. G.; Bartlett, R. J.; Auer, A. A.; Bernholdt, D. B.; Christiansen, O.; Harding, M. E.; Heckert, M.; Heun, O.; Huber, C.; Jonsson, D.; Jusélius, J.; Lauderdale, W. J.; Metzroth, T.; Michauk, C.; O’Neill, D. P.; Price, D. R.; Ruud, R.; Schiffmann, F.; Varner, M. E.; Vázquez, J. *ACES II*, Mainz-Austin: Budapest, Hungary, 2005.
- (47) Janssen, C. L.; Nielsen, I. B.; Leininger, M. L.; Valeev, E. F.; Seidl, E. T., *The Massively Parallel Quantum Chemistry Program (MPQC)*, version 2.3.1; Sandia National Laboratories: 2004.
- (48) Nielsen, I. M. B. *Chem. Phys. Lett.* **1996**, *255*, 210–216.
- (49) Nielsen, I. M. B.; Seidl, E. T. *J. Comput. Chem.* **1995**, *16*, 1301–1313.
- (50) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Rauhut, G.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; An P. Palmieri, A. N.; Pitzer, R.; Schumann, U.; Stoll, H.; A. J. Stone, R. T.; Thorsteinsson, T., *MOLPRO*, version 2006.1; 2006.
- (51) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudling, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenber, J. J.; Zkrzewski, V. G.; Dappich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanykkara, A.; Callacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A., *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.
- (52) Eliel, E. L.; Wilen, S. H.; Mander, L. N., *Stereochemistry of Organic Compounds*; John Wiley & Sons: New York, 1994.
- (53) Klopper, W.; Kutzelnigg, W. *J. Mol. Struct. (THEOCHEM)* **1986**, *28*, 339–356.
- (54) Karton, A.; Martin, J. M. L. *Theor. Chem. Acc.* **2006**, *115*, 330–333.
- (55) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104–6114.
- (56) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
- (57) Tarczay, G.; Császár, A. G.; Klopper, W.; Quiney, H. M. *Mol. Phys.* **2001**, *99*, 1769–1794.
- (58) Handy, N. C.; Yamaguchi, Y.; Schaefer, H. F. *J. Chem. Phys.* **1986**, *84*, 4481–4484.
- (59) Allen, W. D.; Császár, A. G. *J. Chem. Phys.* **1993**, *98*, 2983–3015.

- (60) Clabo, D. A., Jr.; Allen, W. D.; Remington, R. B.; Yamaguchi, Y.; Schaefer, H. F. *Chem. Phys.* **1988**, *123*, 187–239.
- (61) Allen, W. D.; Yamaguchi, Y.; Császár, A. G.; Clabo, D. A., Jr.; Remington, R. B.; Schaefer, H. F. *Chem. Phys.* **1990**, *145*, 427–466.
- (62) Nielsen, H. H. *Rev. Mod. Phys.* **1951**, *23*, 90–136.
- (63) Mills, I. M., *Molecular Spectroscopy: Modern Research*; Rao, K. N., Mathews, C. W., Eds.; Academic Press: New York, 1972; p 115.
- (64) Schuurman, M. S.; Allen, W. D.; Schleyer, P. v. R.; Schaefer, H. F. *J. Chem. Phys.* **2005**, *122*, 104302. Schuurman, M. S.; Allen, W. D.; Schaefer, H. F. *J. Comput. Chem.* **2005**, *26*, 1106–1112.
- (65) Barone, V. *J. Chem. Phys.* **2004**, *120*, 3059–3065.
- (66) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899–926.
- (67) Reed, A. E.; Weinhold, F.; Curtiss, L. A.; Pochatko, D. J. *J. Chem. Phys.* **1986**, *84*, 5687–5705.
- (68) Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066–4073.
- (69) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211–7218.
- (70) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735–746.
- (71) Simmonett, A. C.; Evangelista, F. A.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **2007**, *127*, 014306.
- (72) Zhang, X.; Maccarone, A. T.; Nimlos, M. R.; Kato, S.; Bierbaum, V. M.; Ellison, G. B.; Ruscic, B.; Simmonett, A. C.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **2007**, *126*, 044312.
- (73) Wilke, J. J.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **2008**, *128*, 074308.
- (74) Czakó, G.; Mátyus, E.; Simmonett, A. C.; Császár, A. G.; Schaefer, H. F.; Allen, W. D. *J. Chem. Theory Comput.* **2008**, *4*, 1220–1229.
- (75) Simmonett, A. C.; Schaefer, H. F.; Allen, W. D. *J. Chem. Phys.* **2009**, *130*, 044301.
- (76) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.
- (77) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129–4141.
- (78) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vazquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599–11613.
- (79) Dunning, T. H., Jr. *J. Chem. Phys.* **1970**, *53*, 2823–2833.
- (80) Rasänen, M.; Aspiala, A.; Homanen, L.; Murto, J. *J. Mol. Struct.* **1982**, *96*, 81–100.
- (81) Pross, A.; Radom, L.; Riggs, N. V. *J. Am. Chem. Soc.* **1980**, *102*, 2253–2259.
- (82) Weinhold, F.; Landis, C. L., *Valency and Bonding*; Cambridge University Press: Cambridge, UK, 2005.
- (83) Allinger, N. L.; Fermann, J. T.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **1997**, *106*, 5143–5150.
- (84) Goodman, L.; Gu, H. B.; Pophristic, V. *J. Phys. Chem. A* **2005**, *109*, 1223–1229.
- (85) Lesarri, A.; Cocinero, E. J.; Lopez, J. C.; Alonso, J. L. *J. Am. Chem. Soc.* **2005**, *127*, 2572–2579.
- (86) Carbonniere, P.; Lucca, T.; Pouchan, C.; Rega, N.; Barone, V. *J. Comput. Chem.* **2005**, *26*, 384–388.
- (87) Dobrowolski, J. C.; Jamroz, M. H.; Kolos, R.; Rode, J. E.; Sadlej, J. *ChemPhysChem* **2007**, *8*, 1085–1094.
- (88) Hollas, J. M., *High Resolution Spectroscopy*, 2nd ed; Wiley: New York, 1998.
- (89) Polak, R.; Fiser, J. *J. Chem. Phys.* **2008**, *351*, 83–90.
- (90) Helgaker, T.; Jaszunski, M.; Ruud, K.; Gorska, A. *Theor. Chem. Acc.* **1998**, *99*, 175–182.
- (91) Benedikt, U.; Auer, A. A.; Jensen, F. *J. Chem. Phys.* **2008**, *129*, 064111.
- (92) Provasi, P. F.; Aucar, G. A.; Sauer, S. P. A. *J. Chem. Phys.* **2000**, *112*, 6201–6208.
- (93) Grigolin, P.; Moccia, R. *J. Chem. Phys.* **1972**, *57*, 1369–1376.