Theoretical Study of the Electronic Gas-Phase Spectrum of Glycine, Alanine, and Related Amines and Carboxylic Acids

Anders Osted,*,† Jacob Kongsted,† and Ove Christiansen

Department of Chemistry, University of Århus, Langelandsgade 140, DK-8000 Århus C, Denmark

Received: September 22, 2004

A theoretical study on the origin of the common electronic excitations in amino acids is presented, focusing on the excited states of glycine, alanine and the related substructures formic acid, acetic acid, propionic acid, ammonia, methylamine, and ethylamine. Special attention is given to the valence excitation from the nonbonding lone-pair on the carboxylic oxygen atom to the antibonding π-orbital ($n_O \rightarrow \pi^*_{CO}$) and the first Rydberg excitation from the nonbonding lone-pair on the nitrogen atom ($n_N \rightarrow 3s$). From extensive calculations on formic acid and methylamine, different basis sets and electron correlation treatments are benchmarked using a hierarchy of coupled cluster (CC) methods, consisting of CCS, CC2, CCSD, CCSDR(3), and CC3, in combination with augmented correlation consistent basis sets. The dependence of the excitation energies on the size of the backbone structure in the two groups of molecules is investigated, and 0−0 transition energies for the $n_O \rightarrow \pi^*_{CO}$ and $n_N \rightarrow 3s$ transitions are calculated for the smallest molecules. Excellent agreement with experimental values is found where secure experimental assignments are available. A few outstanding problems in the experimental assignments found in the literature are described for both the carboxylic acids and the amines. Final predictions for vertical excitation energies are given for all molecules, including glycine and alanine where no gas-phase experimental results are available. Finally, calculations on protonated amino acids are presented showing an isolation of the $n_O \rightarrow \pi^*_{CO}$ from higher lying states by as much as 1.9 eV for alanine.

I. Introduction

Amino acids are the building blocks from which all proteins are built. Only 20 different amino acids are found in the naturally occurring proteins on earth. All are α-amino acids, and all but glycine have the α-carbon as a stereogenic center giving two enantiomeric forms (L- and D-forms). The enantiomers of a given amino acid share most physical properties, for example, electronic excitation energies, but differ in the direction in which they rotate plane-polarized light, as evident from optical rotatory dispersion. The backbone structure of all α-amino acids contains two functional groups, namely, the unsaturated carboxylic acid group and the saturated amino group attached to the α-carbon. The aim of this study is to investigate theoretically the origin of the electronic excitations common to all amino acids due to the carboxylic acid group and the amino group. More specific, we shall present calculations of the gas-phase electronic spectrum for glycine and alanine up to around 8 eV and selected states above this limit. To put the calculated results in context and to verify the quality of the calculated procedures, we report also calculations on electronic spectra for a number of related molecules corresponding to the amino acid subgroups consisting of carboxylic acids and amines. This includes formic acid, acetic acid, propionic acid, ammonia, methylamine, and ethylamine.

The motivation for the present work originates in experimental work on ultrafast laser spectroscopy on some of the above-mentioned compounds.[1,2] These studies stimulated a number of questions concerning which states are present in which energy regions in both vacuum and in solution, as well as questions on what takes place after the photoexcitation. Many such questions are not convincingly answered in the literature. In this work, we focus on the theoretical calculation of UV−VUV absorption of the compounds in the gas phase. In later studies, we wish to build upon this work in studying the effect of solvents and aspects of the dynamics following photoexcitation. Furthermore, we also wish to return to theoretical studies of the interactions between light and amino acids in other contexts. Indeed, the recent circular dichroism (CD) experiments on amino acids[3] and the reports on enantiomer excess of L amino acids over D amino acids found in meteorites,[4−7] were another motivation for initiating studies of these molecules. The later fact has added momentum to the idea that the homochirality of biomolecules has its origin in matter of extraterrestrial origin and that this matter may be enantiomerically enriched from interaction with circular polarized light in outer space.[8−10] Exactly how this should occur is, however, still not completely well established, and there are other suggestions for the origin of the homochirality of naturally occurring amino acids.[11−13]

Some features of the energetically lower lying electronic transitions of both the amines and the carboxylic acids are well understood in the gas phase through studies of the smallest compounds within each group of molecules. The lowest excitation energy within carbonyl compounds is known to be a $n_O \rightarrow \pi^*_{CO}$ transition from the lone-pair on the carbonyl oxygen to the antibonding π_{CO}-valence orbital. Though the analogous excitation within the electric dipole approximation is symmetry forbidden in formaldehyde, this transition is symmetry allowed in the carboxylic group. The weak transition has been measured in several studies of smaller carboxylic acids in the gas

* To whom correspondence should be addressed. E-mail: osted@theory.ki.ku.dk.

† Permanent address: Department of Chemistry, H. C. Ørsted Institute, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark.

10.1021/jp045697f CCC: $30.25
Published on Web 02/02/2005

phase.[14−22] These studies also discuss the transitions at higher energies, and the lowest lying of these appear to be Rydberg transitions from the oxygen lone-pair to 3s, 3p, and 3d Rydberg orbitals ($n_O \rightarrow 3s$, $n_O \rightarrow 3p$, and $n_O \rightarrow 3d$) and valence transitions from the bonding π and bonding σ orbitals to the antibonding π-valence orbital ($\pi_{C=O} \rightarrow \pi^*_{CO}$ and $\sigma_{C-O} \rightarrow \pi^*_{CO}$). However, the exact order and location of the transitions differ within the different studies. Calculations of the gas-phase transitions of formic acid below 11 eV have been reported,[23,24] appointing much the same ordering of the lowest transitions as the experiments. The lowest excitation has also been calculated for microsolvated (by water) formic acid showing a blue shift of the transition energy.[25] Aloisio et al. have reported studies of the geometry dependence of hydrated formic acid in small water clusters.[26] Finally, conformational studies of the excited state of the $n_O \rightarrow \pi^*_{CO}$ transition shows that the potential energy curve of the excited state is rather different from the ground state.[19,27]

Turning to the amines, the energetically lower part of the gas-phase spectrum is dominated by Rydberg transitions from the lone-pair electrons located on the nitrogen atom to 3s, 3p, 3d, 4s, and 4p Rydberg orbitals ($n_N \rightarrow 3s$, $n_N \rightarrow 3p$, $n_N \rightarrow 3d$, $n_N \rightarrow 4s$, and $n_N \rightarrow 4p$). Experimentally, most studies have concerned the ammonia molecule[28−32] which is the starting point for all amines but also studies of methylamine and ethylamine have been reported.[28,33,34] Theoretical studies have also mostly concerned the ammonia spectrum and the assignment of the excited states.[35−37]

The work performed on the amino acids is far less extensive than the work on the group of carboxylic acids and amines. To our knowledge, no one has reported on calculations on the excitation energies and no experiments have been reported on gas-phase spectra either. However, some experimental work on amino acids in solution or as thin layers has been performed. Neta et al. have measured the absorption spectrum of some amino acids in aqueous solution at different pH values,[38] and Vinogradov et al. report the absorption spectrum of thin layers of aliphatic amino acids and their di- and tripeptides.[39] Nishino et al. have determined the UV and CD spectra of the $n_O \rightarrow \pi^*_{CO}$ transition of L- and D-alanine, among other amino acids, in aqueous solution at different pHs.[3] The CD spectrum of several amino acids including alanine in hexafluoroisopropanol (HFIP) has been reported by Snyder et al.[40]

Direct comparison of our gas-phase calculated results with the available condensed-phase experiments for glycine and alanine is problematic. The interaction with the surroundings will affect the energetic location and maybe also the ordering of the states. Furthermore, in solution the amino acids will be in different forms depending on the pH of the solution because of the basic/acidic behavior of the amino/carboxylic group. For the smaller carboxylic acids and amines, experimental results and other calculations are available. The secure experimental assignments are in excellent agreement with our calculations. In a few cases, experimental assignments in the literature are far outside the expected error bars of the calculations as obtained from the systematic trends in the calculations.

Our calculations of the vertical excitation energies apply the hierarchy of coupled cluster (CC) methods, consisting of the CCS, CC2,[41] CCSD,[42] CCSDR(3),[43] and CC3[44] models. The unique feature of this CC hierarchy and its efficient implementation for transition energies, transition properties, and excited-state properties[45,46] is that it allows for a *systematic* investigation of the convergence with respect to the treatment of electron correlation and basis set. This allows us to perform thorough investigations for the smaller systems within the groups of carboxylic acids and amines and design a proper compromise for the level of electron correlation and basis set used also for the larger amino acids. Former calculations on benzene,[45,47] *s*-tetrazine,[48] pyrrole,[49] furan,[50] and pyrimidine[51] have shown the CC method to be very useful for high-accuracy studies of electronic excitation energies of molecules of this size. A major difference between this and the former studies is the lack of symmetry in the amino acid structures, making the calculations much more computationally expensive.

With reference to the Franck−Condon approximation, the calculated vertical excitation energies are most often considered the same as, and are therefore compared with, the experimental determined energy with maximum absorption ϵ_{max} in the relevant band. This is an approximation. For example, the calculated vertical excitations contain no corrections due to the difference in zero-point vibrational energy between the electronic ground state and the typically less bound excited electronic state, and the experimental band structures may be broad, complex, and overlapping. Thus, this particular way of comparing theoretical and experimental results introduces a potential inaccuracy of a few tenths of an electronvolt. In fact, this additional inaccuracy may be quite large compared to the present standards for accurate calculations of theoretical vertical excitation energies. However, efficient CC methods for calculating equilibrium structures and harmonic vibrational frequencies have been developed,[52,53] which enable us to calculate 0−0 transition energies. When experimental 0−0 transition energies are available, a rigorous comparison of equivalent quantities between theory and experiment may be performed. However, these calculations are still rather demanding and have therefore only been performed for a few of the lowest electronic states in the smallest molecules.

