

# Inner-Shell Excitation Spectroscopy of the Peptide Bond: Comparison of the C 1s, N 1s, and O 1s Spectra of Glycine, Glycyl-Glycine, and Glycyl-Glycyl-Glycine

Michelle L. Gordon,<sup>†</sup> Glyn Cooper,<sup>†</sup> Cynthia Morin,<sup>†</sup> Tohru Araki,<sup>†</sup> Cássia C. Turci,<sup>‡</sup> Konstantin Kaznatcheev,<sup>§</sup> and Adam P. Hitchcock<sup>\*,†</sup>

Department of Chemistry, McMaster University, Hamilton, ON, L8S 4M1, Canada, Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21910, Brazil, and Canadian Light Source, University of Saskatchewan, Saskatoon, SK S7N 5C6 Canada

Received: February 20, 2003; In Final Form: June 5, 2003

Oscillator strengths for C 1s, N 1s, and O 1s excitation spectra of gaseous glycine and the dipeptide, glycyl-glycine, have been derived from inner-shell electron energy-loss spectroscopy recorded under scattering conditions where electric dipole transitions dominate (2.5 keV residual energy,  $\theta \approx 2^\circ$ ). X-ray absorption spectra of solid glycine, glycyl-glycine, glycyl-glycyl-glycine, and a large protein, fibrinogen, were recorded in a scanning transmission X-ray microscope. The experimental spectra are assigned through interspecies comparisons and by comparison to ab initio computed spectra of various conformations of glycine and glycyl-glycine. Inner-shell excitation spectral features characteristic of the peptide bond are readily identified by comparison of the spectra of gas-phase glycine and glycyl-glycine. They include a clear broadening and a  $\sim 0.3$  eV shift of the C 1s  $\rightarrow \pi^*_{\text{C=O}}$  peak and introduction of a new pre-edge feature in the N 1s spectrum. These effects are due to 1s  $\rightarrow \pi^*_{\text{amide}}$  transitions introduced with formation of the peptide bond. Similar changes occur in the spectra of the solids. The computational results support the interpretation of the experimental inner-shell spectra and provide insight into electron density distributions in the core excited states. Possible conformational dependence of the inner-shell excitation spectra was explored by computing the spectra of neutral glycine in its four most common conformations, and of glycyl-glycine in planar and two twisted conformations. A strong dependence of the computed C 1s, N 1s, and O 1s spectra of glycyl-glycine on the conformation about the amide linkage was confirmed by additional ab initio calculations of the conformational dependence of the spectra of formamide.

## 1. Introduction

Proteins are chains of amino acids linked together by peptide bonds. Each protein is characterized by its primary structure (the amino acid sequence), its secondary structure (chain folding), and often a tertiary structure (subunit assembly). Although the biological activity of a protein is a function of all structural aspects, as well as associated entities such as metal ions and cofactors, at the root of modern biochemistry is a belief that it is the particular amino acids present and the way in which they are linked together in sequence that gives each protein its unique function and character. Thus, understanding the bonding between specific pairs of amino acids is one component of understanding the sequence-structure relationships of proteins.

Inner-shell spectroscopy, using either X-ray absorption (NEXAFS) or electron impact (ISEELS), is a useful technique for understanding bonding in complex molecules.<sup>1–3</sup> It is a site-specific probe with well-understood selection rules. Inner-shell spectra can be calculated with reasonable accuracy.<sup>3</sup> The goal of this paper is to use comparisons of the inner-shell excitation spectra of glycine and glycyl-glycine in the gaseous and solid states, along with conformation-specific ab initio calculations to probe in detail the links between inner-shell spectral

transitions and molecular structure, specifically those changes that relate to formation of the peptide bond.

Glycine, also denoted  $\alpha$ -aminoacetic acid, is the simplest amino acid. In the solid phase, glycine exists in a “zwitterionic” form, where the acidic hydrogen is transferred to the basic amino group. Upon heating, glycine vaporizes and converts to its nonionic form. In aqueous solution at low pH, glycine is found in its cationic form, at neutral pH the zwitterionic form is dominant, and at high pH the anion is the dominant species. The simplest species which incorporates a peptide bond is glycyl-glycine, in which two glycine molecules are combined in a head-to-tail arrangement with elimination of a water molecule. In this work we have measured the inner-shell (C 1s, N 1s, O 1s) spectra of two gaseous species, glycine (Gly) and glycyl-glycine (Gly-Gly), and the inner-shell (C 1s, N 1s, O 1s) spectra of four solid species, glycine, glycyl-glycine, the tripeptide glycyl-glycyl-glycine (Gly-Gly-Gly), and a protein, fibrinogen. Comparisons of the spectra of the gaseous and solid species, and experimental results with ab initio calculations, have provided insight into the electronic structure changes that accompany peptide bond formation.

Glycine and glycyl-glycine have several functional groups and low symmetries and could thus exhibit complex spectroscopy. They are nitrogen-containing organic compounds which are of interest owing to potential technical uses and their biological importance. In addition to our interest in exploring fundamental aspects of inner-shell excitation spectroscopy, a

\* To whom correspondence should be addressed. E-mail: aph@mcmaster.ca.

<sup>†</sup> McMaster University.

<sup>‡</sup> Universidade Federal do Rio de Janeiro.

<sup>§</sup> University of Saskatchewan.

motivation for the present study is to support emerging efforts to differentiate and thus possibly map the locations of peptides and possibly small proteins using NEXAFS microscopy.<sup>4–7</sup> The C 1s NEXAFS spectra of the 20 common amino acids (as solids deposited from acidic solution) were reported recently.<sup>8</sup> The next step after characterizing the individual amino acids is to examine the effect of peptide bond formation on the inner-shell spectrum. Further work on characterizing inner-shell spectra of larger peptides in the gas and condensed phases is planned.

In the present work, electron energy-loss spectroscopy (EELS) has been used to measure the C 1s, N 1s, and O 1s inner-shell excitation spectra of gaseous glycine and glycyl–glycine. All the spectra were acquired under small momentum-transfer conditions where electron scattering is dominated by electric dipole transitions and therefore simulates photoabsorption. The photoelectron spectra of gas-phase amino acids and small peptides have been reported.<sup>9–11</sup> Also previously reported are the C 1s, N 1s, and O 1s X-ray absorption spectra of surface-adsorbed glycine;<sup>12,13</sup> all 20 of the common amino acids as thin films;<sup>8</sup> the O 1s spectra of thin films of selected amino acids;<sup>14</sup> and the C 1s spectra of selected amino acids and several small peptides (Gly–Tyr, His–Phe, Gly–Trp, Arg–Gly, and Gly–Gly).<sup>15</sup> Quantum chemical calculations have been reported for surface-adsorbed glycine<sup>12,13</sup> and all of the 20 common amino acids.<sup>8,16,17</sup> To our knowledge, this is the first report of the inner-shell spectra of a *gaseous* amino acid or peptide. Gas-phase spectroscopy of these species is difficult due to their tendency to decompose on heating. Several amino acids have been found not to decompose upon heating,<sup>18–20</sup> which makes them amenable to study in the gas phase. Comparisons of the inner-shell excitation spectra of the gas phase and solid phase of the same species are of interest in order to understand the sensitivity of different types of inner-shell excitation features to the local environment.<sup>21,22</sup>

## 2. Experimental Section

**2.1. Samples.** Glycine ( $C_2H_5NO_2$ ,  $NH_2CH_2CO_2H$ ), glycyl–glycine ( $C_4H_8N_2O_3$ ,  $NH_2CH_2CONHCH_2CO_2H$ ), and glycyl–glycyl–glycine ( $C_6H_{11}N_3O_4$ ,  $NH_2CH_2CONHCH_2CONHCH_2CO_2H$ ) were obtained commercially from Sigma Aldrich in the form of crystalline powders, each with a stated purity better than 99%. They were used without further purification. Plasminogen-depleted human plasma fibrinogen was obtained from Calbiochem. It is reported to be >95% pure, clottable by thrombin, and homogeneous as judged by SDS–PAGE. It was used without further purification.

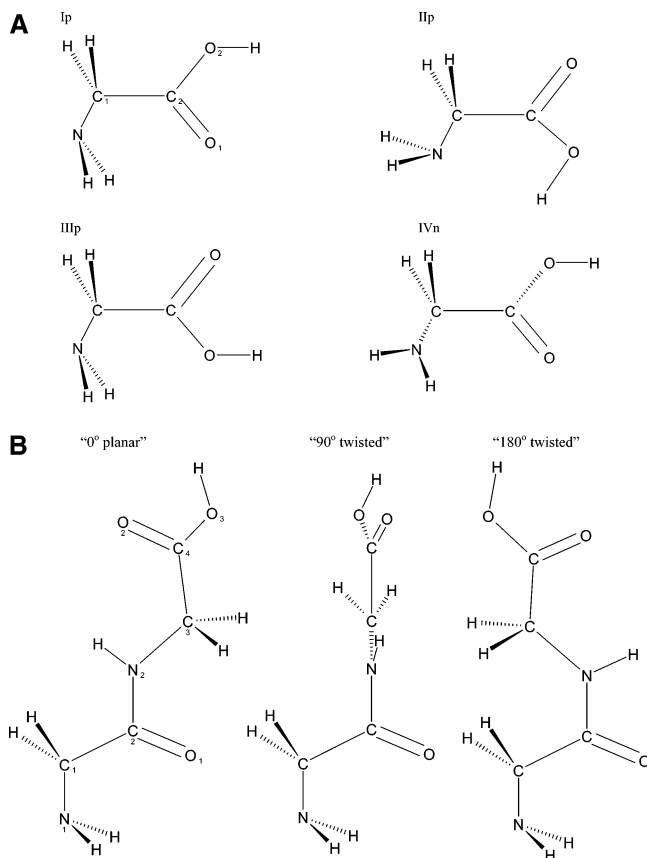
**2.2. Electron Energy-Loss Spectroscopy.** Electron energy-loss spectra of gaseous glycine and glycyl–glycine in the regions of C 1s, N 1s, and O 1s excitation were acquired with a gas-phase ISEELS spectrometer that has been described in detail previously.<sup>1,23</sup> The glycine and glycyl–glycine samples were placed inside a small aluminum tube attached directly to the collision cell of the spectrometer. The cell was then heated with an internally mounted quartz–halogen bulb, while the end plates of the gas cell were water-cooled to trap the vaporized sample and avoid insulating deposits on sensitive parts of the spectrometer. Multiple spectra were recorded to confirm reproducibility and to check the heat sensitivity of the samples. The spectrometer was operated under conditions of small momentum transfer—small scattering angle ( $\sim 2^\circ$ ) and high electron impact energy (2.5 keV + energy loss)—where electric dipole transitions dominate. The spectrometer resolution depends on the electron beam current and analyzer pass energy but is typically 0.75 eV full width at half-maximum (fwhm) at a beam current of 20  $\mu A$  and 0.5 eV fwhm at a beam current of  $\sim 1 \mu A$ .

The energy scales of all spectra were calibrated by acquiring the spectra of a stable mixture of the analyte molecule and a reference compound. The C 1s and O 1s spectra of both molecules were calibrated using the C 1s  $\rightarrow \pi^*$  transition of  $CO_2$  (290.74(4) eV)<sup>24,25</sup> and the O 1s  $\rightarrow \pi^*$  transition of  $CO_2$  (535.4(2) eV).<sup>26</sup> The N 1s spectra were calibrated using the N 1s  $\rightarrow \pi^*$  (401.10(2) eV) transition of  $N_2$ .<sup>24,25</sup>

The excitation spectra associated with a particular inner-shell edge are isolated from the underlying valence-shell and core ionization continua by subtracting a smooth curve determined from a curve fit of the function  $a(E - b)^c$  to the pre-edge experimental signal. The background subtracted spectra are converted to absolute oscillator strength scales using previously described methods.<sup>27</sup>

**2.3. NEXAFS Spectroscopy of Solids.** The NEXAFS spectra were recorded using the bending magnet scanning transmission X-ray microscope at the Advanced Light Source.<sup>28,29</sup> Milligram amounts of the pure materials were dissolved to form dilute aqueous solutions. Although the pH was not controlled with buffers to avoid any background, the pH of these dilute solutions was such that the species are expected to be predominantly in zwitterion form. A small drop ( $\sim 0.5 \mu L$ ) of this solution was placed on a thin (100 nm)  $Si_3N_4$  window and allowed to dehydrate to form a film sufficiently thin ( $\sim 100$  nm) to allow transmission measurements with an optical density of about 1 at the strong C 1s peak. The glycyl–glycine and triglycine samples did not exhibit any polarization effects, and the spectra of all regions of the sample were the same. However, the dried glycine sample was highly crystalline. Regions which appeared uncrystallized were chosen for measurements, but the spectrum was found to vary from location to location on account of different crystallite orientations which produced different spectra. A second sample was prepared as a saturated solution in a wet cell with X-ray-transparent silicon nitride windows. Although the sample actually partly solidified due to incomplete sealing and partial drying, the solid material was a randomly oriented gel, showing identical spectra in all areas. The C 1s and N 1s spectra of glycine presented here are those from that sample. They agree with spectra derived from the crystalline sample as a weighted average of regions which had extremely  $\sigma$ -dominated and extremely  $\pi$ -dominated spectra. Since the randomly oriented solid deposited in the wet cell was still very hydrated, it was not possible to obtain a reliable O 1s spectrum. Therefore, the O 1s spectrum of glycine was taken from the symmetry-weighted average of the dried crystalline sample. That spectrum is in reasonable agreement with an estimate of the glycine contribution to the mixed glycine water spectrum recorded on the wet cell sample. At high radiation doses, all three of the fully dried samples exhibited additional spectral features associated with radiation damage. The spectra presented here were recorded using acquisition methods (rapid scanning of lines or areas using a defocused beam) which kept the dose at a level where no significant radiation damage occurred. Interestingly, the damage rate for the wet glycine sample was significantly lower than that for the dry glycine samples. The NEXAFS spectra of glycine were calibrated by measuring its spectrum simultaneously with the spectra of gaseous  $CO_2$ ,  $N_2$ , and  $O_2$ , introduced into the STXM chamber. The energy scales for the spectra of the other solid samples were then calibrated relative to glycine spectra recorded in a short time interval, over which the energy scale of the instrument is known to be stable.