II. Theory and Computational Details

The response methodology[54,55] allows a simple extension of standard wave function theories to the treatment of electronic transitions. Applying response theory to the CC formalism, the quality of the results very much depends on two aspects. First, the ground state of the system must be dominated by a single Slater determinant as conventional CC methods use a single reference state, most often a Hartree−Fock (HF) state. This criterion is fulfilled for all the systems considered in this paper. Second, excitations that are qualitatively described as single-electron excitations are described highly accurately whereas simultaneously excitation of two electrons is described significantly less accurately. Analyzing the percentage contribution of the CCSD response eigenvectors from single excitations for all the systems under consideration shows that in all cases the contribution is more than 90% as it should be for fairly pure single excitations. From these considerations, we can expect that the electronic spectra are rather accurately determined from CC response calculations.

The CC methodology allows for a systematic investigation of the convergence with respect to the treatment of the electronic correlation using the hierarchy of iterative CC models CCS, CC2,[41] CCSD,[42] and CC3[44] of increasing accuracy. However, the models become more computationally expensive when moving up in the hierarchy. The noniterative CCSDR(3) model for calculating excitation energies has been designed to account for the triple excitation corrections at reduced cost (a factor of 10−20) compared to the CC3 model. The CCSDR(3) and the CC3 models have been shown to give similar results,[43] and we therefore choose to use primarily the CCSDR(3) model in the

hierarchy of CC models to account for the triple excitation corrections in this work. Comparisons of full configuration-interaction (FCI) calculations and CC calculations including triple excitations by CC3 or CCSDR(3) have shown a mean-(max) error of approximately 0.03(0.1) eV.[43,56] Finally, previous predictions of 0−0 excitation energies calculated using CC theory have agreed with experimental results within an accuracy of about 0.1 eV.[45,47,49−51]

The calculated CC excitation energies have been analyzed from a breakdown of the resulting response eigenvectors to see which Hartree−Fock molecular orbitals (MOs) are involved in the primary excitations. Symmetry has been used (when present) together with basis MO theory to predict which valence and Rydberg states are present in each symmetry, and together with calculated properties of the excited states such as the second-order moment of charge we use the above MO analysis to assign specific excitation characters to the calculated energies. For example, we summarize this information in the simplified $n_O \rightarrow \pi^*_{CO}$ assignment for the excitation from the nonbonding lone-pair on the carboxylic oxygen atom to the antibonding carboxyl valence state. This is done in order to give some qualitative information on the excited states, but configurational mixing is always present and in some cases such representations are actually too simple.

The geometries for the structures of the amines, the carboxylic acids, and the amino acids used in this study have been obtained from geometry optimizations using the Gaussian 98 program package.[57] All geometries are optimized at the MP2 level of theory using the cc-pVTZ basis set.[58] From these optimized structures, vertical excitation energies (ω), oscillator strengths (f), and first-order properties for the excited states have been calculated with the CC response program[45,46] in the Dalton program package.[59] The set of correlation consistent basis sets x-aug-cc-pVXZ, x = -,d and X = D,T, of Dunning[58,60] have been employed. To reduce the computational expenses, the core electrons were kept frozen in the correlated calculations, using canonical HF orbitals. We performed test calculations which showed that the error introduced was negligible with the number of digits reported here.

Geometry optimizations and vibrational frequency calculations are carried out using CC methods for the ground and selected excited states in a few of the smallest structures, namely, formic acid, ammonia, and methylamine. The aug-cc-pVD(T)Z basis set (see later description) has been used in these calculations. The purpose of these optimizations is to study the effects of geometrical relaxation and zero-point vibrations. The calculations are carried out using analytical gradient techniques[52,53] as implemented in the ACESII[61] program package. All electrons were correlated in these calculations.

III. Results and Discussion

A. Testing Basis Set and Coupled Cluster Method. We have performed several calculations using the correlation consistent basis sets (x-aug-)cc-pVXZ (x = -,d and X = D,T) of Dunning[58,60] on formic acid and methylamine. We use these subsystems of alanine to determine a proper basis set and coupled cluster method as each of them contain one of the two primary excitations of interest in alanine, namely, $n_O \rightarrow \pi^*_{CO}$ and $n_N \rightarrow 3s$.

In Figure 1, we illustrate the dependence of the basis set used for CCSD calculations of the five lowest electronic excitation energies in formic acid. From Figure 1, we see that in order to reproduce the results obtained from the largest basis set with a proper accuracy the basis set has to be of the aug-cc-pVXZ

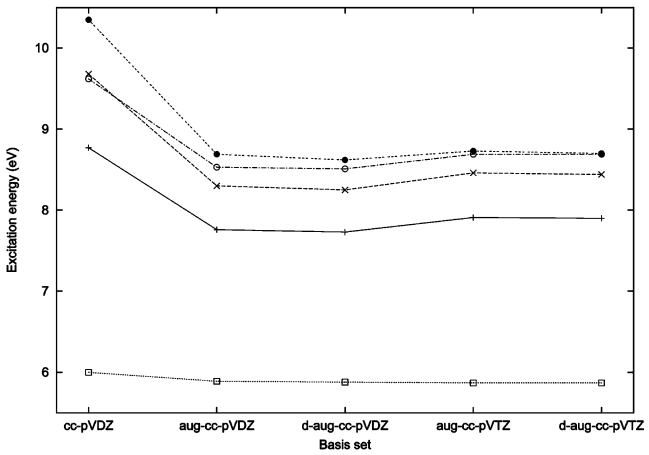

Figure 1. The energetically lowest excitation energies of formic acid within each symmetry plotted as a function of the basis set used in the calculation. All calculations have been performed within the CCSD electron correlation method. $2^1A'$ (— + —); $3^1A'$ (— — × — —); $4^1A'$ (- - - ● - - -); $1^1A''$ (···□···); $2^1A''$ (· — · ○ — · —).

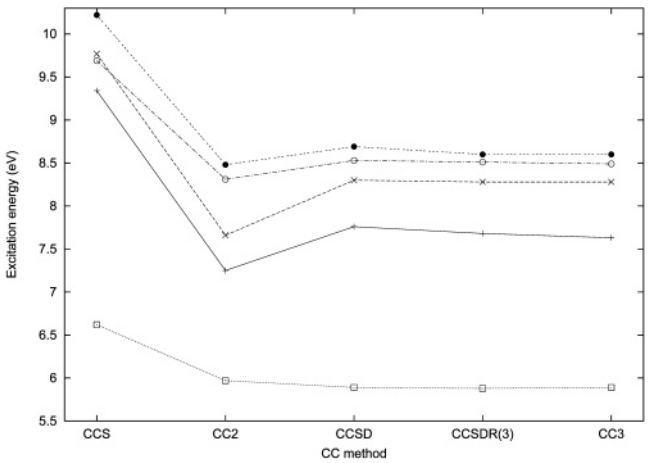

Figure 2. The energetically lowest excitation energies of formic acid within each symmetry plotted as a function of different coupled cluster truncations. All calculations have been performed using the aug-cc-pVDZ basis set. $2^1A'$ (— + —); $3^1A'$ (— — × — —); $4^1A'$ (- - - ● - - -); $1^1A''$ (···□···); $2^1A''$ (· — · ○ — · —).

type. Using the aug-cc-pVDZ basis set, the first and second valence excitations, $n_O \rightarrow \pi^*_{CO}$ ($1^1A''$) and $\pi_{C=O} \rightarrow \pi^*_{CO}$ ($4^1A'$), are reproduced within less than 0.5%. Also for the three other excited states a reproduction of the d-aug-cc-pVTZ results is seen within 2% using this basis set. From Figure 1, it is seen that the excitation energies for the Rydberg states ($2^1A'$, $3^1A'$, and $2^1A''$) depend much more on the choice of basis set than the excitation energy for valence states. The Rydberg states are very diffuse and they differ in the number of valence electrons from the ground state, and larger basis sets are therefore required for an accurate representation. The increase of about 0.15 eV going from the d-aug-cc-pVDZ to the d-aug-cc-pVTZ quality basis set has been observed for Rydberg states in many other molecules.[45,47,49−51]

In Figure 2, we illustrate how the five lowest excitation energies of formic acid depend on the treatment of electronic correlation. All the calculations have been performed using the aug-cc-pVDZ basis set. From Figure 2, it is seen that CCS greatly exaggerates the vertical excitation energy. The CC2 method corrects this, but for some states the excitation energy is underestimated significantly. The CCSD model shows a mean(max) deviation of 0.25(1.0)% as compared to the triple-

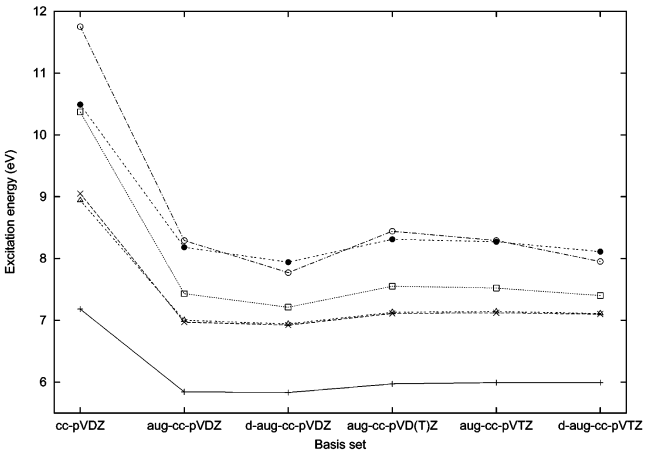

Figure 3. The energetically lowest excitation energies of methylamine within each symmetry plotted as a function of the basis set used in the calculation. All calculations have been performed within the CCSD electron correlation method. $2^1A'$ (— + —); $3^1A'$ (— — × — —); $4^1A'$ (···□···); $5^1A'$ (— • ○ThinSpace— •); $1^1A''$ (- - - △ - - -); $2^1A''$ (- - ● - -).

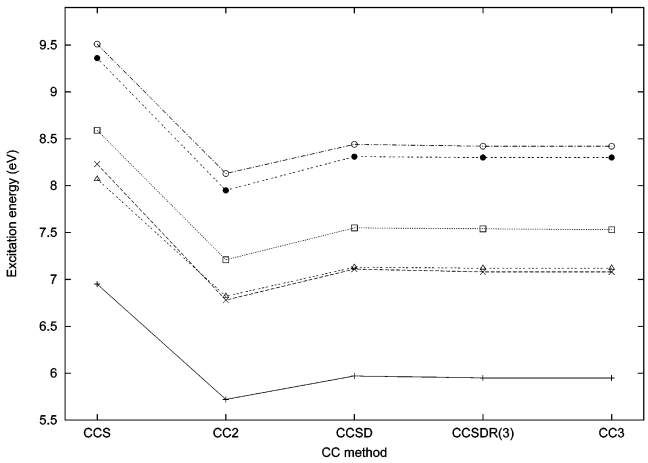

Figure 4. The energetically lowest excitation energies of methylamine within each symmetry plotted as a function of different coupled cluster truncations. All calculations have been performed using the aug-cc-pVD(T)Z basis set. $2^1A'$ (— + —); $3^1A'$ (— — × — —); $4^1A'$ (···□···); $5^1A'$ (— • ○ThinSpace— •); $1^1A''$ (- - - △ - - -); $2^1A''$ (- - ● - -).

corrected results. From Figure 2, it is seen that the Rydberg excitation energies are also very sensitive to the representation of electron correlation.

We conclude that the aug-cc-pVDZ basis set with the CCSD model gives results for the excitation energies of the carboxylic acid group with an accuracy of order 0.2 eV, and better (about 0.1 eV) for the $n_O \rightarrow \pi^*_{CO}$ and $\pi_{C=O} \rightarrow \pi^*_{CO}$ valence excitations. We shall later return to even more accurate estimates. The oscillator strengths calculated using this basis set and coupled cluster method also agree fairly well with the results using much larger basis sets. The deviations are up to 20% in some cases, but these deviations are observed on small numbers. Thus, this type of accuracy is acceptable for the oscillator strength where an order of magnitude and the relative strength between states are often sufficient.

In Figure 3, we show the dependence on the choice of basis set for the six lowest lying CCSD excitation energies of methylamine. From Figure 3, it is seen that in order to obtain the $n_N \rightarrow 3s$ transition energies within 0.1 eV of the d-aug-cc-pVTZ results one has to apply the aug-cc-pVTZ basis set if we choose to use the same type of basis set for all the atoms in the molecule. However, the excitations in amines in the energy region of interest are Rydberg transitions from the lone-pair electrons on nitrogen to nitrogen atomic-like orbitals. This implies that we might get similar results when applying the larger aug-cc-pVTZ basis set to the nitrogen atom and keeping the smaller aug-cc-pVDZ on the carbon and hydrogen atoms. For this type of basis set, we use the abbreviation aug-cc-pVD(T)Z (aD(T)). Results for this basis set are included in Figure 3 and deviate less than 0.5% for the three lowest excitations and less than 6% for the higher transitions from the d-aug-cc-pVTZ results.

In Figure 4, the six lowest excitation energies within the two symmetries of methylamine are plotted as a function of electron correlation models. All calculations have been performed using the aug-cc-pVD(T)Z basis set. As in the case of the carboxylic acid group, it is seen that CCS has very large errors, which are overcorrected by CC2. The CCSD model reproduces the triple-corrected results with mean(max) deviations of less than 0.2-(0.3)%. From Figure 3 and Figure 4, we conclude that in order to get results of the excitation energies in amines that are generally within 0.2 eV of the results obtained using the d-aug-cc-pVTZ basis set and CC3 method we have to use at least the aug-cc-pVD(T)Z basis set within the CCSD model. At this level, the $n_N \rightarrow 3s$ transition energies should have an accuracy of about 0.1 eV.

The benchmark calculations on methylamine and formic acids can be used to design well-balanced compromises between computational cost and accuracy for calculations on the larger systems. Choosing the aug-cc-pVDZ basis set for hydrogen, carbon, and oxygen atoms together with the aug-cc-pVTZ basis set for the nitrogen atoms in combination with the CCSD model will presumable give results which deviate less than 1% as compared to the triple-corrected d-aug-cc-pVTZ results for the $n_O \rightarrow \pi^*_{CO}$, $n_N \rightarrow 3s$, and $\pi_{C=O} \rightarrow \pi^*_{CO}$ transitions and within a few percent for the remaining Rydberg excitations considered in this study. This particular compromise, denoted CCSD/aD-(T), shall be used in many of the calculations in the following sections.

Finally, comparing Figure 1 and Figure 3, we note that the excitations within the two groups start at about the same energy (6 eV). However, it is seen that in the interval from 6 to 8 eV the transitions following are mainly from the amino group as only one transition in this interval comes from the acid group. This means that if both groups are present in a molecule (as in an amino acid) we would expect that most transitions following the $n_O \rightarrow \pi^*_{CO}$ and $n_N \rightarrow 3s$ transitions originate from the amino group if no interaction occurs between the groups changing the energy levels of the electronic states considerably.

To get a better understanding of the nature of the lowest lying transitions in the amino acids glycine and alanine, we have performed CCSD/aD(T) calculations of the vertical excitation energies ω below and near 8 eV of the carboxylic acids and amines. As the size of these structures allows for calculations using the d-aug-cc-pVTZ (daT) basis set within the CCSD model, we calculate a basis set correction Δ_B as the difference in the values of an excitation energy calculated using CCSD/daT and CCSD/aD(T). (For ethylamine, we had to use a combination of the aug-cc-pVTZ (aT) and daT basis sets instead of the daT basis set on all atoms. Tests performed on methylamine showed that using the daT basis set on N and H atoms and the aT basis set on C atoms gave results identical to results obtained using the daT on all atoms within the number of digits reported here.) The triple-corrected CCSDR(3) model becomes very computationally expensive for all but the smallest molecules when using the large daT basis set, and we calculate

TABLE 1: CCSD/aD(T) Excitation Energies (ω in eV) and Oscillator Strengths (f) of Formic Acid, Acetic Acid, and Propionic Acid[a]

		theoretical vertical excitations						experimental data	
		ω	Δ_T[b]	Δ_B[c]	ω^{Est}	f	0−0 excitation, ω^{Est}	ϵ_{max}	ϵ_{0-0}
					HCOOH				
$1^1A''$	$n_O \rightarrow \pi^*_{CO}$	5.89	−0.01	−0.02	5.86	0.00068	4.71	5.7−5.8[d,e,f]	4.64[g] 4.78[h]
$2^1A'$	$n_O \rightarrow 3sa'$	7.76	−0.08	0.14	7.82	0.02069		7.5−7.8[e,i,j]	
$3^1A'$	$n_O \rightarrow 3pa'$	8.30	−0.02	0.14	8.42	0.06442		8.3[f] 8.9−9.0[e,i,j]	8.1[f] 8.84[j]
$4^1A'$	$\pi_{C=O} \rightarrow \pi^*_{CO}$	8.69	−0.09	0.01	8.61	0.18799		8.29[j] 8.4[e,f]	8.11[j]
$2^1A''$	$\pi_{C=O} \rightarrow 3sa'$	8.53	−0.02	0.16	8.67	0.00396			8.76[j]
					CH_3COOH				
$1^1A''$	$n_O \rightarrow \pi^*_{CO}$	5.98	−0.02	−0.11	5.85	0.00032		5.8−5.9[e,k,l]	
$2^1A'$	$n_O \rightarrow 3sa'$	7.27	−0.07	0.11	7.31	0.04990		7.1[e] 7.21[k]	
$3^1A'$	$n_O \rightarrow 3pa'$	8.17	−0.07	0.09	8.19	0.01931		8.5[e]	
$2^1A''$	$\pi_{C=O} \rightarrow 3sa'$	8.26	−0.03	0.21	8.44	0.00323			
					CH_3CH_2COOH				
$1^1A''$	$n_O \rightarrow \pi^*_{CO}$	6.01	−0.02	−0.01	5.98	0.00014		6.02[l]	
$2^1A'$	$n_O \rightarrow 3sa'$	7.27	−0.06	0.18	7.39	0.05016			
$3^1A'$	$n_O \rightarrow 3pa'$	8.10	−0.09	0.13	8.14	0.00812			
$2^1A''$	$\pi_{C=O} \rightarrow 3sa'$	8.24	−0.03	0.17	8.38	0.00365			

[a] Additivity of basis set, Δ_B, and triple excitation, Δ_T, corrections are assumed in constructing the best estimates, ω^{Est}, for the excitation energies. The excitation energies are ordered after increasing ω^{Est} values. [b] Δ_T = CCSDR(3)/aD(T)-CCDS/aD(T). [c] Δ_B = CCSD/daT-CCSD/aD(T). [d] Reference 18. [e] Reference 20. [f] Reference 17. [g] Reference 19. [h] Reference 15. [i] Reference 16. [j] Reference 22. [k] Reference 14. [l] Reference 21.

a triple correction using the aD(T) basis set and define Δ_T as the difference in the calculated value of an excitation energy using CCSDR(3)/aD(T) and CCSD/aD(T) results. From these corrections, we calculate a best estimate (ω^{Est}) as the sum of the CCSD/aD(T) calculated excitation energy, the basis set correction, and the triple correction.