**2.4. Computational Methods.** To gain insight into how the electronic structure changes that occur with peptide bond formation are reflected in inner-shell excitation spectra, calcula-

**SCHEME 1: Conformers of (A) Glycine and (B) Gly–Gly**

tions were performed using GSCF3 (Gaussian self-consistent field version 3),<sup>30,31</sup> which is an ab initio code designed specifically for inner-shell excitation and ionization calculations. The program uses the Hartree–Fock SCF approach to solve for the energies and molecular orbitals of the system under investigation. The basis sets used are those of Huzinaga et al.<sup>32</sup> The improved virtual orbital (IVO) method,<sup>31,33</sup> which explicitly takes into account the core hole in the Hartree–Fock approximation, is used to perform quantum calculations on core-excited molecules. In this approach, the core electron is removed directly from an inner-shell orbital specified by the user. The virtual orbitals of this system provide a good approximation to the term values of the core excitation features at that site. A separate calculation is performed for each distinct chemical site in each molecule (see Scheme 1).

There are several stable conformers of glycine in the gas phase. Theoretical studies have found as many as eight conformers,<sup>34–39</sup> which are labeled Ip, Iip, IIIp, IVn, Vn, VIp, VIIp, and VIIIp (p denotes a  $C_s$ -symmetry conformer, while n denotes a  $C_1$ -symmetry conformer). However, only a few of these have been observed experimentally. Ip and Iip have been identified by microwave spectroscopy,<sup>40–44</sup> while the presence of at least one other conformer has been inferred from infrared spectroscopy of glycine trapped in inert gas matrices.<sup>44</sup> To investigate the sensitivity of inner-shell excitation spectra to the different glycine conformers, and also, if necessary, to take into account their respective contributions to the experimental spectra reported in the present work, we have performed separate GSCF3 calculations on the four conformers of glycine predicted by the most elaborate levels of theory<sup>37–39</sup> to have the highest relative abundances at the estimated temperature of the gas in the present electron energy-loss experiments ( $\sim 400$  K): Ip,

53%; Iip, 9%; IIIp, 30%; and IVn, 7%. According to the calculated total energies, the other four conformers should comprise only  $\sim 2\%$  of the molecular population.

For glycyl–glycine, many more molecular conformations are possible. For this computational study, we have restricted the part of the molecule containing the amine and amide groupings ( $H_2N-CH_2-CONH$ ) to have a conformation similar to that of the Ip conformer of glycine, since this is the lowest energy configuration. This allows some degree of intramolecular hydrogen bonding between the oxygen of the amide group and the H's of the  $NH_2$  group. We have then performed calculations on structures with three different orientations of the carboxyl part of the molecule ( $CH_2COOH$ ) relative to the amide and amine parts. X-ray diffraction, in conjunction with theoretical Hartree–Fock calculations,<sup>45</sup> indicates that the preferred orientation is one where the molecular "backbone" (NCCNCCOH) is planar and the  $CH_2COOH$  group is positioned on the side of the molecule opposite to the  $CH_2NH_2$  group. Intramolecular hydrogen bonding can then take place between the C=O oxygen of the carboxyl group and the H of the NH amide group. We have called this orientation the "0° planar" conformation. We have then constructed two other possible conformations, one where the  $CH_2COOH$  group is rotated about the  $NH-CH_2$  bond by 90° ("90° twisted"), and the other where it is twisted by 180° ("180° twisted"). We have performed GSCF3 calculations on each of these conformers in order to investigate possible conformation-dependent aspects of the inner-shell excitation spectra of molecules containing peptide bonds. This aspect of our work provides a qualitative indication of the sensitivity of ISEELS and NEXAFS spectra to conformations about the amide linkages in peptides. Since the secondary structure and thus details of protein folding depend on the conformation at successive amide linkages,<sup>46</sup> there would be potential applications to structural investigations of peptides and possibly proteins, if there is a conformational sensitivity of the inner-shell excitation spectra.

Table 1 lists the geometries and the basis sets used in the GSCF3 theoretical calculations. The calculations are performed in three steps. In step one, the eigenvectors (MOs) and eigenvalues of the ground state are calculated, and the core MO that will lose the electron is determined. In the second step, the core ion state is computed by removing the user-specified core electron and allowing the system to relax and reorganize in the presence of the core hole. The difference in the total energy of the core-ionized and ground states gives the core level ionization potential (IP), with a typical accuracy of 1 eV. The third step determines the core excitation energies and transition probabilities in terms of the IVO approximation.<sup>31,33</sup> The absolute accuracy of the computed core excitation energies depends on the size of the basis set used. However, the core state term values ( $TV = IP - E$ ) are more accurate and relatively independent of basis set choices. The core excitation term values and optical oscillator strengths are generated by the third step of the GSCF3 calculation. These are used to generate simulated core excitation spectra by summing Gaussian lines at an energy given by the term value, an area given by the oscillator strength for excitation to each improved virtual orbital, and a width chosen as a function of the term value. In the discrete region, the chosen line width is that of the instrument, while larger line widths are used for the lifetime-broadened continuum resonances. The widths used are summarized in footnotes to the tables which summarize the computational results. In comparing the computed and experimental spectra, the calculated transition energies

TABLE 1: Geometries and Basis Sets<sup>a</sup> for GSCF3 Computations

A. Geometry of Conformers of Glycine								
atom	position (Å)			atom	position (Å)			
	x	y	z		x	y	z	
Ip Conformer								
C <sub>1</sub>	0.56356	-0.85777	0.00000	H	0.56620	2.3388	0.00000	
N	-0.41686	-1.9219	0.00000	H	1.2144	-0.95036	0.87443	
C <sub>2</sub>	0.00000	0.55239	0.00000	H	1.2144	-0.95036	-0.87443	
O <sub>1</sub>	0.99612	1.4717	0.00000	H	-1.0198	-1.8383	-0.81094	
O <sub>2</sub>	-1.1735	0.84380	0.00000	H	-1.0198	-1.8383	0.81094	
IIp Conformer								
C <sub>1</sub>	-0.82840	-0.59630	0.00000	H	1.41116	-0.48834	0.00000	
N	0.03356	-1.7821	0.00000	H	-1.4767	-0.55570	0.87852	
C <sub>2</sub>	0.00000	0.69372	0.00000	H	-1.4767	-0.55570	0.87852	
O <sub>1</sub>	1.32388	0.48908	0.00000	H	-0.11649	-2.3606	0.81697	
O <sub>2</sub>	-0.51003	1.78733	0.00000	H	-0.11649	-2.3606	-0.81697	
IIIp Conformer								
C <sub>1</sub>	-0.85697	-0.61761	0.00000	H	1.76895	1.2365	0.00000	
N	-0.20502	-1.9118	0.00000	H	-1.5113	-0.53499	0.87259	
C <sub>2</sub>	0.00000	0.63996	0.00000	H	-1.5113	-0.53499	-0.87259	
O <sub>1</sub>	1.32951	0.37375	0.00000	H	0.39374	-2.0061	0.81226	
O <sub>2</sub>	-0.44912	1.7630	0.00000	H	0.39374	-2.0061	-0.81226	
IVn Conformer								
C <sub>1</sub>	0.72591	-0.70797	0.15628	H	-2.3908	-0.09728	-0.11449	
N	1.8704	0.099651	-0.23160	H	0.64571	-1.5901	-0.48259	
C <sub>2</sub>	-0.53718	0.10503	0.01983	H	0.73561	-1.0613	1.19996	
O <sub>1</sub>	-1.6258	-0.68999	-0.09721	H	1.7979	1.01960	0.19012	
O <sub>2</sub>	-0.59237	1.3122	0.05774	H	2.7315	-0.32842	0.08722	
B. Geometry of Conformers of Glycine–Glycine								
atom	position (Å)			atom	position (Å)			
	x	y	z		x	y	z	
“0° Planar” Conformer								
C <sub>1</sub>	0.000	0.000	0.000	H	1.523	2.023	0.000	
C <sub>2</sub>	1.532	0.000	0.000	H	5.577	3.895	0.000	
O <sub>1</sub>	2.237	-1.000	0.000	H	-0.573	-1.990	0.778	
N <sub>1</sub>	-0.552	-1.361	0.000	H	-0.573	1.990	-0.778	
N <sub>2</sub>	2.093	1.203	0.000	H	-0.189	0.598	0.880	
C <sub>3</sub>	3.532	1.380	0.000	H	-0.189	0.598	-0.880	
C <sub>4</sub>	3.928	2.858	0.000	H	3.700	0.760	0.880	
O <sub>2</sub>	3.153	3.781	0.000	H	3.700	0.760	-0.880	
O <sub>3</sub>	5.278	2.976	0.000					
90° Twisted Conformer								
C <sub>1</sub>	-1.389	-0.647	0.000	H	0.000	1.830	-0.863	
C <sub>2</sub>	0.000	0.000	0.000	H	0.000	5.239	2.020	
O <sub>1</sub>	1.062	-0.608	0.000	H	-1.067	-2.693	0.778	
N <sub>1</sub>	-1.314	-2.114	0.000	H	-1.067	-2.693	-0.778	
N <sub>2</sub>	0.000	1.327	0.000	H	-1.812	-0.185	0.880	
C <sub>3</sub>	0.000	2.096	1.230	H	-1.812	-0.185	-0.880	
C <sub>4</sub>	0.000	3.603	0.964	H	0.880	1.605	1.644	
O <sub>2</sub>	0.000	4.112	-0.128	H	0.880	1.605	1.644	
O <sub>3</sub>	0.000	4.280	2.138					
180° Twisted Conformer								
C <sub>1</sub>	-1.389	-0.647	0.000	H	0.863	1.830	0.000	
C <sub>2</sub>	0.000	0.000	0.000	H	-2.020	5.239	0.000	
O <sub>1</sub>	1.062	-0.608	0.000	H	-1.067	-2.693	0.778	
N <sub>1</sub>	-1.314	-2.114	0.000	H	-1.067	-2.693	-0.778	
N <sub>2</sub>	0.000	1.327	0.000	H	-1.812	-0.185	0.880	
C <sub>3</sub>	-1.230	2.096	0.000	H	-1.812	-0.185	-0.880	
C <sub>4</sub>	-0.964	3.603	0.000	H	-1.644	1.605	0.880	
O <sub>1</sub>	0.128	4.112	0.000	H	-1.644	1.605	-0.880	
O <sub>2</sub>	-2.138	4.280	0.000					

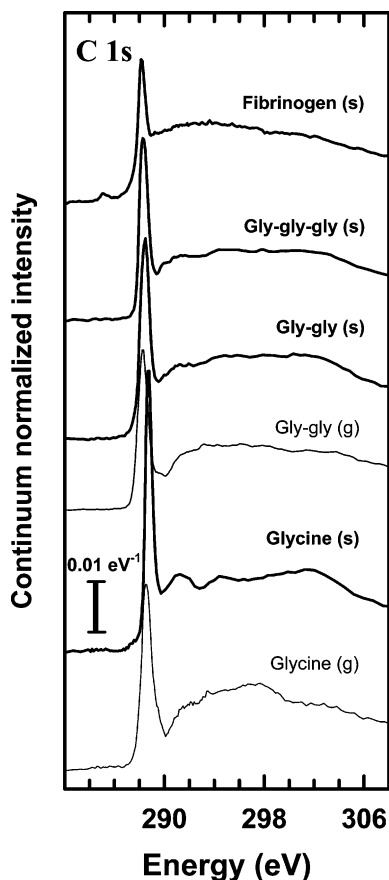
<sup>a</sup> Basis sets (Huzinaga-contracted Gaussian-type orbitals): HTS8X, 411121 21111 for core hole; PGM6G, xxx for 2\* d-polarization (xxx = element-specific) on core hole; HTS6X 6, 63 5 for non-core hole C, N, O; HTS3X 1, 6 for H. <sup>b</sup> Bond lengths of Gly–Gly were taken from average of peptide links<sup>46</sup> plus the glycine geometry for the non-peptide part of the molecule. The conformation of the lowest energy planar conformer was determined by reference to the geometry shown in the X-ray diffraction/theoretical study by Abramov et al.<sup>45</sup>

are retained, but a rigid shift of the computed energy scale relative to the experimental energy scale is built into the plot presentation. These shifts, which range from 1.2 to 3.3 eV, are indicated in the figure captions. They are typical of the inaccuracies found in other applications of GSCF3 to core excitation.

### 3. Results and Discussion

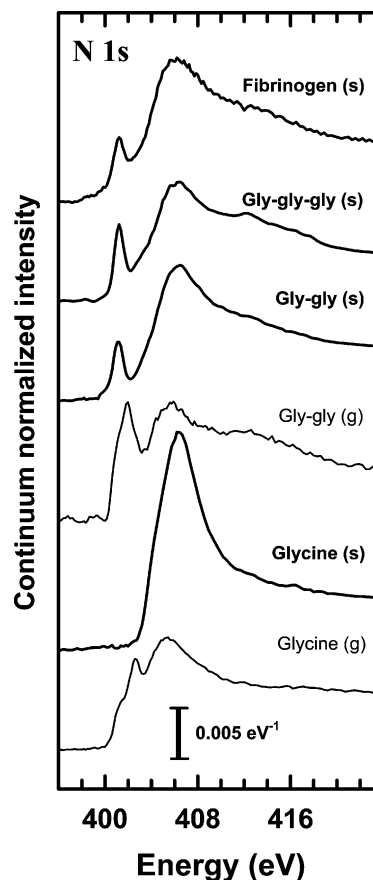
#### 3.1. Comparison of Gas- and Condensed-Phase Spectra.

Figure 1 compares the C 1s spectra of solid glycine, glycyglycine, glycyglycylglycine, and fibrinogen with the C 1s ISEELS spectra of gaseous glycine and glycyglycine. The



**Figure 1.** Comparison of the C 1s NEXAFS spectra of solid glycine, solid Gly–Gly dimer, solid Gly–Gly–Gly trimer, and fibrinogen, recorded in a scanning transmission X-ray microscope (STXM), with the C 1s spectra of gas-phase glycine and glycy–glycine, recorded by inner-shell electron energy-loss spectroscopy (ISEELS) in the dipole regime. The intensity scale of the background-subtracted spectra is in terms of oscillator strength per carbon atom, and the data are plotted with offsets for clarity.