To test how well Δ_T calculated in this way recovers Δ_T calculated using a larger basis set and a more expensive triple model, we performed two tests on the smallest structure within each group of molecules, namely, formic acid and ammonia. First, we calculated the excitation energies of these two molecules using the iterative CC3/aD(T) model and compared the results with those of the CCSDR(3)/aD(T) model. In the case of ammonia, the CCSDR(3) model reproduced the energies of CC3 perfectly, and in the case of formic acid they deviate by less than 1% in the worst case.

Second, we calculated the triple correction of ammonia and formic acid using the large daT basis set and compared the results with those obtained with the aD(T) basis set. In the case of ammonia, the triple corrections were almost the same, and as the corrections are very small we conclude that the triple correction calculated using the smaller basis set gives almost identical results as obtained from calculations using both larger basis sets and more expensive triple models.

For formic acid, the triple correction calculated from the smaller basis set underestimates the triple correction. However, the triple correction calculated with the smaller basis set lowers the excitation energy as does the correction calculated using the larger basis set. For example, the triple correction of the $n_O \rightarrow \pi^*_{CO}$ transition, which is the worst case, was calculated to be −0.04 using the daT basis set whereas −0.01 was obtained using the aD(T) basis set. Similarly, for the $n_O \rightarrow 3s$ transition the daT triple correction was calculated to be −0.13 and only −0.08 was obtained using the aD(T) basis set. For both basis sets, the difference in contribution is relatively small and therefore contributes only moderately to the uncertainty in the final estimate.

We conclude that the estimated excitation energies (ω^{Est}) obtained from adding the CCSD/aD(T) result with the basis set correction and triple correction are very good approximations to the real CCSDR(3)/daT values, deviating no more than 0.5% from these values for both amines and carboxylic acids. However, the values are obtained at a much lower computational cost. Finally, the size of the corrections can be taken as a rather conservative error estimate for the final estimate.

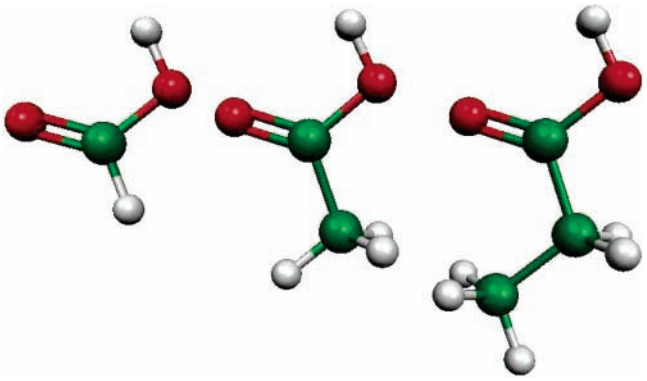

Figure 5. The carboxylic acids. Structures from the left are formic acid, acetic acid, and propionic acid.

B. The Carboxylic Acids. In Table 1, we have collected the three lowest excitation energies in the A' symmetry and the two lowest in the A'' symmetry of formic acid and the two lowest in both symmetries of acetic acid and propionic acid. These three carboxylic acids can all be constructed as substructures of alanine. The carboxylic acids studied are shown in Figure 5. All the geometries have C_s-symmetry with the carboxyl group lying in the mirror-plane. In Table 1, ω refers to the CCSD/aD(T) calculated vertical excitation energies and ω^{Est} is the estimated result obtained from adding Δ_T and Δ_B to the CCSD/aD(T) value. The term f is the CCSD/aD(T) oscillator strength of the transition. Note that the excitation energies in Tables 1 and 2 are ordered according to the ω^{Est} values.

From Table 1, we first of all note that the four lowest (noncorrected) excitation energies of the three acid molecules are all assigned to the same transitions. Comparing the CCSD/aD(T) excitation energies belonging to the same assigned transition within the different structures, it is seen that the energies change when going from formic acid to acetic acid by some tenth of an electronvolt. However, the energies are seen to converge when going from acetic acid to propionic acid as the energies are almost the same for these molecules. The same

trend is observed for the estimated excitation energies (ω^{Est}) except for the $n_O \rightarrow \pi^*_{CO}$ transition when going from acetic acid to propionic acid. Because of a considerable negative basis set correction to the transition in acetic acid, the energy of the transition becomes very close to that of formic acid.

In Table 1, we have collected experimental data from several groups. Comparing the calculated vertical excitation energies with the energy of maximum absorption from experiment, we find good agreement for the two lowest lying excitation energies, with the calculated results being within 0.1 eV of the band maxima. We should remember that the calculated values do not include any correction for nuclear motion, for example, the difference in zero-point energy in the ground- and excited state. This is a major factor of uncertainty in the comparison. Ari et al.[20] measured the $n_O \rightarrow \pi^*_{CO}$ transition in both formic acid and acetic acid and found ϵ_{max} at 5.8 eV in both cases, showing great resemblance with the estimated (ω^{Est}) result in Table 1. The value found for this transition in acetic acid and propionic acid by Hintze et al.[21] also compares well with our results.

In Table 1, we also report the calculated lowest singlet 0−0 excitation energy of formic acid obtained using CCSD/aD(T) methods for calculating equilibrium structures and harmonic vibrational frequencies. The CCSD/aD(T) 0−0 excitation energy has been corrected by the small Δ_T and Δ_B corrections used for the vertical excitation energy. Due to rather large structural changes[19,27] when going from the ground-state geometry to the first excited singlet state geometry, the 0−0 transition occurs at an energy around 1 eV lower than the vertical excitation energies calculated. This is in agreement with the large difference between ϵ_{max} and the 0−0 transition found experimentally. We find a perfect match (within 0.1 eV) when comparing our calculated value with the experimental data also listed in Table 1. This adds credibility to the calculated results presented here since the comparison between theory and experiment for 0−0 transition energies is free of ambiguities.

In summary, we have so far seen that the calculations and the experimental results support each other for these lower lying states.

Proceeding to energetically higher lying states, we find our best estimate for the vertical excitation energy of the second Rydberg excitation $n_O \rightarrow 3pa'$ of formic acid deviates as much as half an electronvolt from the energy of maximum absorption reported in recent experimental studies.[20,22] This discrepancy is far outside the expected error bars of the theoretical calculations of vertical excitation energies. We have in the estimated vertical excitation energies accounted for the effects of rather large basis sets including diffuse functions as well as accurate inclusion of triple excitations. Both effects were modest. The deviation is furthermore likely to be too large to be ascribed to nuclear motion. For example, there is no similar discrepancy for the $n_O \rightarrow 3s$ state which is a Rydberg excitation from the same orbital. Thus the assignment of a peak close to 9 eV to this state cannot be supported from our calculations, and in fact the experimental assignment is far from trivial in this region of energy as discussed below.

In carbonyl compounds, a very strong $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition occurs around 8 eV[62,63] which dominates the spectrum and complicates the task of assigning the other transitions in this region as these are far less intensive. Leach et al.[22] assigned for formic acid a progression of bands in the interval 8−8.8 eV with 0−0(ϵ_{max}) absorption located at 8.107(8.29) eV to the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transitions. A similar assignment was proposed by Ari et al.[20] who reported ϵ_{max} at 8.4 eV. The $n_O \rightarrow 3p$ transition was assigned by Leach et al.[22] to begin at 8.839 eV and by Ari et al.[20] to have a maximum absorption at 8.9 eV. However, Fridh[17] assigned the progression in the 8−8.8 eV interval to the $n_O \rightarrow 3p$ Rydberg series lying on top of the strong $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition with a maximum absorption around 8.3−8.5 eV whereas ϵ_{max} of $\pi_{C=O} \rightarrow \pi^*_{CO}$ was found at 8.4 eV. Our estimated result of 8.42 eV for the $n_O \rightarrow 3pa'$ transition obviously agrees much better with the work of Fridh than with the more recent experiments by Ari et al.[20] and by Leach et al.[22]

Our calculated $\omega(\omega^{Est})$ value of the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition in formic acid is 8.69(8.61) eV with an oscillator strength of 0.19. The calculated oscillator strength agrees with the value of 0.2 reported by Leach et al.[22] The situation with many states overlapping in the same energy region obviously complicates the interpretation of experiments including the assignment of maximum absorption peaks to definite electronic transitions and integrated absorption to oscillator strengths, and the agreement for the oscillator strength may be fortuitous.

The basis set correction for the $\pi_{C=O} \rightarrow \pi^*_{CO}$ excitation is negligible, and the difference between ω and ω^{Est} comes from Δ_T. This is different from the other excitation energies reported in Table 1 where Δ_B is the dominating part of the correction. From the tests on Δ_T previously discussed, we expect this correction to be underestimated and the true excitation energy may be located toward slightly lower energies. This is actually the case as the CCSDR(3)/daT value is 8.58 eV. Analyzing the response eigenvector corresponding to this excitation, we find a considerably mixing of the dominating $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition with the $n_O \rightarrow 3pa'$ among other single excitations. This is also observed in the semiempirical calculations by Demoulin.[24] Demoulin also finds the $n_O \rightarrow 3pa'$ transition to be located at lower energy than the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition. Comparing the calculation of the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition of formic acid in this work with experimental data, we find that the calculated vertical energy lies about 0.2 eV on the high side of the reported interval 8.3−8.4 eV.[14,17,20,22] However, an overestimation of this size can be explained as primarily due to nuclear motion, in particular the neglect of expected reduction in zero-point vibration corrections due to the bonding to antibonding transition, as well as the overlapping of electronic bands in the spectra.

Returning to the excitation energies reported in Table 1, we find that the second transition of $^1A''$ symmetry in the carboxylic acids is a $\pi_{C=O} \rightarrow 3s$ transition. The estimated ω^{Est} value is higher than the calculated ω value because of a significant positive basis set correction. This actually causes the ω^{Est} value of the $\pi_{C=O} \rightarrow 3s$ transition to change order with the ω^{Est} value of the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition in Table 1 (as compared to the ω values) as this transition was lowered by a considerable triple correction. The ω^{Est} values are almost degenerate in energy. Comparing the ω^{Est} value for the $\pi_{C=O} \rightarrow 3s$ transition of formic acid with the value of 8.76 eV for the 0−0 transition presented by Leach et al.,[22] we find again some disagreement with the interpretation of this experiment. The calculated vertical excitation energy lies at a lower energy than the experimental 0−0 transition energy, and this is not what should be expected as geometrical relaxation in the excited state and zero-point vibrational energy corrections should decrease the theoretical value even further.

Concerning other theoretical work on formic acid, we find some agreement on the assignments found in this work with significantly older results. Iwata and Morokuma[23] used a two-configuration electron−hole method to calculate the vertical excitation energies of the $n_O \rightarrow \pi^*_{CO}$ and $\pi_{C=O} \rightarrow \pi^*_{CO}$ transitions. They found 5.83 and 9.84 eV, respectively, with

TABLE 2: CCSD/aD(T) Excitation Energies (ω in eV) and Oscillator Strengths (f) of Ammonia, Methylamine, and Ethylamine[a]

		theoretical vertical excitations						experimental data	
		ω	Δ_T[b]	Δ_B[c]	ω^{Est}	f	0−0 excitation, ω^{Est}	ϵ_{max}	ϵ_{0-0}
					NH_3[d]				
2^1A_1	$n_N \rightarrow 3sa_1$	6.64	−0.03	0.01	6.62	0.08571	5.69	6.39[e,f,g]	5.72[e] 5.73[h]
1^1E	$n_N \rightarrow 3pe$	8.22	−0.01	−0.04	8.17	0.00688		8.18[e] 7.93[f]	7.34[h] 7.46[e]
3^1A_1	$n_N \rightarrow 3pa_1$	9.50	−0.02	−0.88	8.60	0.00263		(8.26[f])	7.92[h]
4^1A_1	$n_N \rightarrow 4sa_1$	10.38	−0.02	−1.21	9.15	0.01504		9.11[i]	8.66[h]
2^1E	$n_N \rightarrow 3de$	10.44	−0.02	−1.22	9.20	0.05980		9.27[i]	8.69[h]
3^1E	$n_N \rightarrow 4pe$	11.41	−0.03	−1.75	9.63	0.00710		9.75[i]	8.84[h]
					CH_3NH_2				
$2^1A'$	$n_N \rightarrow 3sa'$	5.97	−0.01	0.02	5.98	0.02380	5.16	5.7[j] 5.77[e]	5.17[e] 5.18[j]
$3^1A'$	$n_N \rightarrow 3pa'$	7.11	−0.03	−0.01	7.07	0.01713	}		
$1^1A''$	$n_N \rightarrow 3pa''$	7.13	−0.01	−0.02	7.10	0.00196	}	7.13[e] 7.2[j]	6.22[j] 6.24[e]
$4^1A'$	$n_N \rightarrow 3pa'$	7.55	−0.01	−0.15	7.39	0.03948	}		
$5^1A'$	$n_N \rightarrow 4sa'$	8.44	−0.02	−0.49	7.93	0.00338			
$2^1A''$	$n_N \rightarrow 4pa''$	8.31	−0.01	−0.20	8.10	0.00014		8.7[j]	7.88[j]
					$CH_3CH_2NH_2$				
$2^1A'$	$n_N \rightarrow 3sa'$	6.04	−0.03	0.03	6.04	0.03648		5.8[j] 5.83[e]	5.21[j] 5.32[e]
$1^1A''$	$n_N \rightarrow 3pa''$	6.99	−0.02	0.00	6.97	0.00502	}	7.0[j] 7.01[e]	
$3^1A'$	$n_N \rightarrow 3pa'$	7.07	−0.03	−0.01	7.03	0.01480	}		
$4^1A'$	$n_N \rightarrow 3pa'$	7.44	−0.05	−0.07	7.32	0.00303			
$5^1A'$	$n_N \rightarrow 4sa'$	8.12	−0.04	−0.31	7.77	0.04002	}	7.9[j]	
$2^1A''$	$n_N \rightarrow 4pa''$	8.12	−0.02	−0.14	7.96	0.00005	}		

[a] Additivity of basis set, Δ_B, and triple excitation, Δ_T, corrections are assumed in constructing the best estimates, ω^{Est}, for the excitation energies. The excitation energies are ordered after increasing ω^{Est} values. [b] Δ_T = CCSDR(3)/aD(T)-CCDS/aD(T). [c] Δ_B = CCSD/daT-CCSD/aD(T). [d] Calculations are done in C_s, but C_{3v} nomenclature is used. Oscillator strength is the sum of the two degenerate states. [e] Reference 28. [f] Reference 29. [g] Reference 30. [h] Reference 32. [i] Reference 31. [j] Reference 34.

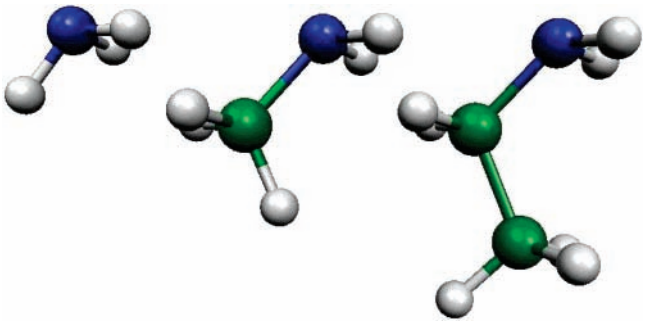

Figure 6. The amines. Structures from the left are ammonia, methylamine, and ethylamine.

the lower excitation energy in good agreement with this work. Older calculations by Demoulin[24] show much resemblance with the results of this work with regard to the ordering of states but typically more than half an electronvolt difference for the absolute values of the excitation energies.

C. The Amines. In Table 2, we have collected the excitation energies below and around 8 eV of ammonia, methylamine, and ethylamine, the amines which can be constructed as substructures of alanine. All structures have C_s-symmetry with a mirror-plane through the backbone structure of the molecule. The structures of the amines are illustrated in Figure 6. Throughout we handle ammonia as an amine and we note that the calculations performed on ammonia have been carried out in C_s-symmetry. In Table 2, ω refers to the CCSD/aD(T) calculated vertical excitation energies and ω^{Est} to the estimate obtained from adding the basis set and triple corrections to the CCSD/aD(T) result.

We have reported the five lowest excitation energies for ammonia. The states 3*pe*, 3*de*, and 4*pe* are double degenerated, and the oscillator strengths reported for these states are the sum of the oscillator strengths for the degenerated states. Comparing the CCSD/aD(T) calculated vertical excitation energies ω with the ω^{Est} values for the two lowest excitations of ammonia, we find that the changes are moderate (within 0.05 eV), showing that these two states in ammonia are described to a satisfactory degree of accuracy within the CCSD/aD(T) model. The ω^{Est} values are around 0.2 eV higher than the position of the experimental band maxima. This can be explained by the neglect of the nuclear motion in the ground state and the excited state in the theoretical vertical calculations.

For the $n_N \rightarrow 3sa'$ transition, we have also calculated the 0−0 excitation energy. The number listed in Table 2 is corrected with the small numbers for Δ_B and Δ_T used to correct the vertical excitation energies. From Table 2, it is seen that the 0−0 excitation energy is almost 1 eV lower than the calculated vertical excitation energy. This can be understood from the large geometrical changes in the ammonia molecule when going from the pyramidal ground-state structure to the 3s Rydberg state which is planar as are all the Rydberg states of ammonia considered in this study.[37] This is also in agreement with the large difference between the experimental values of ϵ_{max} and ϵ_{0-0}, and we find that the calculated 0−0 excitation energy matches the experimental data perfectly as the theoretical and experimental values differ by less than 0.05 eV. As for the carboxylic acids, this observation adds credibility to the calculated results since the comparison between theory and experiment for 0−0 transition energies is free of ambiguities.

Considering the transitions to energetically higher lying excited states, we observe from Table 2 that considerable basis set corrections are added to get our best estimate. This clearly shows that the higher excited states of ammonia are poorly described with the aD(T) basis set. The corresponding excited states in the larger amine molecules are found to have faster basis set convergence and to be much better represented in the aD(T) basis set. To investigate if the ω^{Est} values are converged for ammonia, we did some calculations of the excitation energies of ammonia using the t-aug-cc-pVTZ (201 basis functions) and the d-aug-cc-pVQZ (291 basis functions) basis sets. The calculations showed that the ω^{Est} values in Table 2 only deviated within 0.04 eV from the t-aug-cc-pVTZ results and within 0.08

eV from the d-aug-cc-pVQZ results. As in the case when going from the x-aug-cc-pVDZ series to the x-aug-pVTZ series of basis sets, as observed in Figure 3, the excitation energies are raised when going from the x-aug-cc-pVTZ series of basis sets to the d-aug-cc-pVQZ basis set and the d-aug-cc-pVQZ results therefore lie at higher energies than the ω^{Est} values. However, as the ω^{Est} values deviate less than 0.1 eV from the results obtained from calculations with many more basis functions, we conclude that the ω^{Est} values are close to the basis set limit also for ammonia and therefore give accurate results.

Using the t-aug-cc-pVTZ basis set, the $n_N \rightarrow 3de$ transition energy becomes slightly lower than the $n_N \rightarrow 4sa_1$ transition energy and the transitions therefore change places in the CCSD/t-aug-cc-pVTZ calculated spectrum. However, the transitions are nearly degenerate around 9.16 eV.