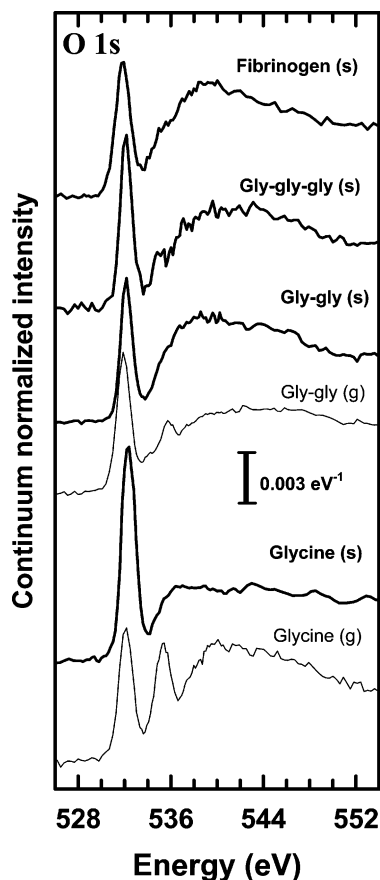
gas-phase spectra (all edges) can be obtained in electronic form from <http://unicorn.mcmaster.ca/corex/cedb-title.html> for those readers who wish to see the finer details of these spectra. The C 1s spectra of glycine,<sup>8,12,13,15</sup> glycy–glycine,<sup>15</sup> and fibrinogen<sup>47</sup> have been reported previously, the latter in comparison to the C 1s spectrum of albumin.<sup>47,48</sup> All spectra have been background subtracted and normalized to the continuum oscillator strength for a single C atom. Offsets are used for clarity. All six spectra are dominated by the C 1s(C=O)  $\rightarrow$   $\pi^*_{C=O}$  transition in the range of 288.2–288.6 eV. There is a  $\sim$ 0.3 eV shift to lower energy between the spectra of solid glycine and solid Gly–Gly or Gly–Gly–Gly, and a further shift to lower energy of  $\sim$ 0.15 eV between Gly–Gly–Gly and fibrinogen. These shifts reflect the changes from only a carboxyl group in glycine to a mixed (carboxyl, amide) situation in Gly–Gly or Gly–Gly–Gly, to one with a large number of amide groups and only one carboxyl group in a protein. A similar shift can be seen in a previous comparison of the C 1s spectra of glycine and Gly–Gly–Gly,<sup>15</sup> and the origin of that shift in terms of change from a carboxyl to an amide group was noted in that work. Several other features differ between the C 1s spectra of a large protein and the simplest dipeptide, Gly–Gly. The small peak at 285.1 eV in fibrinogen arises from amino acid residues with phenyl groups. There is a weak shoulder below the  $\pi^*_{C=O}$  resonance and a filling-in of the 289–291 eV region in the spectrum of fibrinogen relative to the small peptides, again



**Figure 2.** Comparison of the N 1s NEXAFS spectra of solid glycine, solid Gly–Gly dimer, solid Gly–Gly–Gly trimer, and fibrinogen, recorded by STXM, and the gas-phase spectra of Gly and Gly–Gly, recorded by ISEELS. The intensity scale of the background-subtracted spectra is in terms of oscillator strength per nitrogen atom, and the data are plotted with offsets for clarity.

features attributable to the R groups.<sup>8</sup> The intensity of the amide  $\pi^*$  is much weaker relative to the continuum in fibrinogen on account of dilution of the peptide bond signal with the contributions of the R groups.

Figure 2 compares the N 1s spectra of solid glycine, glycy–glycine, Gly–Gly–Gly, and fibrinogen with the N 1s ISEELS spectra of gaseous glycine and glycy–glycine. To our knowledge, only the N 1s spectrum of albumin<sup>47–49</sup> has been reported previously. The spectra have been background subtracted and normalized to the continuum oscillator strength for a single N atom. Offsets are used for clarity. In contrast to the C 1s spectra, there are rather dramatic differences between the gas-phase and solid-state N 1s spectra. The N 1s spectra of both gaseous species have relatively strong contributions from excitations to virtual levels of Rydberg character. Although occasionally disputed,<sup>22</sup> Rydberg levels of a gas are believed to be broadened into excitonic states<sup>21</sup> or eliminated in the solid state. The latter is clearly the case in glycine(s). The peak at 401.5 eV in the spectra of solid Gly–Gly, Gly–Gly–Gly, and fibrinogen is due to the N 1s  $\rightarrow$   $\pi^*_{C=O}(C=ONH)$  transition, i.e., a N 1s  $\rightarrow$   $\pi^*_{amide}$  transition, indicating a partial  $\pi^*$  character at the amide nitrogen. This feature is absent in the N 1s spectrum of solid glycine since there is no amide bond. Since most nitrogen atoms in a protein are amide nitrogen sites, the N 1s spectrum of fibrinogen should be similar to the N 1s spectra of solid Gly–Gly and Gly–Gly–Gly, as observed. We note that, in the N 1s spectrum of albumin,<sup>48,49</sup> the  $\pi^*_{amide}$  signal is actually much stronger than that in fibrinogen. This is one of a number of observations of



**Figure 3.** Comparison of the O 1s spectra of solid glycine, solid Gly–Gly dimer, solid Gly–Gly–Gly trimer, and fibrinogen, recorded by STXM, and the gas-phase spectra of Gly and Gly–Gly, recorded by ISEELS. The intensity scale of the background-subtracted spectra is in terms of oscillator strength per oxygen atom, and the data are plotted with offsets for clarity.

rather large variations in the N 1s spectroscopy of proteins and related species, which may be associated with a strong sensitivity of the N 1s spectrum to the environment of the amide bonds in the sample (see also section 3.3 and the note at end of the paper).

Figure 3 compares the O 1s spectra of solid Gly, Gly–Gly, Gly–Gly–Gly, and fibrinogen with the O 1s ISEELS spectra of gaseous glycine and glycylic acid. To our knowledge, O 1s spectra have not been presented for any of these materials. The spectra have been background subtracted and normalized to the continuum oscillator strength for a single O atom. Offsets are used for clarity. There are rather dramatic differences between the gas-phase and solid-state spectra, particularly between the O 1s spectrum of gaseous neutral glycine, which shows two prominent discrete resonances, and the other four species, which are dominated by, or exclusively exhibit, a single low-lying  $\pi^*_{C=O}$  resonance. The two peaks of comparable intensity in Gly(g) are characteristic of gaseous carboxylic acids<sup>50</sup> and are associated with the existence of two oxygen atoms in very different environments (C=O vs C–OH). These features coalesce into a single peak, with a transition energy close to the lower energy peak, in the carboxylate and dimer forms of the molecules.<sup>51</sup> Thus, the absence of a double-peaked structure in the O 1s spectrum of Gly(s) indicates that the oxygen atoms are either in a dimer structure (if neutral) or, more likely, in a carboxylate environment in solid glycine, consistent with the expected zwitterionic form. The much higher intensity relative to the continuum in Gly(s) relative to Gly(g) is a consequence of both oxygen atoms in Gly(s) being in the same

**TABLE 2: Energies, Term Values, and Proposed Assignments for C 1s Spectral Features of Glycine (Gas, Solid)**

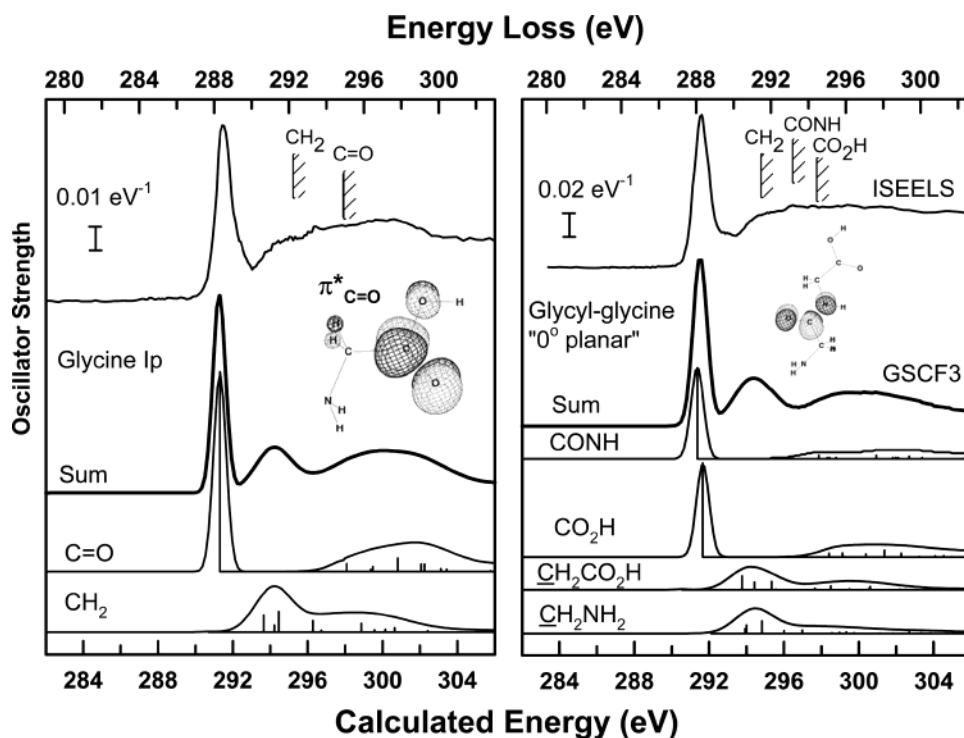
$E$ (eV), $\pm 0.08$	gas <sup>a</sup>		$E$ (eV), $\pm 0.08$	solid	
	term value <sup>b</sup> (eV) $T_{CH_2}$	$T_{COOH}$		assignment (final orbital) CH <sub>2</sub>	COOH
288.56 <sup>c</sup>		6.5	288.73 <sup>c</sup>		$\pi^*_{C=OOH}$
289.3 (sh)	3.0		289.3 (sh)	3p	
290.8 (sh)	1.5		291.3	$\sigma^*_{C-H}$ , $\sigma^*_{C-N-H}$	$\sigma^*_{N-H}$
292.3 <sup>d</sup>				IP	
294 (br)	–1.7		294.3		$\sigma^*_{O-H}$
295.1 <sup>d</sup>					IP
297		–2		$\sigma^*_{C-C}$ , $\sigma^*_{C-N}$	$\sigma^*_{C-OH}$
303 (br)		–8	301	$\sigma^*_{C-C}$	$\sigma^*_{C-O}$

<sup>a</sup> The gas-phase spectra can be obtained in electronic form from <http://unicorn.mcmaster.ca/corex/cedb-title.html>. <sup>b</sup> Term value = ionization potential (IP) – excitation energy. <sup>c</sup> Calibration: gas, –2.18(4) eV relative to CO<sub>2</sub>  $\pi^*$ ; solid, –4.01(6) relative to CO<sub>2</sub>, 3s ( $\nu = 0$ ). <sup>d</sup> From XPS.<sup>52</sup>

environment. An O 1s(C–OH)  $\rightarrow \pi^*_{COOH}$  peak is expected in Gly–Gly (g) and is clearly visible at 535.7 eV. Its intensity is much reduced since the hydroxyl oxygen is only one of the three oxygen atoms. This peak is also shifted 0.4 eV to higher energy than the corresponding feature in Gly(g) for reasons not presently understood. The O 1s spectra of solid Gly–Gly and Gly–Gly–Gly are similar, with each exhibiting a single peak at 532.2 eV, which is dominated by the O 1s  $\rightarrow \pi^*_{CONH}$  transition but has minor contributions from O 1s  $\rightarrow \pi^*_{COO^-}$  transitions. The O 1s  $\rightarrow \pi^*_{CONH}$  peak in fibrinogen is shifted further to lower energy (531.6 eV), and its intensity is significantly reduced relative to the continuum, probably due to contributions from side-chain oxygen atoms in saturated bonding environments.

**3.2. Detailed Analysis of Gas-Phase Spectra Aided by GSCF3 Calculations.** **3.2.1. C 1s Spectra.** Figure 4 plots the C 1s spectra computed by GSCF3 for the Ip conformer of glycine and the “0° planar” conformation of glycylic acid, which are expected to be the lowest energy conformers of each species. The site-specific components of the calculations are also shown. The experimental C 1s spectra of glycine and glycylic acid are also included for comparison. The other glycine conformers studied theoretically, IIp, IIIp, and IVn, show only small differences in peak positions and intensities (these are discussed in section 3.3). Note that the energies and intensities are on absolute scales for both the computed and experimental results (offsets are used for clarity). In contrast to Figures 1–3, the intensity scale in Figures 4–6 is oscillator strength for the full molecule, not per atom. The energies, term values, and proposed assignments for the features in the C 1s spectrum are presented in Table 2 (for glycine) and in Table 3 (for glycylic acid). The experimental term values ( $TV = IP - E$ ) were determined from published ionization potentials (IPs)<sup>52</sup>—292.3 eV for the methyl and 295.1 eV for the carboxylic carbon of glycine—or estimated from the IPs of carbon atoms in similar molecular structures. Table 8 summarizes the calculated transition energies, oscillator strengths, and orbital characters of selected low-energy states of glycine in the Ip conformation. Table 9 summarizes the calculated transition energies, oscillator strengths, and orbital characters of selected low-energy states for the “0° planar” conformation of glycylic acid.

The C 1s spectrum of glycine shows a strong peak at 288.6 eV, corresponding to the C 1s(COOH)  $\rightarrow \pi^*_{COOH}$  transition associated with the carboxyl group. The  $\pi^*$  resonance in acetic



**Figure 4.** C 1s oscillator strength spectra of gaseous glycine and glycyl-glycine compared to computed spectra of the Ip conformer of glycine and the “0° planar” conformer of glycyl-glycine. The geometry of the “0° planar” conformer was estimated by analogy with that of the Ip conformer of glycine, along with information from the crystal structure and the theoretically optimized geometry.<sup>45</sup> The hatched line indicates the C 1s IP, as determined by X-ray photoelectron spectroscopy or estimated from values for similar structures.<sup>52</sup> The spectral contributions of the chemically different sites are plotted. The computed spectra are plotted on absolute oscillator strength scales (total for the full molecule) with offsets. The computed energy scales are plotted with a shift relative to experiment of 3.0 (Gly) and 3.3 eV (Gly-Gly) in order to align the first peak. The sketches are for the (C 1s(COOH)<sup>-1</sup>,  $\pi^*_{\text{COOH}}$ ) (Gly) and the (C 1s(CONH)<sup>-1</sup>,  $\pi^*_{\text{CONH}}$ ) (Gly-Gly) molecular orbitals.

**TABLE 3: Energies, Term Values, and Proposed Assignments for C 1s Spectral Features of Glycyl-Glycine (Gas, Solid)**

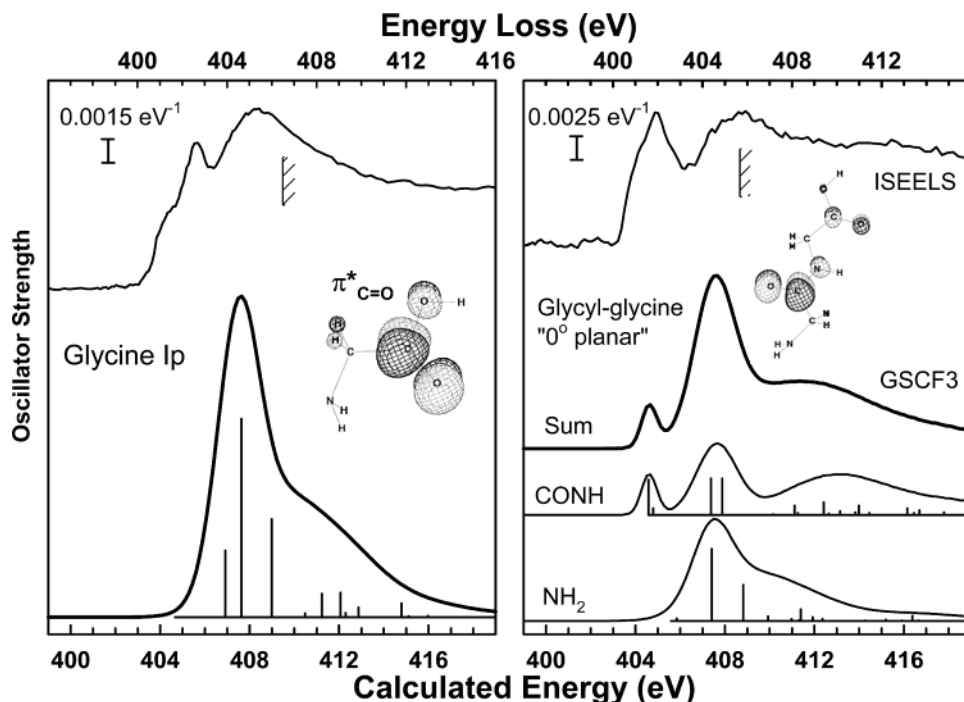
$E$ (eV), $\pm 0.08$	gas <sup>a</sup>				$E$ (eV), $\pm 0.08$	solid			
	term value <sup>b</sup> (eV)					assignment (final orbital)			
	$T_{\text{CH}_2\text{NH}_2}$	$T_{\text{CONH}}$	$T_{\text{CH}_2\text{COOH}}$	$T_{\text{COOH}}$		$\text{CH}_2\text{NH}_2$	CONH	$\text{CH}_2\text{COOH}$	COOH
288.26 <sup>c</sup>		5.7			288.48 <sup>c</sup>	$\pi^*_{\text{CONH}}$			
288.6 (fit) <sup>d</sup>				6.5					$\pi^*_{\text{COOH}}$
289.7	2.6		2.6			3p		3p	
291 (sh)	1.3		1.3		291	$\sigma^*_{\text{CH}}$ , $\sigma^*_{\text{NH}}$ , $\sigma^*_{\text{CC}}$		$\sigma^*_{\text{CH}}$ , $\sigma^*_{\text{NH}}$ , $\sigma^*_{\text{CC}}$	
292.3 <sup>e</sup>						IP		IP	
293.1				2.0					$\sigma^*_{\text{OH}}$
294.0 <sup>e</sup>							IP		
295.1 <sup>e</sup>									IP
296 (br)				-1	294.5	$\sigma^*_{\text{CC}}$	$\sigma^*_{\text{CC}}$	$\sigma^*_{\text{CC}}$	$\sigma^*_{\text{COH}}$
298		-4			298		$\sigma^*_{\text{C=O}}$		
303 (br)				-8	302				$\sigma^*_{\text{C-O}}$

<sup>a</sup> The gas-phase spectra can be obtained in electronic form from <http://unicorn.mcmaster.ca/corex/cedb-title.html>. <sup>b</sup> Term value = ionization potential (IP) - excitation energy. <sup>c</sup> Calibration: gas, -2.48(4) eV relative to CO<sub>2</sub>  $\pi^*$ ; solid, -4.26(6) relative to CO<sub>2</sub> 3s ( $\nu = 0$ ). <sup>d</sup> This value was taken from fitting the first peak to two components. <sup>e</sup> Estimated from XPS of analogous species.<sup>52</sup> Trigly(s) peaks observed at 288.33(8) (calibration, 4.11(6) relative to CO<sub>2</sub> 3s ( $\nu = 0$ )); 290.0, 291.2, 294.8, 301.

acid<sup>50</sup> is located at 288.7 eV, which is close to the  $\pi^*$  resonance position for glycine measured in the present work. In general, this position is consistent with trends in positions of C 1s  $\rightarrow \pi^*_{\text{C=O}}$  transitions as a function of electronegativity of the substituents, which have been characterized recently by solid-state and computational studies by Urquhart and Ade<sup>53</sup> and by gas-phase spectroscopy and computation by Lessard et al.<sup>54</sup>

Since Gly-Gly contains three carbon atoms which are in similar environments to counterparts in glycine and only one carbon atom which differs strongly, namely the carbon at the amide group, we should expect only small changes in the overall spectrum. Relative to the spectrum of glycine, the C 1s spectrum of glycyl-glycine exhibits a significantly broader  $\pi^*$  peak,

which is shifted about 0.3 eV toward lower energy. This shift is a manifestation of peptide bond formation in which a carboxyl group is replaced by an amide group. Since the carbon atom at the amide (peptide) bond is surrounded by an oxygen and one nitrogen atom, rather than the two oxygen atoms at the carboxylic acid carbon, the C 1s(CONHR) core level is less tightly bound, causing a shift of the corresponding  $\pi^*$  resonance toward lower energies. The 0.3 eV shift observed between the carboxyl  $\pi^*_{\text{C=O}}$  peak in Gly and the amide  $\pi^*_{\text{CONH}}$  peak in Gly-Gly is fully consistent with the recently reported correlations of the energy of C 1s(C=O)  $\rightarrow \pi^*_{\text{C=O}}$  transitions as a function of the electronegativity of the immediate environment.<sup>53,54</sup>



**Figure 5.** N 1s oscillator strength spectra of gaseous glycine and glycyl-glycine compared to computed spectra of the Ip conformer of glycine and the “0° planar” conformer of glycyl-glycine. See caption to Figure 4 for further details. The computed energy scales are plotted with a shift relative to experiment of 3.0 (Gly) and 3.0 eV (Gly-Gly) in order to align the  $\sigma^*_{\text{C-N}}$  resonance at 405 eV. The sketches are for the  $(\text{N } 1s)^{-1}, \pi^*_{\text{COOH}}$  (Gly) and the  $(\text{N } 1s)^{-1}, \pi^*_{\text{CONH}}$  (Gly-Gly) molecular orbitals.

**TABLE 4: Energies, Term Values, and Proposed Assignments for N 1s Spectral Features of Glycine (Gas, Solid)**

gas <sup>a</sup>		solid		assignment (final orbital)
$E$ (eV), $\pm 0.1$	term value <sup>b</sup> (eV)	$E$ (eV), $\pm 0.1$		
401.3	4.3			3s
402.5 <sup>c</sup>	3.1			3p
405.3	0.3	405.3 (sh)		$\sigma^*_{\text{C-N}}$ , $\sigma^*_{\text{N-H}}$
405.6 <sup>d</sup>				IP
		406.7 <sup>c</sup>		$\sigma^*_{\text{C-N}}$
417 (w)	-11			shake-up

<sup>a</sup> The gas-phase spectra can be obtained in electronic form from <http://unicorn.mcmaster.ca/corex/cedb-title.html>. <sup>b</sup> Term value = ionization potential (IP) - excitation energy. <sup>c</sup> Calibration: gas, +1.4(1) eV relative to  $\text{N}_2$   $\pi^*$ ; solid, -0.6 eV relative to 3p ( $\nu = 0$ ) of  $\text{N}_2$ . <sup>d</sup> From XPS.<sup>52</sup>

With regard to the computed spectra, the  $\pi^*$  features in the C 1s spectra are reproduced reasonably well once a rigid energy shift between calculation and experiment is made. In particular, the calculation for Gly-Gly predicts a separation of 0.3 eV between the C 1s(CONH)  $\rightarrow$   $\pi^*_{\text{CONH}}$  and C 1s(COOH)  $\rightarrow$   $\pi^*_{\text{COOH}}$  transitions. The general agreement between the theoretical calculations and the experimental ISEELS data at higher energies in the C 1s (and N 1s and O 1s) core ionization continua is less satisfactory, although there are correlations between the more intense computed states and the near-threshold broad resonances, which are interpreted as excitations to quasi-bound  $\sigma^*$  states. The positions of these features can be approximately correlated with local bond lengths.<sup>2</sup> A detailed discussion of the present computational results with respect to the conformations of glycine and glycyl-glycine is given in section 3.3.

The shoulders at 289.3 eV in the C 1s spectrum of glycine and at 289.7 eV in the C 1s spectrum of glycyl-glycine are tentatively assigned as C 1s  $\rightarrow$  3p (Rydberg) transitions. The theoretical GSCF3 calculations for both glycine and glycyl-glycine predict the lowest energy C 1s excitations to virtual

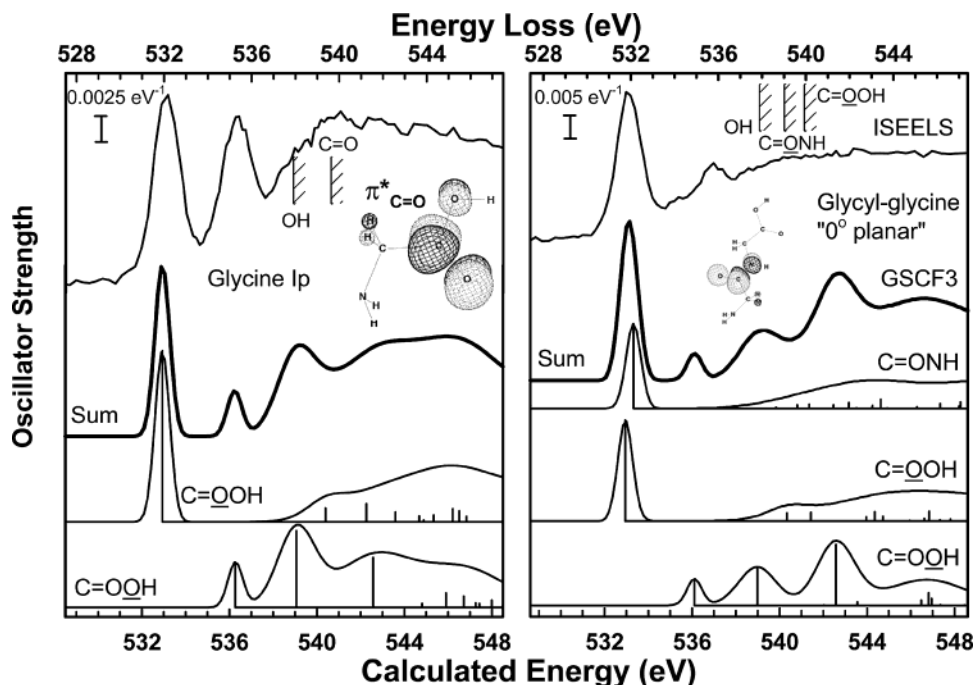
**TABLE 5: Energies, Term Values, and Proposed Assignments for N 1s Spectral Features of Glycyl-Glycine (Gas, Solid)**

gas <sup>a</sup>		solid		assignment (final orbital)	
$E$ (eV), $\pm 0.1$	term value <sup>b</sup> (eV)	$E$ (eV), $\pm 0.1$		$\text{NH}_2$	CONH-R
401.0 (sh)	4.6			3s	
401.9 <sup>c</sup>		401.5 <sup>c</sup>	4.0		$\pi^*_{\text{CONH}}$
402.8 (sh)	2.8			3p	
405.6 <sup>d</sup>				IP	
405.7	-0.1	406.7		$\sigma^*_{\text{C-N}}$ , $\sigma^*_{\text{N-H}}$	
405.9 <sup>d</sup>				IP	
408.6	-3			$\sigma^*_{\text{N-H}}$	
413	-7.4	413		$\sigma^*_{\text{CONH}}$	

<sup>a</sup> The gas-phase spectra can be obtained in electronic form from <http://unicorn.mcmaster.ca/corex/cedb-title.html>. <sup>b</sup> Term value = ionization potential (IP) - excitation energy. <sup>c</sup> Calibration: gas, +0.8(1) eV relative to  $\text{N}_2$   $\pi^*$ ; solid, -5.8 eV relative to 3p ( $\nu = 0$ ) of  $\text{N}_2$ . <sup>d</sup> Estimated from IPs of similar species.<sup>52</sup> Trigly peaks observed at 401.5 (-5.8 eV relative to 3p ( $\nu = 0$ ) of  $\text{N}_2$ ), 406.6, 413.

valence levels of  $\sigma$  symmetry to be of  $\sigma^*_{\text{C-H}}/\sigma^*_{\text{C-N-H}}/\sigma^*_{\text{N-H}}$  character. These can be assigned to the intensity at 290.8 eV in the glycine spectrum and at 290.7 eV in the glycyl-glycine spectrum. The peak in the glycine spectrum at 294 eV can be assigned as a C 1s  $\rightarrow$   $\pi^*_{\text{CH}_2\text{NH}_2}$  transition, with some C 1s  $\rightarrow$   $\sigma^*_{\text{O-H}}$  character, the peak at 297.0 eV as C 1s  $\rightarrow$   $\sigma^*_{\text{C-OH}}$  and C 1s  $\rightarrow$   $\sigma^*_{\text{C-C}}$  transitions, and the peak at 303 eV as mainly C 1s  $\rightarrow$   $\sigma^*_{\text{C-O}}$  with small contributions from  $\sigma^*_{\text{C-C}}$  transitions. In the glycyl-glycine spectrum, the transitions analogous to the last three features in the glycine spectrum are located at 293.1, 296, and 303 eV, respectively. In the case of glycyl-glycine, there is relatively more spectral intensity in the  $\sim 291$ – $293$  eV energy region than there is in the glycine spectrum. This can be rationalized by the GSCF3 calculations, which indicate that there are additional C 1s  $\rightarrow$   $\sigma^*_{\text{C(amide)-C(NH)}}$  and C 1s  $\rightarrow$   $\sigma^*_{\text{C(COOH)-NH}}$  transitions in this region for glycyl-glycine.