Comparing the ω^{Est} values of the transition energies for the higher excited states of ammonia, we generally find that the calculated energies deviate up to about 0.1 eV from most of the experimental ϵ_{max} values. For the $n_N \rightarrow 3pa'$ transition, an excitation of 8.26 eV has sometimes been cited for ϵ_{max}. This value does not agree with the present calculations, and it does not agree with the trends for the experimental difference between the experimental ϵ_{max} and ϵ_{0-0} values of order 0.6−0.9 eV for the Rydberg excitations with the same cation core. Indeed, it is not clear how a 8.26 eV value for ϵ_{max} can be extracted from the experiment where there is significant overlap between the electronic bands.[29]

Calculations of the electronic spectrum of ammonia have been reported in the literature for more than two decades.[35,36] In 1991, Chantranupong et al. reported calculations on both vertical and adiabatic excitation energies of ammonia using a CI method with a basis set including additional s, p, and d Rydberg atomic orbitals (AOs) and a total of 52 basis functions. They obtained for the vertical (adiabatic) excitation energies 6.56 eV (5.88 eV), 7.89−8.01 eV (7.10−7.21 eV), 8.46 eV (7.65 eV), 9.02 eV (8.26 eV), 9.07 eV (8.25−8.36 eV), and 9.41 eV (8.50−8.64 eV) for the $n_N \rightarrow 3sa_1$, $n_N \rightarrow 3pe$, $n_N \rightarrow 3pa_1$, $n_N \rightarrow 4sa_1$, $n_N \rightarrow 3de$, and $n_N \rightarrow 4pe$ transitions, respectively.[37] We see that for all states large differences between vertical and 0−0 transition energies are obtained, as also found in our calculations. Over the energy range discussed in this paper, we find a much better agreement with experiments than the previous calculations, which of course is a result of the much more advanced basis sets and correlation methods used in this work, compared to the much earlier calculations. Seen in the light of the tests we performed using the very large d-aug-cc-pVQZ basis set showing that the ω^{Est} results are close to the basis set limit, we expect the results of ammonia presented here to be accurate to about 0.1 eV.

Concerning methylamine and ethylamine in Table 2, we find that the CCSD/aD(T) calculated vertical excitation energies compare much better with the ω^{Est} values than in the case of ammonia. For excitation energies of transitions to Rydberg states of principal quantum number 3, ω and ω^{Est} differ less than 0.2 eV. For transitions to states of principal quantum number 4, the deviation becomes more significant. This is also what we expect as the states become more diffuse and therefore need more diffuse basis functions in the description when increasing the principal quantum number. The better representation of the Rydberg states in the larger molecules comes from the additional basis functions from the additional atoms present.

From Table 2, we first of all note that the ordering of the transitions in methylamine and ethylamine is the same. However, some configurational mixing is found in the transitions and especially the second $n_N \rightarrow 3pa'$ and the $n_N \rightarrow 4sa'$ transitions are heavily mixed. The assignments express only qualitative information and are in many cases an oversimplification. The assignments used here agree concerning the basis set correction as discussed above, but they disagree with respect to the oscillator strength where the two transitions interchange oscillator strengths when going from methylamine to ethylamine.

Comparing the calculated lowest vertical excitation energy for methylamine and ethylamine with the energy of maximum absorption found in experiments, agreement within 0.3 eV is found, which is similar to the case for the lowest excitation energy in ammonia. For methylamine, the 0−0 excitation energy has been explicitly calculated for the $n_N \rightarrow 3sa'$ transition. The experimental 0−0 energy is within 0.02 eV of the theoretical prediction. As previously discussed, the comparison of 0−0 transition energies between theoretical calculations and experimental spectra is free of the problems encountered in a similar comparison for vertical excitation energies. It is gratifying to see this reflected in a much better agreement between calculations and experiment. We interpret this as supporting the high accuracy (for the lowest transition energy an accuracy within 0.1 eV as compared to the exact theoretical result) of the calculated vertical energy difference between the electronic ground state and an excited electronic state *both* at the ground-state geometry. However, this theoretical quantity is simply not what can be read directly from the experimental energy of maximum absorption.

Concerning the next excitation energies within the two amines, we observe some differences between the ω and the ω^{Est} values. For methylamine, the ordering of the states according to the ω^{Est} values and the CCSD/aD(T) is different from the ω ordering of states as the $n_N \rightarrow 4sa'$ is found at lower energy than the $n_N \rightarrow 4pa''$ transition only for the ω^{Est} values (as is expected as the angular momentum quantum number is larger in the latter case, typically giving a larger energy of the orbital). For ethylamine a similar observation is made.

Experimentally, a second absorption maximum is found at 7.1−7.2 eV for methylamine.[28,34] We believe that this absorption band has contributions from several transitions which we have illustrated with the bracket in Table 2. From Table 2, we see that the two strongest transitions in methylamine contributing to this absorption maximum are the two $n_N \rightarrow 3pa'$ transitions with oscillator strengths of 0.017 and 0.039 located at energies of 7.07 and 7.39 eV. These transitions are followed by relatively weak transitions. This qualitatively agrees with the spectrum of methylamine in ref 34 where the absorption rises to a maximum located at 7.13 eV and then falls off again for higher energies showing an almost Gaussian behavior for the absorption curve.

The ethylamine case is different. In Table 2, we have illustrated the contributing transitions to the second absorption maximum with a bracket. From Table 2, we see that only the first $n_N \rightarrow 3pa'$ transition contributes to the absorption maximum with maximum at 7.0 eV[34] and that the other contributing transitions around this energy are relatively weak transitions. The stronger $n_N \rightarrow 4sa'$ transition follows at a higher energy. Qualitatively, this results in a spectrum with an absorption curve which rises to a maximum. However, after this first maximum the absorption does not decrease like a Gaussian as found for methylamine. Dependent on the overlap with the previous transitions, the curve rises in some way to another maximum. This is actually what is observed in the ethylamine spectrum in ref 34. The absorption maximum, which is located at lower energy than was the case for methylamine and not quite as strong an absorption either, is followed by a small decrease in

TABLE 3: CCSD/aD(T) Vertical Excitation Energies (ω in eV) and Oscillator Strengths (f) for Alanine and Glycine with and without Protonation of the Amino Group[a]

	$\omega(f)$ at the indicated excitations								
molecule	$n_O \rightarrow \pi^*_{CO}$	$n_N \rightarrow 3s$	$n_O \rightarrow 3s$	$n_N \rightarrow 3p$	$n_N \rightarrow 3p$	$n_N \rightarrow 3p$	$n_O \rightarrow 3p$	$\pi_{C=O} \rightarrow 3s$	$\pi_{C=O} \rightarrow \pi^*_{CO}$
Gly	5.88(0.00053)	6.32(0.01247)	7.24(0.04608)	7.48(0.01535)	7.61(0.03162)	7.81(0.01996)	8.00(0.03399)	8.36(0.00174)[b]	8.80(0.07514)[c]
Ala	5.96(0.00079)	6.64(0.01417)	7.26(0.08320)	7.46(0.00341)	7.55(0.00335)	7.88(0.00094)	8.05(0.01643)	8.30(0.00648)	8.79(0.04528)
Gly^+	5.95(0.00013)		7.64(0.00611)					8.69(0.01179)[d]	8.70(0.14457)
Ala^+	6.28(0.00057)		8.22(0.04093)					9.26(0.00023)	8.52(0.11642)

[a] Many transitions of the neutral glycine and alanine molecules with excitation energies larger than 8 eV have been omitted from the table. [b] The $\pi_{C=O} \rightarrow 3s$ transition is found to mix with other transitions, and considerable contributions from this transition are also found in a transition occurring at $\omega(f) = 8.27(0.01672)$ eV. [c] The $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition in Gly is found to mix with other transitions, and considerable contributions from this transition are also found in a transition at $\omega(f) = 8.57(0.03918)$ eV. [d] The $\pi_{C=O} \rightarrow 3s$ transition in Gly^+ is found to mix with the same transition as in Gly. This also explains the larger oscillator strength compared to the same transition in the other molecules.

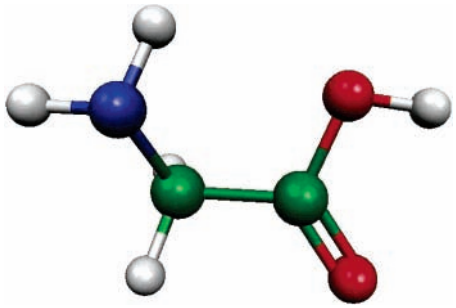

Figure 7. The structure of the amino acid glycine.

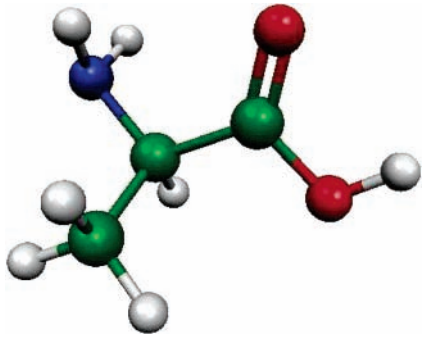

Figure 8. The structure of the amino acid L-α-alanine.

absorption, and then the curve starts to rise to a new maximum located somewhere around 7.8−7.9 eV.

In summary, we find that the calculations presented in Table 2 explain many features of the experimental UV−VUV spectra of methylamine and ethylamine presented in the literature,[28,34] including aspects in which the two molecules differ.

D. The Amino Acids. Having investigated the nature of the carboxylic group and the amino group in detail in the previous sections, we now discuss the calculated CCSD/aD(T) vertical excitation energies of glycine (Gly), alanine (Ala), the protonated glycine (Gly^+), and the protonated alanine (Ala^+). The structures of the neutral amino acids Gly and Ala are shown in Figures 7 and 8, respectively. The calculated excitation energies are listed in Table 3. From the assignments of the transitions in Table 3, it is seen that this gives excitation energies for transitions for which we have established that the CCSD/aD(T) method gives results which are comparable with the results obtained using the daT basis set and corrected for triple excitations. We therefore suspect the presented spectrum of the amino acids in Table 3 to be fairly accurate. Especially, as we have seen from the previous calculations, the performance of the CCSD/aD(T) method becomes better with the size of the molecule as the additional basis functions from the additional atoms contribute to the description of the diffuse states in the molecule. Thus, an accuracy of 0.1−0.2 eV for the theoretical vertical transitions is expected, but the difference between the theoretical vertical transitions and the experimental band maxima may add additional uncertainty in relation to direct comparison with experimental band maxima. To obtain the $\pi_{C=O} \rightarrow 3sa$ and the $\pi_{C=O} \rightarrow \pi^*_{CO}$ excitation energies, a large number of excitation energies above 8 eV have been calculated but are omitted from Table 3.