**Figure 6.** O 1s oscillator strength spectra of gaseous glycine and glycyglycine compared to computed spectra of the Ip conformer of glycine and the “0° planar” conformer of glycyglycine. See caption to Figure 4 for further details. The computed energy scales are plotted with a shift relative to experiment of 1.0 (Gly) and 1.2 eV (Gly–Gly) in order to align the first peak. The sketches are for the  $(O\ 1s(COOH)^{-1}, \pi^*_{COOH})$  (Gly) and the  $(O\ 1s(CONH)^{-1}, \pi^*_{CONH})$  (Gly–Gly) molecular orbitals.

**TABLE 6: Energies, Term Values, and Proposed Assignments for O 1s Spectral Features of Glycine (Gas, Solid)**

gas <sup>a</sup>		solid	
$E$ (eV), $\pm 0.1$	term value <sup>b</sup> (eV) $T_{C=O}$ $T_{O-H}$	$E$ (eV), $\pm 0.1$	assignment (final orbital) C=O   O–H
532.1 <sup>c</sup>	6.1	532.4 <sup>c</sup>	$\pi^*_{C=O}$ ( $\pi^*_{C=O}$ in carboxylate form)
535.3		536.8	$\pi^*_{C=O}$
538.2 <sup>d</sup>			IP
540.0 <sup>d</sup>			IP
540	–1.8		$\sigma^*_{O-H}$
544 (sh)	–4	545	$\sigma^*_{C-O}$

<sup>a</sup> The gas-phase spectra can be obtained in electronic form from <http://unicorn.mcmaster.ca/corex/cedb-title.html>. <sup>b</sup> Term value = ionization potential (IP) – excitation energy. <sup>c</sup> Calibration: gas, –3.31(1) eV relative to CO<sub>2</sub>  $\pi^*$ ; solid, +1.60(8) relative to O<sub>2</sub>  $\pi^*$ . <sup>d</sup> From XPS.<sup>52</sup>

**3.2.2. N 1s Spectra.** Figure 5 plots the N 1s spectra computed by GSCF3 for the Ip conformer of glycine and the “0° planar” conformation of glycyglycine. The computed N 1s spectra of the other glycine conformers, Iip, IIIp, and IVn, show only small differences in peak positions and intensities. The experimental spectra of glycine and glycyglycine are also included for comparison. Note that both energies and intensities (full molecule) are on absolute scales for both the computed and experimental results (offsets are used for clarity). The energies, term values, and proposed assignments for the N 1s spectral features of glycine are presented in Table 4, while the values for glycyglycine are presented in Table 5. Table 8 summarizes the calculated transition energies, oscillator strengths, and orbital characters of selected low-energy states of glycine, while Table 9 summarizes the corresponding quantities for the “0° planar” conformation of glycyglycine. A detailed discussion of the present results with respect to the conformations of glycine and glycyglycine is given in section 3.3.

The lowest energy feature, observed at 401.3 eV (shoulder) in the experimental N 1s spectrum of glycine, is associated with the N 1s  $\rightarrow$  3s (Rydberg) transition. A feature associated with the 3s Rydberg transition is observed in the glycyglycine spectrum, but is relatively less intense. The feature located at 402.5 eV in glycine is assigned as the N 1s  $\rightarrow$  3p (Rydberg) transition. Note that, in Figure 2, the N 1s spectrum of solid glycine does not show any peaks in the 401–403 eV energy region, confirming these assignments since Rydberg transitions are quenched in the solid state. Our calculations for glycine do not show the Rydberg features since these are only predicted by the GSCF3 method if a much larger basis set is used and the IVO approximation is not used. Relative to glycine, the N 1s spectrum of Gly–Gly(g) differs most around 402 eV. In particular, there is a strong, narrow peak at 401.9 eV in Gly–Gly(g). This is assigned to the N 1s(amide)  $\rightarrow$   $\pi^*_{CONH}$  transition, reflecting the fact that there is a significant contribution of the N 2p atomic orbital to the  $\pi^*_{CONH}$  molecular orbital. Similar N 1s  $\rightarrow$   $\pi^*_{CONH}$  transitions are seen in the N 1s spectrum of formamide,<sup>55</sup> and a corresponding peak is very prominent in the N 1s spectra of the solid species (Gly–Gly, Gly–Gly–Gly, fibrinogen) as well as albumin.<sup>47–49</sup>

The broader peaks located at 405.3 eV in glycine and 405.7 eV in glycyglycine are assigned as N 1s  $\rightarrow$   $\sigma^*_{C-N}$  and N 1s  $\rightarrow$   $\sigma^*_{N-H}$  transitions. The shorter length of the partially conjugated amide C–N bond gives rise to a shift to higher transition energy for  $\sigma^*_{CONH}$  compared with  $\sigma^*_{C-NH_2}$ . There are broad peaks at 405.7, 408.6, and 413 eV in the N 1s spectrum of Gly–Gly. These features are attributed to N 1s  $\rightarrow$   $\sigma^*_{C-NH_2}$ ,  $\sigma^*_{C-NH}$ , and  $\sigma^*_{CONH}$  transitions. Since the N 1s  $\rightarrow$   $\sigma^*_{CONH}$  transition is not present in glycine, the Gly–Gly spectrum shows higher intensity in these regions than the glycine spectrum (this is readily observed in the difference spectrum plotted in Figure 10).

**3.2.3. O 1s Spectra.** Figure 6 plots the O 1s spectra computed by GSCF3 for the Ip conformer of glycine and the “0° planar” conformation of glycyglycine. The computed O 1s spectra

**TABLE 7: Energies, Term Values, and Proposed Assignments for Features in the O 1s Spectra of Glycyl–Glycine**

gas <sup>a</sup>				solid			
<i>E</i> (eV), ±0.1	term value <sup>b</sup> (eV)			<i>E</i> (eV), ±0.1	assignment (final orbital)		
	<i>T</i> <sub>CONH</sub>	<i>T</i> <sub>COOH</sub>	<i>T</i> <sub>OH</sub>		C=ONH	C–OOH	O–H
531.7 <sup>c,d</sup> (fit)		7.5				$\pi^*_{\text{COOH}}$	
532.2 <sup>c,d</sup> (fit)	5.8			532.2 <sup>c</sup>	$\pi^*_{\text{CONH}}$		
534.0 (w)	4.0				$\sigma^*_{\text{CNO}}$	$\sigma^*_{\text{CNO}}$	
535.7			4.3	535.2	IP		$\pi^*_{\text{COOH}}$
538 <sup>e</sup>						IP	
539.2 <sup>e</sup>						IP	
539 (br)			1	538.6 (br)			$\sigma^*_{\text{O–H}}$
540.0 <sup>e</sup>							IP
542 (br)			–2	542 (br)			$\sigma^*_{\text{C–OH}}$
546 (br)	–6	–5				$\sigma^*_{\text{C–O}}$	

<sup>a</sup> The gas-phase spectra can be obtained in electronic form from <http://unicorn.mcmaster.ca/corex/cedb-title.html>. <sup>b</sup> Term value = ionization potential (IP) – excitation energy. <sup>c</sup> Calibration: gas, center of first peak at –3.7(1) eV relative to CO<sub>2</sub>  $\pi^*$  (the centroid of the first peak in Gly–Gly is shifted –0.3 eV relative to that in glycine); solid, +1.41(8) relative to O<sub>2</sub>  $\pi^*$ . <sup>d</sup> The comparison to glycine (Figure 3) reveals that the first peaks contains two components. These energies are based on spectral subtraction and fits to two Gaussian lines. <sup>e</sup> Estimated from IPs of similar species.<sup>52</sup> Trigly peaks observed at 532.2 (calibration, +1.36(8) relative to O<sub>2</sub>  $\pi^*$ ), 535.2, 538.8 (br), 542 (br).

**TABLE 8: Selected Eigenvalues, Oscillator Strengths, and Orbital Characters for Computed Core Excited States of the Ip Conformer of Glycine**

site	IP	character	orbital no.	– $\epsilon$ (eV) <sup>a</sup>	<i>f</i> (×10 <sup>–2</sup> )
C <sub>1</sub> (CH <sub>2</sub> )	293.65	$\pi^*(\text{C=OOH})$	21	2.95	0.02
		$\sigma^*(\text{C–H})/\sigma^*(\text{C–N})$	23	0.01	1.60
		$\sigma^*(\text{C–N})$	22	–0.56	0.63
		$\sigma^*(\text{C–N})/\sigma^*(\text{C–H})$	26	–0.80	1.93
		$\pi^*(\text{CH}_2\text{–NH}_2)$	30	–2.63	2.13
C <sub>2</sub> (C=O)	297.78	$\pi^*(\text{C=OOH})$	21	6.47	7.56
		$\sigma^*(\text{O–H})$	22	–0.30	0.72
		$\sigma^*(\text{C–OH})$	33	–3.05	2.47
		$\sigma^*(\text{N–H})$	24	–0.54	0.83
N	406.37	$\pi^*(\text{C=OOH})$	21	1.71	0.00
		$\sigma^*(\text{N–H})$	24	–0.54	0.83
		$\sigma^*(\text{N–H})$	25	–1.26	2.48
		$\sigma^*(\text{C–N})$	30	–2.61	2.45
O <sub>1</sub> (C=O)	540.48	$\pi^*(\text{C=O})$	21	4.20	0.41
		$\sigma^*(\text{O–H})$	24	1.42	1.72
		$\sigma^*(\text{C–OH})$	31	–2.09	2.24
		$\sigma^*(\text{C–O})$	21	5.93	1.53
O <sub>2</sub> (OH)	538.86	$\pi^*(\text{C=O})$	21	5.93	1.53
		$\sigma^*(\text{O–H})$	22	–1.54	0.31
		$\sigma^*(\text{C–OH})$	23	–3.41	0.81

<sup>a</sup> The predicted spectra were generated using Gaussian lines of area given by the computed oscillator strength (*f*) and widths of 0.8 ( $\epsilon < -2$  eV), 2.0 ( $-2.0 < \epsilon < 2.0$  eV), 4.0 ( $2.0 < \epsilon < 8.0$  eV), and 6.0 eV ( $\epsilon > 8.0$  eV).

of other conformers, IIp, IIIp, and IVn, show only small differences in peak positions and intensities. The experimental O 1s spectra of glycine and glycyl–glycine are also included for comparison. Note that both energies and intensities are on absolute scales for both the computed and experimental results (offsets are used for clarity). The energies, term values, and proposed assignments for the spectral features of glycine are presented in Table 6, while the values for glycyl–glycine are presented in Table 7. Table 8 summarizes the computed energies, oscillator strengths, and orbital characters of selected low-energy states of glycine, while Table 9 summarizes the corresponding quantities for the “0° planar” conformation of glycyl–glycine. A detailed discussion of the present results with respect to the conformations of glycine and glycyl–glycine is given in section 3.3.

The O 1s spectrum of glycine(g) shows two strong peaks. The lower energy feature at 532.1 eV is the O 1s(C=O) →  $\pi^*_{\text{COOH}}$  transition, while the higher energy feature at 535.3 eV is the O 1s(OH) →  $\pi^*_{\text{COOH}}$  transition. The latter feature is shifted to considerably higher energy, in part due to the chemical shift

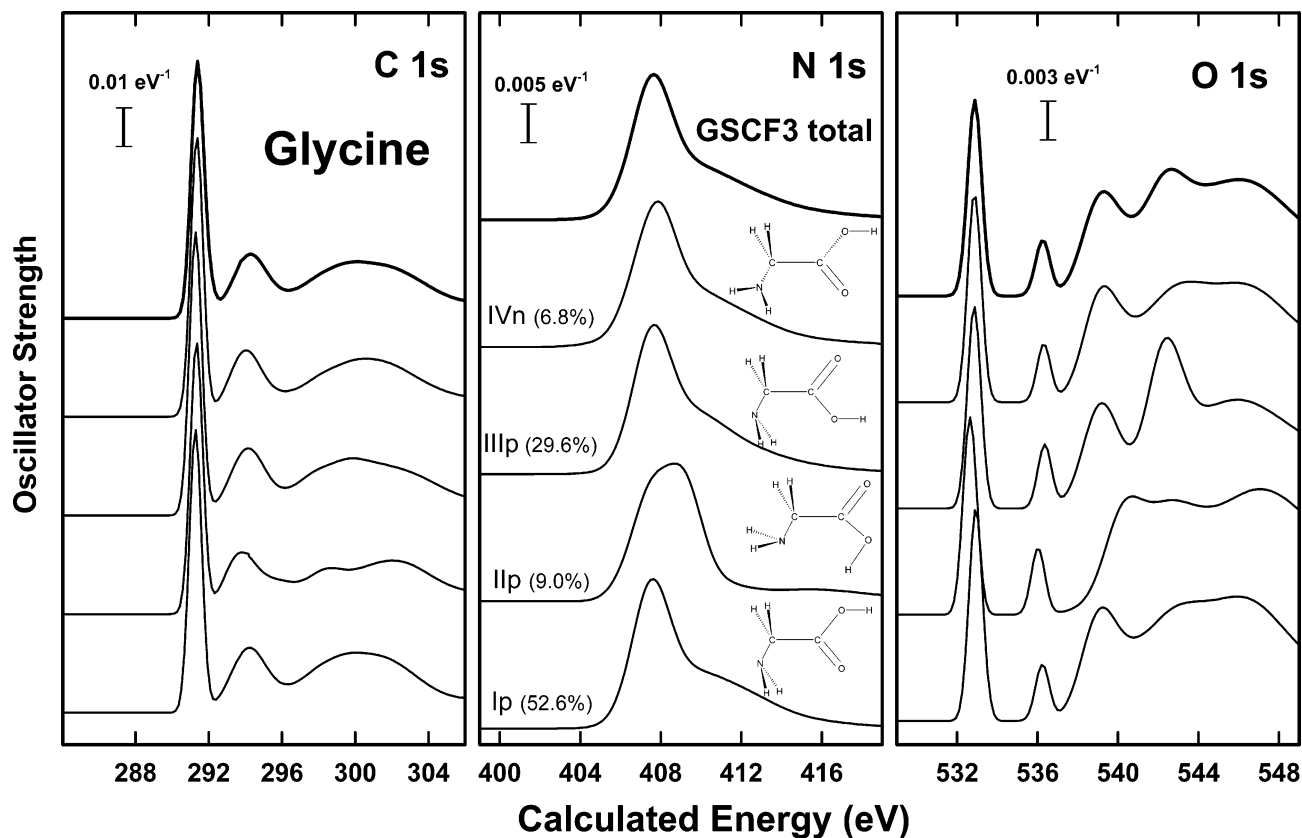
**TABLE 9: Selected Eigenvalues, Oscillator Strengths, and Orbital Characters for Computed Core Excited States of the “0° Planar” Conformer of Glycyl–Glycine**

site	IP	character	orbital no.	– $\epsilon$ (eV)	<i>f</i> (×10 <sup>–2</sup> )
C <sub>1</sub> (CH <sub>2</sub> –NH <sub>2</sub> )	293.63	$\pi^*(\text{CONH})$	37	1.52	0.03
		$\sigma^*(\text{C–NH}_2)$	40	–0.30	0.90
		$\sigma^*(\text{C–H}_2)$	41	–0.37	1.60
		$\sigma^*(\text{C–NH}_2)$	50	–1.20	2.39
		$\pi^*(\text{CONH})$	41	4.68	7.10
C <sub>2</sub> (CONH)	296.04	$\pi^*(\text{CONH})$	36	3.21	0.05
C <sub>3</sub> (CH <sub>2</sub> –COOH)	293.77	$\pi^*(\text{CONH})$	37	0.56	0
		$\sigma^*(\text{C–NH})$	56	0.01	2.67
		$\sigma^*(\text{C–H}_2)$	40	–0.66	1.50
		$\sigma^*(\text{C–C})$	41	–1.57	1.60
		$\pi^*(\text{COOH})$	37	6.73	7.31
C <sub>4</sub> (COOH)	298.39	$\pi^*(\text{COOH})$	37	6.73	7.31
N <sub>1</sub> (NH <sub>2</sub> )	405.79	$\pi^*(\text{CONH})$	37	0.19	0
		$\sigma^*(\text{C–N})/$ $\sigma^*(\text{N–H}_2)$	45	–1.62	2.31
		$\sigma^*(\text{C–N})$	48	–3.03	2.32
		$\pi^*(\text{CONH})$	37	2.27	0.45
		$\pi^*(\text{COOH})$	36	2.06	0.08
N <sub>2</sub> (CONH)	406.86	$\sigma^*(\text{O–H})$	44	–0.52	1.17
		$\sigma^*(\text{C–C})$	57	–1.02	1.16
		$\sigma^*(\text{N–H}_2)$	57	–1.02	1.16
		$\pi^*(\text{CONH})$	37	3.92	1.28
		$\pi^*(\text{COOH})$	37	6.20	1.52
O <sub>1</sub> (CONH)	537.24	$\pi^*(\text{CONH})$	37	3.92	1.28
O <sub>2</sub> (C=OOH)	539.15	$\pi^*(\text{COOH})$	37	6.20	1.52
O <sub>3</sub> (OH)	540.62	$\pi^*(\text{COOH})$	36	4.52	0.40
		$\sigma^*(\text{O–H})$	42	1.64	1.45
		$\sigma^*(\text{C–OH})$	73	–1.95	2.30

<sup>a</sup> The predicted spectra were generated using Gaussian lines of area given by the computed oscillator strength (*f*) and widths of 0.8 ( $\epsilon < -2$  eV), 2.0 ( $-2.0 < \epsilon < 2.0$  eV), 4.0 ( $2.0 < \epsilon < 8.0$  eV), and 6.0 eV ( $\epsilon > 8.0$  eV).

of the O 1s level, and in part because the core hole shielding of an atom distant from the main electron density of the optical molecular orbital is less than that for core excitation at an atom close to the main electron density.<sup>56</sup> According to the GSCF3 results, the broad feature at 540 eV is due to O 1s →  $\sigma^*_{\text{O–H}}$  transitions, while the broad peak at ~544 eV in glycine is associated with O 1s →  $\sigma^*_{\text{C–O}}$  transitions.

The O 1s spectrum of Gly–Gly is quite different from that of Gly. Relative to the O 1s continuum intensity, the intensity of the lower energy peak in Gly–Gly is much increased. This reflects the fact that Gly–Gly has two C=O double bonds versus only one in Gly. Note that the apparently single  $\pi^*$  peak at 532 eV in Gly–Gly consists of an unresolved overlap of the O 1s(carboxyl) →  $\pi^*_{\text{COOH}}$  and the O 1s(amide) →  $\pi^*_{\text{CONH}}$



**Figure 7.** Computed C 1s, N 1s, and O 1s spectra of the four most common conformers of glycine. The top bold curve is a weighted sum using the predicted relative abundances at 400 K.

transitions. As shown in Figure 6, the GSCF3 calculations successfully predict both the spectral overlap and the increased intensity of the lowest  $\pi^*$  peak relative to Gly. However, the peak energy ordering deduced from experiment (via comparison to Gly in the spectral subtraction, see Figure 10) is O 1s(amide)  $\rightarrow \pi^*_{\text{CONH}} < \text{O 1s(carboxyl)} \rightarrow \pi^*_{\text{COOH}}$ , whereas the GSCF3 calculations predict the reverse order. It is interesting to note that, experimentally, both the C 1s and O 1s  $\rightarrow \pi^*_{\text{CONH}}$  transitions occur at slightly lower energies than the C 1s and O 1s  $\rightarrow \pi^*_{\text{COOH}}$  transitions, whereas theoretically, the order reverses between the C 1s and O 1s shells. The transition energies involved are very close though, and depend on both the energies of the final orbitals and the energy required to create the core holes, so it is not surprising that the theoretical method struggles in this case to get the energy ordering correct.

The O 1s(OH)  $\rightarrow \pi^*_{\text{COOH}}$  peak of glycyglycine occurs at 535.7 eV. This feature is weaker and shifted 0.4 eV to higher energy relative to its counterpart in Gly. The calculations indicate this transition should occur at 536 eV, and at the same energy in both molecules. In comparison with glycine, there is one additional feature in the O 1s spectrum of Gly-Gly at 534.0 eV. According to the GSCF3 calculations, this is probably associated with the O 1s  $\rightarrow \sigma^*_{\text{CNO}}$  transitions from both the carbonyl and the hydroxyl oxygen atoms.

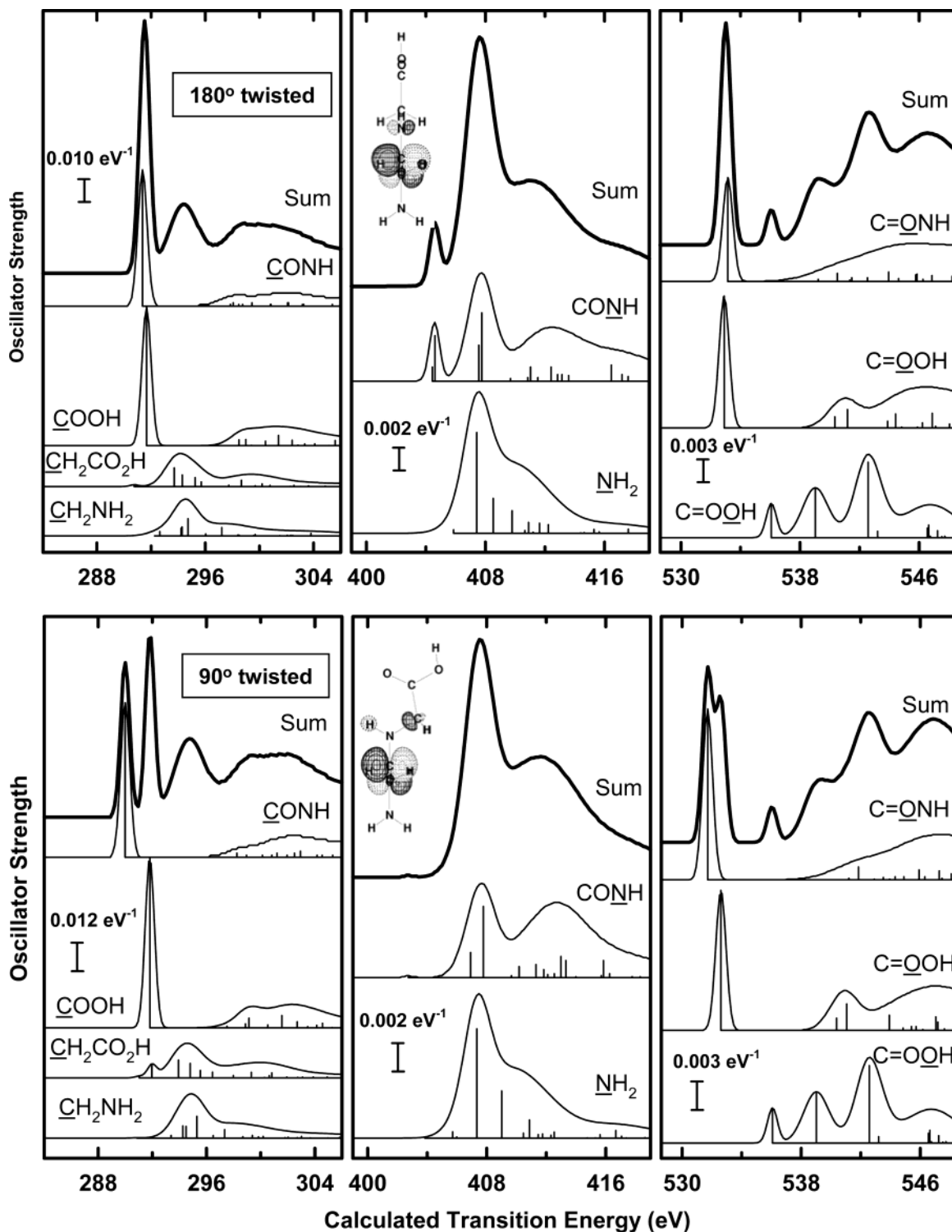
**3.3. Influence of Molecular Conformation on the Glycine and Glycyl-Glycine Inner-Shell Spectra.** Figure 7 presents the results of GSCF3 calculations of the C 1s, N 1s, and O 1s spectra of the Ip, IIp, IIIp, and IVn conformers of glycine (see Figures 4–6 for comparison to experiment). There are few differences between the calculated spectra for the different conformers, suggesting that inner-shell excitation is not very sensitive to the molecular conformation of glycine. This lends confidence that the glycine calculations reproduce the main

spectral features without being unduly influenced by small conformational or symmetry changes. The largest differences occur for conformer IIp, especially in the N 1s and O 1s spectra. This conformer is expected to comprise only  $\sim 9\%$  of the molecular population at the experimental temperature.

Figure 8 presents the results of GSCF3 calculations for C 1s, N 1s, and O 1s excitation of glycyglycine in the “180° twisted” (upper panel) and “90° twisted” (lower panel) conformations described in section 2.4. The corresponding spectra of the “0° planar” conformation have already been presented in Figures 4–6. In this case, there are very significant differences between the calculated spectra for the “90° twisted” geometry and the “0° planar” or “180° twisted” geometries, caused mainly by changing interactions between the  $\pi^*$  systems of the amide and carboxyl groups. The computed spectra of the “0° planar” (Figures 4–6) and “180° twisted” (upper panel of Figure 8) geometries are quite similar, perhaps surprisingly given that the overall geometry has changed so drastically. Although the details of the individual transitions are slightly changed, the resultant spectra are generally very similar to each other because the  $\pi^*$  systems of the amide and carboxyl groups are planar in both cases and can therefore still interact with each other in similar ways. However, the calculated spectra for the “90° twisted” geometry (lower panel of Figure 8) show quite different spectral features, with significant changes in all three core-level spectra. In particular:

(a) The  $\pi^*$  region of the C 1s spectrum shows two distinct peaks, which arise from the C 1s  $\rightarrow \pi^*_{\text{CONH}}$  and the C 1s  $\rightarrow \pi^*_{\text{COOH}}$  transitions having significantly different energies in the “90° twisted” species.

(b) The N 1s spectrum is calculated to have negligible intensity for the N 1s  $\rightarrow \pi^*_{\text{CONH}}$  transition, associated with the fact that the N 2p orbital cannot interact with the  $\pi^*_{\text{CO}}$  orbital.