Concerning the neutral amino acids in Table 3, we find that the transitions up to 8 eV can be assigned to the same transitions in both Gly and Ala, respectively. Generally, we find that the order of the transitions follows the order which was found for the groups of carboxylic acids and amines.

The lowest excitation energy is assigned to the $n_O \rightarrow \pi^*_{CO}$ transition in the carboxylic group. In Table 3, we observe that this transition is very weak as was also the case in the carboxylic acids previously studied. The excitation energy for this transition is located at somewhat the same energy in the amino acids as it was for the carboxylic acids. From the backbone structure of the amino acid, the Gly molecule should be compared with acetic acid whereas the Ala molecule should be compared with propionic acid. We observe that also for the amino acids we find the $n_O \rightarrow \pi^*_{CO}$ transition of the acid group attached to the longest carbon chain to be located at the highest energy.

The second excitation energy is assigned to the $n_N \rightarrow 3s$ transition from the amino group. From Table 2 and Table 3, it is seen that the excitation energy of this transition is located at higher energies in the amino acids than in the amines. Comparing the oscillator strengths of this transition in Gly and Ala in Table 3 with the sizes of the oscillator strengths of methylamine and ethylamine in Table 2, we find that the oscillator strengths are of the same order of magnitude. Similar to the lowest lying transition in the amino acids, we find that the $n_N \rightarrow 3s$ transition in Ala is located at a higher energy than the same transition in Gly, following also the trend observed for the amines that the transition from the amino group attached to the longest carbon chain is located at the highest energy. However, the difference between the excitation energies for this transition is much more pronounced in the amino acid case than for the amines.

Considering the third transition in Table 3, we have assigned it to be a $n_O \rightarrow 3s$ transition. However, some mixing is observed and considerable contributions from $n_N \rightarrow 3p$ excitation are observed too. The excitation energy for this transition is almost the same in the two amino acids as in the case for this transition in acetic acid and propionic acid. Comparing the numbers in Table 1 and Table 3, we find that the oscillator strength for the $n_O \rightarrow 3s$ transition is of the same order of magnitude in the amino acids as in the carboxylic acids.

The next three excitation energies in Gly and Ala come from the $n_N \rightarrow 3p$ transitions from the amino group. Comparing their location with the analogous location in the amines, we note from

Table 2 and Table 3 that the excitation energies of these transitions have moved toward higher energies in the amino acids. That was also the case for the $n_N \rightarrow 3s$ transition discussed previously. The first two of these three transitions, which should be compared to the first $n_N \rightarrow 3pa'$ and the $n_N \rightarrow 3pa''$ transitions in the amines, are located close together as in the amine spectra. The third transition is separated slightly from the other two in the amino acids as in the amines.

Concerning the oscillator strengths of these transitions, we find some disagreements between the observations made for the amines discussed previously and the values reported in Table 3. We primarily compare Gly with methylamine and Ala with ethylamine as they share the same carbon chain backbone structure. First, for Gly we find that the oscillator strength of 0.015 for the excitation energy of 7.48 eV compares well with oscillator strength found for the first $n_N \rightarrow 3pa'$ transition in methylamine in Table 2. However, the oscillator strengths of the other two transitions of 0.032 and 0.020 compare well with the oscillator strength of the second $n_N \rightarrow 3pa'$ transition in methylamine, but not with the weak $n_N \rightarrow 3pa''$ transition.

For Ala, it is the other way around as the oscillator strengths of the three transitions are too weak to compare with the first $n_N \rightarrow 3pa'$ transition in ethylamine. However, the oscillator strength of the first two transitions in Ala compares well with the weak $n_N \rightarrow 3pa''$ and the second $n_N \rightarrow 3pa'$ transition in ethylamine and we note that the very weak oscillator strength of the third of these transitions in Ala compares well with the magnitude of the $n_N \rightarrow 4pa''$ oscillator strength for ethylamine in Table 2. We find significant mixing between contributions from nitrogen p orbitals in all directions as well as from some nitrogen s orbitals. The lack of symmetry in the amino acid structures may lead to difficulties in assigning the excitation energies rigorously to one specific orbital transition.

The excitation energy in Table 3 that has been assigned to a $n_O \rightarrow 3p$ transition has some mixing especially with the $n_N \rightarrow 3p$ transitions. However, as in the case for the previous discussed transitions in Gly and Ala from the carboxylic acid group, the excitation energy as well as the oscillator strengths compare well with the observations made for the carboxylic acids in Table 1. The excitation energy in Gly and Ala is almost the same (8.0 eV), which is expected from the comparisons between acetic acid and propionic acid. The magnitude of the oscillator strengths is also comparable with the oscillator strengths observed in Table 1 for this transition in carboxylic acids.

In Table 3, we have also given the excitation energies of the $\pi_{C=O} \rightarrow 3sa$ and the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transitions in the amino acids. Many excitations above 8 eV have been omitted from Table 3. Comparing the $\pi_{C=O} \rightarrow 3sa$ transition of Gly and Ala in Table 3, we find that both the location around 8.3 eV and the small oscillator strength compare well with the observations made for the carboxylic acids in Table 2. The $\pi_{C=O} \rightarrow 3sa$ transition is found to be located at lower energy when the carbon chain backbone structure becomes larger in both the amino acids as well as the carboxylic acids. Some mixing is found in Gly for this transition, and there are considerable contributions of this component qualitative for a state located at 8.27 eV. However, this transition is much stronger with an oscillator strength of 0.02.

Concerning the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition in the amino acids, we find that it is located at 8.8 eV in both Gly and Ala. Some mixing with other transitions is found for Gly where considerable contributions from the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition are found in a transition located at 8.57 eV with an oscillator strength of 0.04. The oscillator strengths of the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition in the amino acids are somewhat smaller than expected. A decrease in oscillator strength is observed when comparing to the oscillator strength of formic acid in Table 2 and also compared to the protonated amino acids in Table 3. In Gly, the transition is still found to be the strongest of the transitions given in Table 3, but for Ala the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition is found to be a weaker transition than the $n_O \rightarrow 3s$ transition. This is not what we would have expected from the study of formic acid where the $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition is by far the strongest transition in this energy region, and this is also what we find for the protonated amino acid structures in Table 3.

Having discussed the neutral amino acids Gly and Ala, we turn our attention to the protonated structures of the same amino acids. In these structures, a proton has been attached to the amine group electron lone-pair. This captures the electrons in a covalent bonding and removes the transitions in the amino group from this part of the amino acid spectrum.

In Table 3, we have collected the four lowest transitions in the protonated glycine (Gly^+) and alanine (Ala^+) molecules. The assignments of the transitions were much easier in this case as all transitions were located in the carboxylic acid group. The energetically lowest lying transition in these protonated amino acids is still the $n_O \rightarrow \pi^*_{CO}$ transition. Comparing the value of the protonated amino acids with the corresponding value of the neutral amino acids, it is seen from Table 3 that the excitation energies are raised. This is not only the case for the first transition but is a general trend for all the transitions which can be compared in Table 3.

The excitation energy of the valence-state transition $n_O \rightarrow \pi^*_{CO}$ is only slightly affected by the protonation of the amino group. In contrast, major changes are observed for the $n_O \rightarrow 3s$ Rydberg transition. The change in energy of the transition in Ala compared to the protonated Ala^+ seems to be much more pronounced than in the Gly case. For Ala, a change of almost 1 eV is observed upon protonation of the amino group. Considering the oscillator strengths of the two protonated amino acids, we find that the oscillator strength of Ala^+ compares well with the magnitude of the oscillator strengths for this transition in both the neutral Ala molecule in Table 3 as well as the carboxylic acids in Table 1. However, the $n_O \rightarrow 3s$ transition in the Gly^+ molecule is seen to be much weaker than in the neutral Gly molecule and in the carboxylic acids.

As discussed above, the change is even larger for transitions to more diffuse states and we observe from Table 3 that the transitions to the 3*p* Rydberg states are not among the four lowest excitations in the protonated amino acids as they were in the neutral amino acids and in the carboxylic acids. The next transition in the Gly^+ molecule is the $\pi_{C=O} \rightarrow 3s$ transition which is almost degenerate in energy with the very strong $\pi_{C=O} \rightarrow \pi^*_{CO}$ transition. We observe from Table 3 that the $\pi_{C=O} \rightarrow 3s$ transition is located more than 0.5 eV toward higher energies in Ala^+ than in Gly^+. Comparing the oscillator strengths of this transition with the magnitude of the oscillator strengths calculated for the carboxylic acids in Table 1, we find that the value calculated for Gly^+ of 0.012 is stronger and the value of 0.0002 calculated for Ala^+ is weaker than what we found for the carboxylic acids where the oscillator strength was about 0.003. The much larger oscillator strength found for this transition in Gly^+ is most likely due to the mixing of the $\pi_{C=O} \rightarrow 3s$ transition with a much stronger transition as observed in the neutral Gly.

The common form of amino acids in aqueous solution is a zwitterion. The calculations on protonated amino acids show

that amine states occur only at much higher energy when the amino group is protonated. This is likely to be similar in the zwitterionic form of the amino acids. However, it is not appropriate to consider the zwitterionic form in gas-phase calculations and we will therefore not consider the zwitterionic form in explicit calculations here.