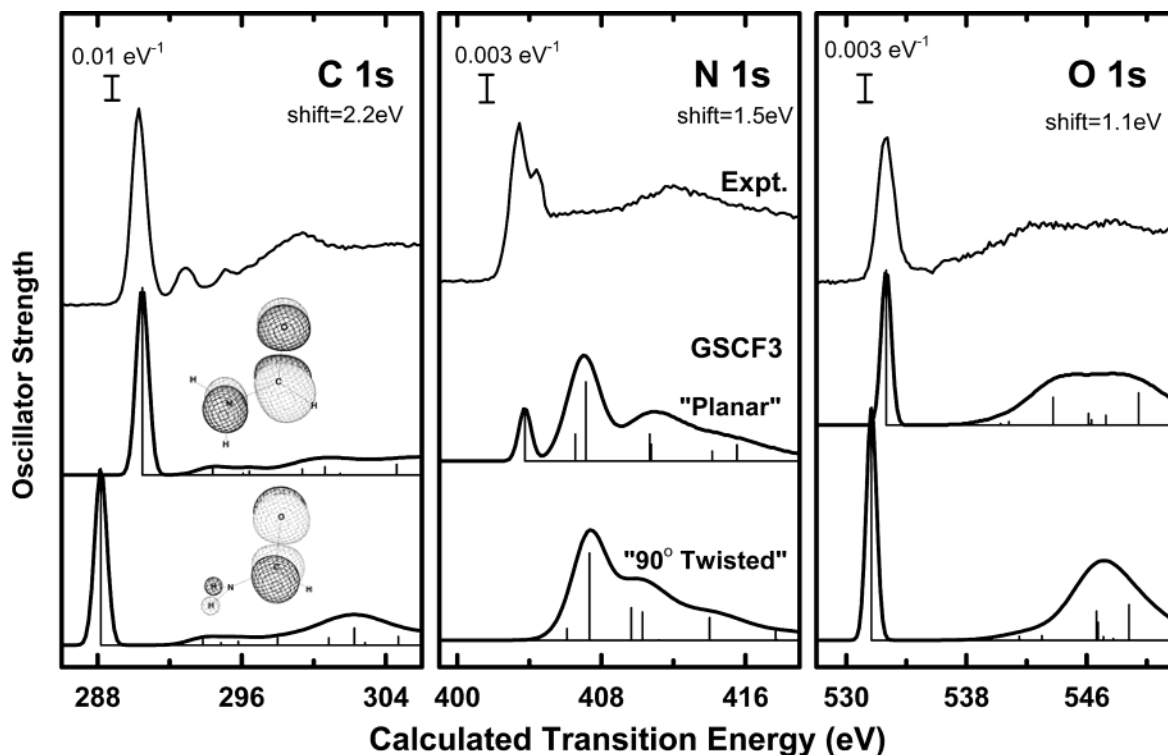


**Figure 8.** (Upper) Computed C 1s, N 1s, and O 1s spectra of glycyl-glycine in the “180° twisted” geometry, in which the top part of the molecule is rotated 180° relative to planar geometry. (Lower) Computed C 1s, N 1s, and O 1s spectra of glycyl-glycine in the “90° twisted” geometry. The MOs sketched are those of the  $\pi^*_{\text{CO}}$  orbital with amide N 1s excitation. The spectra for each distinct site, as well as the weighted sum, are indicated.

(c) The  $\pi^*$  region in the O 1s spectrum exhibits two peaks, for the same reason as in the C 1s signal, although in this case the peaks are closer together than in the C 1s region.

While the disappearance of the N 1s  $\rightarrow \pi^*$  transition in the “90° twisted” geometry was expected from simple orbital overlap considerations, the large changes in the C 1s and O 1s spectra were unexpected. These results indicate that the energy shift associated with delocalization of the  $\pi^*_{\text{CO}}$  orbital onto the amide N atom, with exactly the same electrostatic environment

of the carbonyl group, is as large as shifts associated with changes in the electrostatic environment of a carbonyl group within a planar geometry.<sup>53,54</sup> Because the geometry dependence of the computed C 1s and O 1s spectra of glycyl-glycine is so dramatic, we have investigated further by performing GSCF3 calculations on formamide, the simplest species with a  $\pi^*_{\text{CONH}}$  orbital. We have computed the C 1s, N 1s, and O 1s spectra of HCONH<sub>2</sub> in the equilibrium planar geometry and in a “90° twisted” geometry, where the two H atoms of the NH<sub>2</sub> group



**Figure 9.** Computed C 1s spectra of formamide in “planar” and “90° twisted” (about the C–N bond) geometries (see text for details), compared to the experimental oscillator strength spectrum derived from ISEELS.<sup>55</sup> The MOs sketched are those of the  $\pi^*_{\text{CO}}$  orbital with C 1s excitation in the two geometries.

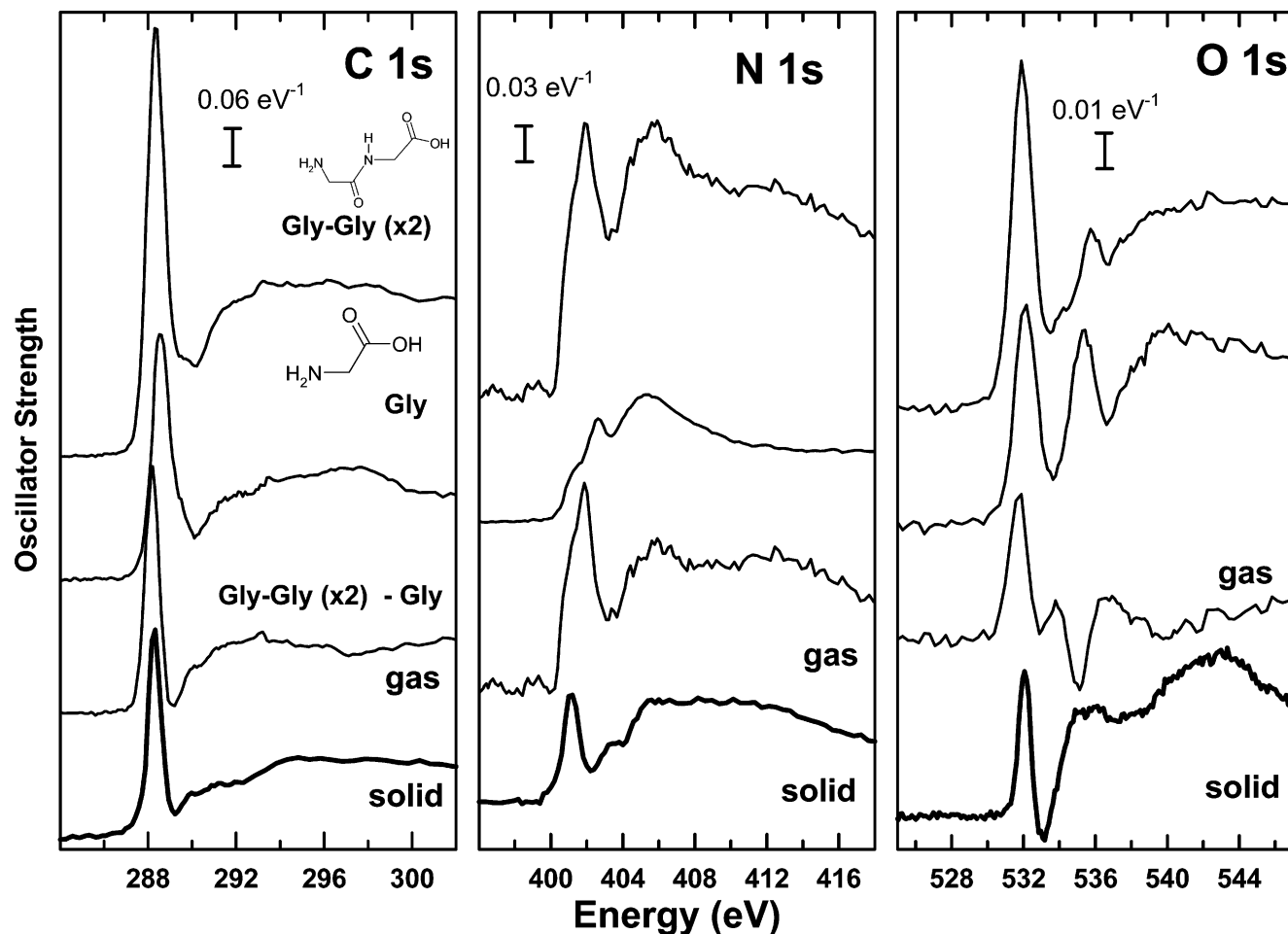
are twisted 90° out of the plane of the HCO moiety. This is equivalent to rotation about the C–N bond, as was done for glycyl–glycine. A computer-optimized geometry using a 6-31G basis set was used. For the 90° twisted geometry, only the twist angle was constrained in the geometry optimization. We have also carried out the core excitation calculation with a range of basis sets. We found the answer to be essentially independent of small changes in bond lengths or basis set details. The results of these calculations are presented in Figure 9 in comparison to the experimental spectrum of formamide.<sup>55</sup> Similar effects are observed in the conformational dependence of the C 1s, N 1s, and O 1s spectra of formamide as were observed in the glycyl–glycine spectra. The twisted geometry gives rise to a C 1s  $\rightarrow \pi^*_{\text{CONH}}$  transition at significantly lower energy than the planar C 1s  $\rightarrow \pi^*_{\text{CONH}}$  transition; the N 1s  $\rightarrow \pi^*_{\text{CONH}}$  transition loses all intensity in the twisted geometry; and the O 1s  $\rightarrow \pi^*_{\text{CONH}}$  transition occurs at lower energy for the twisted case than the planar case. Plots of the  $\pi^*$  orbital for C 1s excited formamide in the planar and twisted geometries (Figure 9) demonstrate that, in the planar formamide molecule, the  $\pi^*$  orbital is both C–N and C–O antibonding, while in the 90° twisted case it has only C–O antibonding character. This results in this orbital being of significantly lower energy (by  $\sim 1$ – $1.5$  eV) in the twisted forms of formamide and glycyl–glycine than in the planar forms, thus giving rise to C 1s  $\rightarrow \pi^*_{\text{CONH}}$  transitions at significantly lower energies. The same features are observed in the orbital pictures for O 1s excitation (not shown).

These computational results indicate there is a significant conformational dependence of the inner-shell spectra (all three core edges) of Gly–Gly and presumably of all peptides, in particular a sensitivity to twist angle at the peptide bond. This suggests that it may be possible to use inner-shell excitation spectra as a probe of the conformation of amide linkages in peptides or even proteins. These effects would be most easy to

see in species where the number of amide groups in a twisted geometry was large compared to those in a planar geometry.

**3.4. Inner-Shell Spectra of the Peptide Bond.** To more clearly reveal the effect of peptide bond formation, Figure 10 compares the C 1s, N 1s, and O 1s spectra of gaseous glycine with the spectrum of glycyl–glycine, along with the difference spectra computed as  $2(\text{Gly–Gly}) - \text{Gly}$  for C 1s and N 1s, but as  $1.5(\text{Gly–Gly}) - \text{Gly}$  for the O 1s spectrum, to account for correct number of atoms of each type. This procedure isolates the spectral signature of the amide C 1s contribution, assuming there is very little difference in peak positions or relative intensities of the  $\alpha$ -C or the terminal COOH carbon signals. These are considered reasonable assumptions, given that these sites are isolated from the change in peptide bond formation by at least one saturated C–C bond.

These difference spectra clearly reveal the effect of peptide bond formation on inner-shell excitation. The C 1s difference signal for the two gas-phase species is dominated by a narrow, strong band at 288.2 eV. The position of this band is in good agreement with that found in the C 1s spectra of proteins,<sup>44,45</sup> where the peptide bond  $\pi^*_{\text{C=O}}$  transition dominates. A very similar difference spectrum is found between the C 1s spectra of solid glycine and solid glycyl–glycine. The situation for the difference in the gas-phase N 1s spectra is somewhat more complicated because the  $\pi^*_{\text{CONH}}$  peak overlaps with Rydberg transitions. However, the difference spectrum is clearly dominated by a strong, narrow band at 401.7 eV. The corresponding difference of the N 1s spectra of the solid species is also very much dominated by the  $\pi^*_{\text{CONH}}$ , which peaks at 401.1 eV. This is about 0.6 eV lower than in the difference of the gas-phase spectra. This suggests that the strongest discrete peak in the N 1s spectrum of Gly–Gly may actually be Rydberg in origin, and the N 1s  $\rightarrow \pi^*_{\text{CONH}}$  peak could actually be the shoulder at 401.1 eV. Generally speaking, the position of these  $\pi^*_{\text{CONH}}$  peaks are in good agreement with the positions of peaks found



**Figure 10.** Experimental C 1s, N 1s, and O 1s oscillator strength spectra of gaseous glycine and glycyl-glycine, and their difference, which is an estimate of the spectral signature of the peptide bond. The weightings used to account for stoichiometry when taking the difference are explained in the text. The bottom curve in each panel (bold) is the corresponding difference signal from the solid-state STXM measurements.

in the N 1s spectra of proteins,<sup>47–49</sup> where the peptide bond N 1s  $\rightarrow \pi^*_{\text{CONHR}}$  dominates relative to the contribution from the terminal  $\text{NH}_2$ . It is fair to say that the appearance of the  $\pi^*_{\text{CONHR}}$  peak with formation of the amide band is the single most characteristic feature of peptide formation. This is particularly clear in the solid state, where the Rydberg transitions do not appear. The difference in the O 1s spectra of the gaseous Gly and Gly-Gly species is the most complicated, but it can be readily interpreted when one realizes there is removal of an OH bond as well as the introduction of the amide group upon peptide bond formation. The negative-going peak at  $\sim 535$  eV arises from loss of one of the two O 1s(OH)  $\rightarrow \pi^*_{\text{COOH}}$  transitions, as well as shifts in the energy of the remaining O 1s(OH)  $\rightarrow \pi^*_{\text{COOH}}$  contribution relative to that in Gly. In the solid, the oxygen atoms are likely in a symmetric environment in the zwitterion form of glycine, so there is no O 1s(OH)  $\rightarrow \pi^*_{\text{COOH}}$  transition. Thus, the difference of the solid spectra is predominantly a single peak at 532.1 eV, corresponding to the O 1s  $\rightarrow \pi^*_{\text{CONH}}$  transition at the peptide bond.