IV. Summary

We have presented calculations of electronic excitation energies for a number of small carboxylic acids (formic, acetic, and propionic acid) and amines (ammonia, methylamine, and ethylamine) in the gas phase as well as for neutral and protonated alanine and glycine. We have for the smaller systems, where experimental results are available, been able to support many experimental assignments as well as questioning a few, where discrepancies between theory and experiment fall outside the error bars of the calculations. The calculations provide trends going from the smaller to the larger carboxylic acids and amines, and furthermore they extend these trends to the neutral amino acids in the gas phase where no experimental information is available. For the compounds and states not yet considered experimentally (including all amino acid states), the present calculations stand as predictions for the position of the vertical excitation energies.

Future work will be directed toward modeling and understanding these electronic transitions for molecules in solution. This requires the considerations of a number of additional theoretical problems, such as the description of the solvent and the possibility for different stable forms depending on the pH of the solution, to mention a few. In any case, the present work serves as a reference for understanding the condensed phase result, and it will be an important reference for future experimental and theoretical studies of electronic absorption in common amino acids.

References and Notes

(1) Thøgersen, J.; Keiding, S. R. University of Aarhus. Private communication.

(2) Thøgersen, J.; Jensen, S. K.; Christiansen, O.; Keiding, S. R. *J. Phys. Chem. A* **2004**, *108*, 7483.

(3) Nishino, H.; Kosaka, A.; Hembury, G. A.; Aoki, F.; Miyauchi, K.; Shitomi, H.; Onuki, H.; Yoshihisa, I. *J. Am. Chem. Soc.* **2002**, *124*, 11618.

(4) Cronin, J. R.; Pizzarello, S. *Science* **1997**, *275*, 951.

(5) Engel, M. H.; Macko, S. A.; Silfer, J. A. *Nature* **1990**, *348*, 47.

(6) Engel, M. H.; Macko, S. A. *Nature* **1997**, *389*, 265.

(7) Pizzarello, S.; Cronin, J. R. *Geochim. Cosmochim. Acta* **2000**, *64*, 329.

(8) Podlech, J. *Angew. Chem., Int. Ed.* **1999**, *38*, 477.

(9) Lerner, N. R. *Geochim. Cosmochim. Acta* **1997**, *61*, 4885.

(10) Peltzer, E. T.; Bada, J. L. *Nature* **1998**, *272*, 443.

(11) Bonner, W. A. *Origins Life Evol. Biosphere* **1988**, *21*, 59.

(12) Bonner, W. A. *Top. Stereochem.* **1991**, *18*, 1.

(13) Toxvaerd, S. *J. Chem. Phys.* **2004**, *120*, 6094.

(14) Barnes, E. E.; Simpsom, W. T. *J. Chem. Phys.* **1963**, *39*, 670.

(15) Ng, T. L.; Bell, S. *J. Mol. Spectrosc.* **1974**, *50*, 166.

(16) Bell, S.; Ng, T. L.; Walsh, A. D. *J. Chem. Soc. Trans. 2* **1975**, *1*, 393.

(17) Fridh, C. *J. Chem. Soc., Faraday Trans. 2* **1978**, *74*, 190.

(18) Singleton, D. L.; Paraskevopoulos, G.; Irwin, R. S. *J. Photochem.* **1987**, *37*, 209.

(19) Ioannoni, F.; Moule, D. C. *J. Phys. Chem.* **1990**, *94*, 2290.

(20) Ari, T.; Güven, M. H. *J. Elec. Spec.* **2000**, *106*, 29.

(21) Hintze, P. E.; Simone, A.; Vaida, V. *Chem. Phys. Lett.* **2001**, *343*, 159.

(22) Leach, S.; Schwell, M.; Dulieu, F.; Chotin, J.-L.; Jochims, H.-W.; Baumgärtel, H. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5025.

(23) Iwata, S.; Morokuma, K. *Theor. Chim. Acta* **1977**, *44*, 323.

(24) Demoulin, D. *Chem. Phys.* **1976**, *17*, 471.

(25) Velardez, G. F.; Ferrero, J. C.; Beswick, J. A.; Daudey, J. P. *J. Phys. Chem. A* **2001**, *105*, 8769.

(26) Aloisio, S.; Hintze, P. E.; Vaida, V. *J. Phys. Chem. A* **2002**, *106*, 363.

(27) Beaty-Travis, L. M.; Moule, D. C.; Lim, E. C.; Judge, R. H. *J. Chem. Phys.* **2002**, *117*, 4831.

(28) Tannenbaum, E.; Coffin, E. M.; Harrison, A. J. *J. Chem. Phys.* **1953**, 311.

(29) Harshbarger, W. R. *J. Chem. Phys.* **1971**, *54*, 2504.

(30) Harshbarger, W. R.; Skerbele, A.; Lassettre, E. N. *J. Chem. Phys.* **1971**, *54*, 3784.

(31) Skerbele, A.; Lassettre, E. N. *J. Chem. Phys.* **1965**, *42*, 395.

(32) Glownia, J. H.; Riley, S. J.; Colson, S. D.; Nieman, G. C. *J. Chem. Phys.* **1980**, *73*, 4296.

(33) Taylor, D. P.; Bernstein, E. R. *J. Chem. Phys.* **1995**, *103*, 10453.

(34) Hubin-Franskin, M.-J.; Delwiche, J.; Giuliani, A.; Ska, M.-P.; Motte-Tollet, F.; Walker, I. C.; Mason, N. J.; Gingell, J. M.; Jones, N. C. *J. Chem. Phys.* **2002**, *116*, 9261.

(35) Runau, R.; Peyerimhoff, S. D.; Buenker, R. J. *J. Mol. Spectrosc.* **1977**, *68*, 253.

(36) Rianda, R.; Frueholz, R. P.; Goddard, W. A., III *Chem. Phys.* **1977**, *19*, 131.

(37) Chantranupong, L.; Hirsch, G.; Buenker, R. J.; Kimura, M.; Dillon, M. A. *Chem. Phys.* **1991**, *154*, 13.

(38) Neta, P.; Simic, M.; Hayon, E. *J. Phys. Chem.* **1970**, *74*, 1214.

(39) Vinogradov, I. P.; Dodonova, N, Ya. *Opt. Spectrosc.* **1970**, *30*, 14.

(40) Snyder, P. A.; Vipond, P. M.; Johnson, W. C., Jr. *Biopolymers* **1973**, *12*, 975.

(41) Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409.

(42) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1981**, *76*, 1910.

(43) Christiansen, O.; Koch, H.; Jørgensen, P. *J. Chem. Phys.* **1996**, *105*, 1451.

(44) Christiansen, O.; Koch, H.; Jørgensen, P. *J. Chem. Phys.* **1995**, *103*, 7429.

(45) Christiansen, O.; Koch, H.; Halkier, A.; Jørgensen, P.; Helgaker, T.; Sánchez de Merás, A. *J. Chem. Phys.* **1996**, *105*, 6921.

(46) Christiansen, O.; Halkier, A.; Koch, H.; Jørgensen, P.; Helgaker, T. *J. Chem. Phys.* **1998**, *108*, 2801.

(47) Christiansen, O.; Stanton, J. F.; Gauss, J. *J. Chem. Phys.* **1998**, *108*, 3987.

(48) Stanton, J. F.; Gauss, J. *J. Chem. Phys.* **1996**, *104*, 9859.

(49) Christiansen, O.; Gauss, J.; Stanton, J. F.; Jørgensen, P. *J. Chem. Phys.* **1999**, *111*, 525.

(50) Christiansen, O.; Jørgensen, P. *J. Am. Chem. Soc.* **1998**, *120*, 3423.

(51) Öhrn, A.; Christiansen, O. *Phys. Chem. Chem. Phys.* **2001**, *3*, 730.

(52) Stanton, J. F. *J. Chem. Phys.* **1993**, *99*, 8840.

(53) Stanton, J. F.; Gauss, J. *Theor. Chim. Acta* **1995**, *91*, 267.

(54) Olsen, J.; Jørgensen, P. In *Modern Electronic Structure Theory II*; Yarkoni, D., Ed.; VCH: New York, 1995.

(55) Christiansen, O.; Jørgensen, P.; Hättig, C. *Int. J. Quantum Chem.* **1998**, *68*, 1.

(56) Christiansen, O.; Koch, H.; Jørgensen, P.; Olsen, J. *Chem. Phys. Lett.* **1996**, *256*, 185.

(57) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.5; Gaussian, Inc.: Pittsburgh, PA, 1998.

(58) Dunning, T. H., Jr. *J. Chem. Phys.* **1988**, *90*, 1007.

(59) Helgaker, T.; Jensen, H. J. Aa.; Jørgensen, P.; Olsen, J.; Ruud, K.; Ågren, H.; Auer, A. A.; Bak, K. L.; Bakken, V.; Christiansen, O.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Enevoldsen, T.; Fernandez, B.; Hättig, C.; Hald, K.; Halkier, A.; Heiberg, H.; Hettema, H.; Jonsson, D.; Kirpekar, S.; Kobayashi, R.; Koch, H.; Mikkelsen, K. V.; Norman, P.; Packer, M. J.; Pedersen, T. B.; Ruden, T. A.; Sanchez, A.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sylvester-Hvid, K. O.; Taylor, P. R.; Vahtras, O. *Dalton, an ab initio electronic structure program*, release 1.2; 2001.

(60) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(61) Stanton, J. F.; Gauss, J.; Watts, J. D.; Lauderdale, W. J.; Bartlett, R. J. *Int. J. Quantum Chem., Symp.* **1992**, *26*, 879.

(62) Herzberg, G. *Molecular spectra and molecular structure, vol. III*: *Electronic spectra and electronic structure of polyatomic molecules*, 1st ed.; D. Van Nostrand: New York, 1966.

(63) Robin, R. B. *Higher excited states of polyatomic molecules vol. III*, 3rd ed.; Academic Press: New York, 1985.