#### 4. Summary

C 1s, N 1s, and O 1s inner-shell electron energy-loss spectra of gaseous glycine and glycyl-glycine have been compared to X-ray absorption spectra of solid glycine, glycyl-glycine, and triglycine. Differences in the gas- and solid-state spectra have been rationalized in terms of differences in the molecular structure of the gas and solid, along with intermolecular solid-state effects. Comparison of the spectra of glycine and glycyl-

glycine clearly identifies changes associated with the structural changes that occur upon peptide bond formation. These include a clear broadening and a  $\sim 0.3$  eV shift of the C 1s  $\rightarrow \pi^*_{\text{C=O}}$  peak and introduction of a new pre-edge feature in the N 1s spectrum, both due to 1s  $\rightarrow \pi^*_{\text{amide}}$  transitions introduced with formation of the peptide bond. Comparison to theoretical calculations supports the spectral interpretation developed from the direct spectral comparison and emphasizes the need to consider shifts in both core and  $\pi^*$  energies in spectral assignments. Theoretical investigations of the dependence of the inner-shell spectra on molecular geometry have revealed a sensitivity to the geometry around peptide bonds which may have potential use as a probe of molecular conformation of the amide groups in peptides or proteins. Although prior comparisons of the C 1s NEXAFS of mono-, di- and triamino acid systems have been presented,<sup>15</sup> that work considered only the C 1s edge, and the spectral interpretation was rather limited compared to the present treatment. Further investigation of the spectra of tri- and larger peptides are needed to reinforce these spectral interpretations and to provide further insight into how NEXAFS spectroscopy can be used to study peptide bond conformation. Such studies will help to guide future applications of NEXAFS microscopy to studies of short-chain peptides in biological systems.

#### 5. Note Added in Proof

After the review of this paper was completed, we were provided with a preprint by Zubavichus et al.<sup>57</sup> of a parallel

study of partial-yield NEXAFS spectra of thin films of Gly, Gly–Gly, Gly–Gly–Gly, and the cyclic dimer, 2,5-diketopiperazine (DKP). Zubavichus et al.<sup>57</sup> interpret the  $\pi^*_{\text{CO}}$  region of their C 1s spectra as consisting of two overlapping peaks at 288.3 and 288.7 eV in all three noncyclic species, and only a single 288.3 eV peak in DKP. On the basis of this, they attribute the 288.3 eV signal to the C 1s  $\rightarrow \pi^*_{\text{CONH}}$  amide signal (as in this work) and the 288.7 eV signal to C 1s  $\rightarrow \pi^*_{\text{COO}^-}$  transitions, with the C 1s  $\rightarrow \pi^*_{\text{COOH}}$  transition considered to overlap the 288.3 eV  $\pi^*_{\text{CONH}}$  amide signal. On the basis of these assignments, they interpret the changes in their spectra in terms of different amounts of carboxylate, carboxylic acid, and amide groups in their sample, which they suggest could be a mixture of neutral and zwitterionic forms. Significantly, their spectrum for glycine is dominated by the 288.3 eV signal. In the N 1s region, Zubavichus et al.<sup>57</sup> observe a strong signal at 401.5 eV in glycine (as well as signal at the same energy, but with increasing intensity, in the N 1s spectra of Gly, Gly–Gly, Gly–Gly–Gly, and DKP). They attribute the 401.5 eV signal in glycine and at least part of that in the other species to N 1s  $\rightarrow \sigma^*_{\text{NH}}$  transitions. This observation is in strong contradiction to the current work, where signal is not detected at this energy in solid glycine. Two possible explanations could be (1) strong polarization effects in our work, which we discount since the N 1s spectra were taken from the gel sample, where no polarization effects were detected in the C 1s region, or (2) formation of some Gly–Gly or other peptides in their sample associated with polymerization in the condensed state. We do note that the shape of the 401.5 eV signal in their N 1s spectrum of glycine is very “ $\pi^*$ ”-like (i.e., similar to that in the peptides) and quite different from the Rydberg signals observed in gaseous Gly and Gly–Gly. It is difficult to reconcile the results of that study with our results since the experimental results differ, and yet the two sample preparations appear to be similar. As Zubavichus et al.<sup>57</sup> also concluded, further studies of different solid-state preparations are required to clarify the role of the ionic form and possible polymerization before one can conclusively resolve the differences in these two studies.

**Acknowledgment.** This research is supported financially by NSERC (Canada), the Canada Research Chair program, FAPERJ, and FUJB (Brazil). We thank Dr. Ikeura-Sekiguchi for assistance with some calculations and gas-phase measurements. M.L.G. thanks the Department of Chemistry for an undergraduate summer research fellowship. Early work on the X-ray spectroscopy component of this study was carried out at the Stony Brook STXM at NSLS beamline X1A with the assistance of Chris Jacobsen and Angelika Osanna. The NEXAFS spectra reported herein were measured at the BL 532 STXM at the Advanced Light Source, which is supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. We thank David Kilcoyne and Tolek Tylicszak for their expert maintenance and development of the BL 532 facility.

## References and Notes

- Hitchcock, A. P. *J. Electron Spectrosc. Relat. Phenom.* **2000**, *112*, 9.
- Stöhr, J. *NEXAFS Spectroscopy*; Springer Tracts in Surface Science 25; Springer: Berlin, 1992.
- Kosugi, N. In *Chemical Applications of Synchrotron Radiation*; Sham, T. K., Ed.; World Scientific: River Edge, NJ, 2002; pp 228.
- Ade, H. In *Vacuum Ultraviolet Spectroscopy II*; Samson, J. A. R., Ederer, D. L., Eds.; Experimental Methods in the Physical Sciences 32; Academic Press: San Diego, 1998; p 225.
- Ade, H.; Urquhart, S. G. In *Chemical Applications of Synchrotron Radiation*; Sham, T. K., Ed.; World Scientific Publishing: River Edge, NJ, 2002; p 285.
- Hitchcock, A. P. *J. Synchrotron Radiat.* **2001**, *8*, 66.
- Meyer-Ilse, W.; Hamamoto, D.; Nair, A.; Lelièvre, S. A.; Denbeaux, G.; Johnson, L.; Pearson, A. L.; Yager, D.; Legros, M. A.; Larabell, C. A. *J. Microsc.* **2001**, *201*, 395.
- Kaznacheyev, K.; Osanna, A.; Jacobsen, C.; Plashkevych, O.; Vahtras, O.; Ågren, H.; Carravetta, V.; Hitchcock, A. P. *J. Phys. Chem. A* **2002**, *106*, 3153.
- Cannington, P. H.; Ham, N. S. *J. Electron Spectrosc. Relat. Phenom.* **1983**, *32*, 139.
- Ricker, G.; Sandorfy, C.; Nascimento, M. A. C. *J. Electron Spectrosc. Relat. Phenom.* **1984**, *34*, 327.
- Bomben, K. D.; Dev, S. B. *Anal. Chem.* **1988**, *60*, 1393.
- Hasselström, J.; Karis, O.; Weinelt, M.; Wassdahl, N.; Nilsson, A.; Nyberg, M.; Pettersson, L. G. M.; Samant, M. G.; Stöhr, J. *Surf. Sci.* **1998**, *407*, 221.
- Nyberg, M.; Hasselström, J.; Karis, O.; Wassdahl, N.; Weinelt, M.; Nilsson, A.; Pettersson, L. G. M. *J. Chem. Phys.* **2000**, *112*, 5420.
- Tanaka, M.; Nakagawa, K.; Koketsu, T.; Agui, A.; Yokoya, A. *J. Synchrotron Radiat.* **2001**, *8*, 1009.
- Boese, J.; Osanna, A.; Jacobsen, C.; Kirz, J. *J. Electron Spectrosc. Relat. Phenom.* **1997**, *85*, 9.
- Carravetta, V.; Plashkevych, O.; Ågren, H. *J. Chem. Phys.* **1998**, *109*, 1456.
- Yang, L.; Plashkevych, O.; Vahtras, O.; Carravetta, V.; Ågren, H. *J. Synchrotron Radiat.* **1999**, *6*, 708.
- Debies, T. P.; Rabalais, J. W. *J. Electron Spectrosc. Relat. Phenom.* **1974**, *3*, 315.
- Klasinc, L. *J. Electron Spectrosc. Relat. Phenom.* **1976**, *8*, 161.
- Cannington, P. H.; Ham, N. S. *J. Electron Spectrosc. Relat. Phenom.* **1979**, *15*, 79.
- Steinberger, I. T.; Teodorescu, C. M.; Gravel, D.; Flesch, R.; Wassermann, B.; Reinhardt, G.; Hutchings, C. W.; Hitchcock, A. P.; Rühl, E. *Phys. Rev. B* **1999**, *60*, 3995.
- Weiss, K.; Bagus, P. S.; Wöll, Ch. *J. Chem. Phys.* **1999**, *111*, 6834.
- Hitchcock, A. P. *Phys. Scr.* **1990**, *T31*, 159.
- Brion, C. E.; Daviel, S.; Sodhi, R. N. S.; Hitchcock, A. P. *AIP Conf. Proc.* **1982**, *94*, 429.
- Sodhi, R. N. S.; Brion, C. E. *J. Electron Spectrosc. Relat. Phenom.* **1984**, *34*, 363.
- Hitchcock, A. P.; Ishii, I. *J. Electron Spectrosc. Relat. Phenom.* **1987**, *42*, 11.
- Hitchcock, A. P.; Mancini, D. C. *J. Electron Spectrosc. Relat. Phenom.* **1994**, *67*, 1.
- Warwick, T.; Ade, H.; Kilcoyne, A. L. D.; Kritscher, M.; Tylicszak, T.; Fakra, S.; Hitchcock, A. P.; Hitchcock, P.; Padmore, H. A. *J. Synchrotron Radiat.* **2002**, *9*, 254.
- Kilcoyne, A. L. D.; Tylicszak, T.; Steele, W. F.; Fakra, S.; Hitchcock, P.; Frank, K.; Anderson, E.; Harteneck, B.; Rightor, E. G.; Mitchell, G. E.; Hitchcock, A. P.; Yang, L.; Warwick, T.; Ade, H. *J. Synchrotron Radiat.* **2003**, *10*, 125.
- Kosugi, N. *Theor. Chim. Acta* **1987**, *72*, 149.
- Kosugi, N.; Kuroda, H. *Chem. Phys. Lett.* **1980**, *74*, 490.
- Huzinaga, S.; Andzelm, J.; Klobokowski, M.; Radzio-Andzelm, E.; Sasaki, Y.; Tatewaki, H. *Gaussian Basis Sets for Molecular Calculations*; Elsevier: Amsterdam, 1984.
- Goddard, W. J.; Hunt, W. A. *Chem. Phys. Lett.* **1969**, *3*, 414.
- Ramek, M.; Cheng, V. K. W.; Frey, R. F.; Newton, S. Q.; Schaefer, L. J. *Theochem, J. Mol. Struct.* **1991**, *81*, 1.
- Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 7917.
- Yu, D.; Armstrong, D. A.; Rauk, A. *Can. J. Chem.* **1992**, *70*, 1762.
- Császár, A. G. *J. Am. Chem. Soc.* **1992**, *114*, 9568.
- Hu, C.-H.; Shen, M. Z.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1993**, *115*, 2923.
- Barone, V.; Adamo, C.; Lelj, F. *J. Chem. Phys.* **1995**, *102*, 364.
- Suenram, R. D.; Lovas, F. J. *J. Mol. Spectrosc.* **1978**, *72*, 372.
- Brown, R. D.; Godfrey, P. D.; Storey, J. W. V.; Bassez, M.-P. *J. Chem. Soc., Chem. Commun.* **1978**, 547.
- Schäfer, L.; Sellers, H. L.; Lovas, F. J.; Suenram, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 6566.
- Suenram, R. D.; Lovas, F. J. *J. Am. Chem. Soc.* **1980**, *102*, 7180.
- Godfrey, P. D.; Brown, R. D. *J. Am. Chem. Soc.* **1995**, *117*, 2019.
- Abramov, Y. A.; Volkov, A.; Wu, G.; Coppens, P. *J. Phys. Chem. B* **2000**, *104*, 2183.
- Ramachandran, G. N.; Saisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283.
- Hitchcock, A. P.; Morin, C.; Heng, Y. M.; Cornelius, R. M.; Brash, J. L. *J. Biomater. Sci. Polym. Ed.* **2002**, *13*, 919.
- Loo, B. W., Jr.; Sauerwald, I. M.; Hitchcock, A. P.; Rothman, S. S. *J. Microsc.* **2001**, *204*, 69.
- Vairavamurthy, A.; Wang, S. *Environ. Sci. Technol.* **2002**, *36*, 3050.

- (50) Ishii, I.; Hitchcock, A. P. *J. Electron Spectrosc. Relat. Phenom.* **1987**, 46, 55.
- (51) Outka, D. A.; Stöhr, J.; Madix, R. J.; Rotermund, H. H.; Hermsmeier, B.; Solomon, J. *Surf. Sci.* **1987**, 185, 53.
- (52) Jolly, W. L.; Bomben, K. D.; Eyermann, C. J. *At. Data Nucl. Data Tables* **1984**, 31, 433.
- (53) Urquhart, S. G.; Ade, H. *J. Phys. Chem. B* **2002**, 106, 8531.

- (54) Lessard, R. J.; Al-Hassan, E.; Tulumello, D.; Cooper, G.; Hitchcock, A. P. Unpublished.
- (55) Ishii, I.; Hitchcock, A. P. *J. Chem. Phys.* **1987**, 87, 830.
- (56) Ishii, I.; McLaren, R.; Hitchcock, A. P.; Robin, M. B. *J. Chem. Phys.* **1987**, 87, 4344.
- (57) Zubavichus, Y.; Zharnikov, M.; Schaporenko, A.; Grunze, M. *J. Electron Spectrosc. Relat. Phenom.* Submitted.